VIASM Vietnam Institute for Advanved Study in Mathematics

ANNUAL MEETING 2013
LECTURE NOTES

HANOI, JULY 2013

Contents

1. Congruent numbers - **John Coates** 1
2. Recent progress on the Gross-Prasad Conjecture - **Wee Tek Gan** 9
3. Parabolic flows in complex geometry - **Duong Hong Phong** 33
4. Perfectoid spaces and the weight-monodromy conjecture, after Peter Scholze - **Takeshi Saito** 53
5. The Tate Conjecture for K3 surphaces, a survey of some recent progress - **Vasudevan Srinivas** 67

CONGRUENT NUMBERS

John Coates

1 Introduction

We say that a positive integer D is *congruent* if it is the area of a right-angled triangle, all of whose sides have *rational* length. For example, the number $D = 5$ is congruent, because it is the area of the right-angled triangle, with sides of lengths $9/6, 40/6, 41/6$. The congruent number problem is simply the problem of deciding which positive integers D are congruent numbers. In fact, no algorithm has ever been proven for infallibly deciding in a finite number of steps whether a given integer D is congruent or not. We can clearly suppose that D is square free, and we shall always assume this in what follows. The origins of this problem are buried deep in antiquity, and the written record goes back at least one thousand years. It is the oldest unsolved major problem in number theory, and possibly in the whole of mathematics. The ancient literature simply wrote down examples of congruent numbers by exhibiting right-angled triangles with the desired area. It also noted certain infinite families of congruent numbers, e.g. $D = n(n^2 - 1)$ is a congruent number for all integers $n \geq 2$, as it is the area of the right-angled triangle whose sides have lengths $2n, n^2 - 1, n^2 + 1$. At some point of time, it was realised that the congruent number problem is really a question about finding non-trivial rational points on an elliptic curve, as is shown by the following elementary lemma, whose proof we omit.

Lemma 1.1 *An integer $D \geq 1$ is congruent if and only if there exists a point (x, y), with x, y in $\mathbb{Q}$ and $y \neq 0$ on the elliptic curve*

$$E^{(D)} : y^2 = x^3 - D^2 x. \tag{1}$$

Fermat was the first to prove that 1 is not a congruent number (for an account of his beautiful proof, see [1]). His highly original argument led to two further developments, which turned out to be of fundamental importance in the history of arithmetic geometry. Firstly, Fermat himself noted that his proof shows that the equation $x^4 + y^4 = 1$ has no solution in rational numbers x, y with $xy \neq 0$, and this is presumably what led him to

J. Coates
Emmanuel College, Cambridge CB2 3AP, England
and
Department of Mathematics, POSTECH, Pohang 790-784, Korea
E-mail: j.h.coates@dpmms.cam.ac.uk

his assertion that the same statement should hold when the exponent 4 is replaced by any integer $n \geq 3$. Secondly, in 1924, Mordell showed that a beautiful generalization of Fermat's argument proves that, for any elliptic curve E defined over $\mathbb{Q}$, the abelian group $E(\mathbb{Q})$ of points on E with rational coordinates is always finitely generated. Interest in the congruent number problem was further enhanced by the discovery of the conjecture of Birch and Swinnerton-Dyer, when it was quickly realised that one part of the problem, which probably had already been noted in the classical literature, is perhaps the simplest and most down to earth example of this conjecture. Recall that the complex L-series of the elliptic curve (1) is defined, in the half plane for which the real part of s is greater than 3/2, by the Euler product

$$L(E^{(D)}, s) \prod_{(q,2D)=1} (1 - a_q q^{-s} + q^{1-2s})^{-1},$$

where, for q a prime not dividing $2D$, the integer a_q is such that the number of solutions of the congruence $y^2 \equiv x^3 - D^2 x$ modulo q is equal to $q - a_q$. The analytic continuation and functional equation of this particular L-series has been known since the time of Kronecker (essentially because the elliptic curve (1) admits complex multiplication). Put $\Lambda(E^{(D)}, s) = (2\pi)^{-s}\Gamma(s)L(E^{(D)}, s)$. Then it can be shown that $\Lambda(E^{(D)}, s)$ is entire, and satisfies the functional equation

$$\Lambda(E^{(D)}, s) = w(E^{(D)})N(E^{(D)})^{1-s}\Lambda(E^{(D)}, 2 - s),$$

where $N(E^{(D)})$ denotes the conductor of $E^{(D)}$, and where the root number $w(E^{(D)})$ is equal to $+1$ if $D \equiv 1, 2, 3 \mod 8$, and is equal to -1 if $D \equiv 5, 6, 7 \mod 8$. In particular, it follows that $L(E^{(D)}, s)$ will have a zero of odd multiplicity at the point $s = 1$ if and only if $D \equiv 5, 6, 7 \mod 8$. Since it is easy to see that a rational point (x, y) on the curve (1) has finite order if and only if $y = 0$, it follows that the conjecture of Birch and Swinnerton-Dyer predicts that:-

Conjecture 1.2 Every positive integer D with $D \equiv 5, 6, 7 \mod 8$ is a congruent number.

The search for a proof of this general conjecture is unquestionably one of the major open problems of number theory. Of course, there are congruent numbers which are not in the residue classes of 5, 6, or 7 modulo 8, the smallest of which is 34 (it is the area of the right angled triangle whose sides have lengths 225/30, 272/30, 353/30).

The first important progress on the above conjecture was made by Heegner [3], in a paper which was neglected when it was initially published, but is now justly celebrated.

Theorem 1.3 (Heegner) *Let N be a square free positive integer, with precisely one odd prime factor, such that $N \equiv 5, 6, 7 \mod 8$. Then N is congruent.*

Considerable efforts were made by many number-theorists to extend this theorem to integers N with more than one prime factor, but until now nothing was established beyond the case of N with at most two odd prime factors (see [4]). However, very recently, Y. Tian [5,6] has at last found the new ideas needed to establish the desired generalization.

Theorem 1.4 (Tian) *Let p_0 be any prime number satisfying $p_0 \equiv 3, 5, 7 \mod 8$, and let M be any square free integer of the form $M = p_0 p_1 \dots p_k$, where $k \geq 1$, and $p_i \equiv 1 \mod 8$ for $1 \leq i \leq k$. Define $K_M = \mathbb{Q}(\sqrt{-2M})$, and let C_M denote the ideal class group of K_M. Assume that $2C_M \cap C_M[2]$ has order 1 or 2, according as $M \equiv 3, 5 \mod 8$ or $M \equiv 7 \mod 8$. Let $N = M$ or $2M$ be such that $N \equiv 5, 6, 7 \mod 8$. Then N is congruent, and $L(E^{(N)}, s)$ has a simple zero at $s = 1$.*

As is explained in detail at the end of [6], a well known argument then establishes the following corollary for the first time.

Corollary 1.5 (Tian) *For each integer $k \geq 1$, there exist infinitely many square free congruent numbers, with exactly k odd prime factors, in each of the residue classes $5, 6, 7\, mod\, 8$.*

Let E be the elliptic curve

$$y^2 = x^3 - x. \tag{2}$$

We remark that the field $\mathbb{Q}(E[4])$ generated by the coordinates of the 4-division points on E is in fact the field $\mathbb{Q}(\mu_8)$ given by adjoining the 8-th roots of unity to $\mathbb{Q}$. Thus a prime p will split completely in the field $\mathbb{Q}(E[4])$ precisely when $p \equiv 1\, mod\, 8$. Curiously, this fact seems to be related to the need to take the primes $p_i \equiv 1\, mod\, 8$, for $1 \leq i \leq k$, in the above theorem (see also Theorem 3.1 at the end of this lecture). Note also that the elliptic curve $E^{(N)}$ is the twist of the elliptic curve E by the quadratic extension $\mathbb{Q}(\sqrt{N})/\mathbb{Q}$. Define $E(\mathbb{Q}(\sqrt{N}))^-$ to be the subgroup of $E(\mathbb{Q}(\sqrt{N}))$ consisting of all points P such that the non-trivial element of the Galois group of $\mathbb{Q}(\sqrt{N})/\mathbb{Q}$ acts on P by -1. Then we can identify $E^{(N)}(\mathbb{Q})$ with $E(\mathbb{Q}(\sqrt{N}))^-$, and the construction of rational points on $E^{(N)}$ by both Heegner and Tian proceeds by constructing points in $E(\mathbb{Q}(\sqrt{N}))^-$.

2 Tian's induction argument

The key new idea in Tian's work is an induction argument on the number of prime factors of N, which we now briefly explain (for full details, see [6]). Let E be the elliptic curve (2), which has conductor 32. Define $\Gamma_0(32)$ to be the congruence subgroup consisting of all matrices in $SL_2(\mathbb{Z})$ with bottom left hand entry divisible by 32. The modular curve $X_0(32)$ is defined over $\mathbb{Q}$, and its complex points are given by

$$X_0(32)(\mathbb{C}) = \Gamma_0(32) \setminus (\mathfrak{H} \cup \mathbb{P}^1(\mathbb{Q})),$$

where $\mathfrak{H}$ denotes the upper half plans, and $\mathbb{P}^1(\mathbb{Q})$ the projective line over $\mathbb{Q}$. Then $X_0(32)$ has genus 1 and the cusp at infinity $[\infty]$ is a rational point (in fact, it can easily be shown that $X_0(32)$ is isomorphic over $\mathbb{Q}$ to the elliptic curve with equation $y^2 = x^3 + 4x$). There is then a degree 2 rational map defined over $\mathbb{Q}$

$$f : X_0(32) \to E, \tag{3}$$

with $f([\infty]) = O$, where O is the point at infinity on E. Let K be any imaginary quadratic field, which we always assume to be embedded in $\mathbb{C}$. If z is any point in $\mathfrak{H} \cap K$, we write P_z for the corresponding point on $X_0(32)$. The classical theory of complex multiplication, going back to the 19th century, then tells us that the point $f(P_z)$ belongs to $E(K^{ab})$, where K^{ab} denotes the maximal abelian extension of K. Moreover, it also provides an explicit description of the action of the Galois group of K^{ab} over K on this point.

We now explain Tian's induction argument in the case of certain square free N lying in the residue class of 5 modulo 8. Similar arguments (see [6]), with slightly different details, are valid for the residue classes of 6 and 7 modulo 8. Thus we assume from now on that p_0 is a prime with $p_0 \equiv 5\, mod\, 8$, and define

$$N = p_0 p_1 \dots p_k \tag{4}$$

where k is any integer ≥ 0, and $p_1, \ldots, p_k$ are distinct prime numbers satisfying $p_i \equiv 1 \, mod \, 8$ for $i = 1, \ldots, k$. We then define

$$K_N = \mathbb{Q}(\sqrt{-2N}). \tag{5}$$

Let

$$\mathfrak{z}_N = \sqrt{-2N}/8 \in \mathfrak{H} \cap K_N,$$

and write P_N for the corresponding point on $X_0(32)$. Consider the point on E defined by

$$w_N = f(P_N) + (1 + \sqrt{2}, 2 + \sqrt{2}); \tag{6}$$

here $(1+\sqrt{2}, 2+\sqrt{2})$ is a point on E of exact order 4. The reason for adding this point of order 4 is the following. Define H_N to be the Hilbert class field of K_N. Then it can easily be shown that w_N has coordinates in H_N, whereas $f(P_N)$ itself only has coordinates lying in a ramified extension of K_N. Define

$$J_N = K_N(\sqrt{N}).$$

The classical theory of genera shows that J_N is a subfield of the Hilbert class field H_N, and we than define the point u_N in $E(J_N)$ by

$$u_N = Tr_{H_N/J_N}(w_N) \tag{7}$$

where, of course, the trace map is taken on the elliptic curve E. In fact, it is easily seen using the theory of complex multiplication that

$$u_N \in E(\mathbb{Q}(\sqrt{N})). \tag{8}$$

When $k = 0$, Heegner [3] showed, just using the theory of complex multiplication, that u_N does not belong to $E(\mathbb{Q}(\sqrt{N}))^-$, whereas $2u_N$ does belong to this subgroup. As we shall now explain, by making use of L-functions, Tian [6] beautifully extended this to all $k \geq 1$, establishing first the following unconditional result. Here $E[2]$ denotes the group of points of order 2 on E. It is well known and easy to see that $E[2]$ is also the full torsion subgroup of $E^{(R)}(\mathbb{Q})$ for any square free positive integer R.

Theorem 2.1 *For all $k \geq 1$, we have*

$$u_N \in 2^{k-1}E(\mathbb{Q}(\sqrt{N}))^- + E[2]. \tag{9}$$

We now outline the main new ingredients in the proof of this theorem. Define

$$F_N = K_N(\sqrt{p_0}, \ldots, \sqrt{p_k}).$$

By the classical theory of genera, F_N is a subfield of H_N, and the Artin map defines an isomorphism from $2C_N$ to the Galois group of H_N over F_N, where, as earlier, C_N denotes the ideal class group of K_N. Define $\mathcal{D}_N$ to be the set of all those positive divisors of N, which are divisible by the prime p_0. Of course, for each $M \in \mathcal{D}_N$, we have the Heegner point u_M, defined exactly as above with N replaced by M. On the other hand, for each such $M \in \mathcal{D}_N$, we can also consider another Heegner point defined by

$$u_{N,M} = Tr_{H_N/K_N(\sqrt{M})}(w_N).$$

Once again, it can be shown that $u_{N,M}$ belongs to $E(\mathbb{Q}(\sqrt{M}))$, and plainly $u_N = u_{N,N}$. The following easily proven averaging lemma is the starting point for a proof of the above theorem by induction on k.

Lemma 2.2 . *Define* $\mathfrak{W}_N = \sum_{M \in \mathcal{D}_N} u_{N,M}$. *Then, if* $k \geq 2$, *we have* $\mathfrak{W}_N = 2^k v_N$, *and, if* $k = 1$, *we have* $\mathfrak{W}_N = 2v_N + \#(2C_N)(0,0)$, *where, in both cases,* $v_N = Tr_{H_N/F_N}(w_N)$. *Morever,* $v_N \in F_N^+ = \mathbb{Q}(\sqrt{p_0}, \dots, \sqrt{p_k})$.

We note first that, when $k = 1$ one verifies directly using the theory of complex multiplication that (9) is valid, so that the induction starts. Now let us suppose $k > 1$, and make the inductive hypothesis that, for all $M \in \mathcal{D}_N$ with $M \neq N$, we have

$$u_M \in 2^{k(M)-1}E(\mathbb{Q}(\sqrt{M}))^- + E[2], \tag{10}$$

where $k(M)$ now denotes the number of prime factors of M (in the special case when $k(M) = 0$, i.e. when $M = p_0$, this is understood to mean that $2u_M$ lies in $E(\mathbb{Q}(\sqrt{M}))^- + E[2]$, and this is easily verified to be true). So far, complex L-functions have not been used anywhere in our argument, and Tian's marvellous idea is to now exploit them to deduce information about the Heegner points $u_{N,M}$ from (10). We write χ_M for the non-trivial character of the the quadratic extension $K_N(\sqrt{M})/K_N$ (this extension is non-trivial because M is always divisible by the prime p_0). Let $L(E/K_N, \chi_M, s)$ be the complex L-series of E over K, twisted by the unramified abelian character χ_M of K. A simple argument with the properties of L-function under induction, applied to the quartic extension $K_N(\sqrt{M})/\mathbb{Q}$, then establishes the following lemma.

Lemma 2.3 *For each* $M \in D_N$, *we have* $L(E/K_N, \chi_M, s) = L(E^{(2N/M)}, s)L(E^{(M)}, s)$.

Combining this lemma with the generalized Gross-Zagier formula of Zhang proven in [7] (which is needed, rather than the classical Gross-Zagier formula, because the discriminant of K_N has the factor 2 in common with the conductor of E), Tian then establishes the following result, in which $\widehat{h}$ denotes the canonical Neron-Tate height function on $E(\bar{\mathbb{Q}})$. For each positive square free integer R, define

$$L^{(alg)}(E^{(R)}, 1) = L(E^{(R)}, 1)\sqrt{R}/\Omega,$$

where $\Omega = 2.62206\dots$ is the least positive real period of the Neron differential on E. It is well known that $L^{(alg)}(E^{(R)}, 1)$ is a rational number. Moreover, $L^{(alg)}(E^{(2)}, 1) = 1/2$.

Theorem 2.4 *Assume that* $M \in D_N$, *and that the Heegner point* u_M *is not torsion. Then* $E^{(M)}(\mathbb{Q})$ *has rank 1, and we have*

$$\widehat{h}(u_{N,M})/\widehat{h}(u_M) = L^{(alg)}(E^{(2N/M)}, 1)/L^{(alg)}(E^{(2)}, 1). \tag{11}$$

It can perfectly well happen that $L^{(alg)}(E^{(2N/M)}, 1) = 0$ (for example, when $N/M = 17$), but in this case we conclude from the above theorem that $u_{N,M}$ is itself a torsion point in $E^{(M)}(\mathbb{Q})$, and so belongs to $E[2]$ by the remark made earlier. On the other hand, if $u_{N,M}$ is not torsion, the Gross-Zagier theorem tells us that $L(E/K_N, \chi_M, s)$ has a simple zero at $s = 1$, and so we conclude from Lemma 2.3 that $L(E^{(M)}, s)$ also has a simple zero at $s = 1$ and thus again invoking the Gross-Zagier theorem, it follows that u_M is not torsion. We are therefore in a position where we can use (11) to estimate the height of $u_{N,M}$. To do this, we invoke the following theorem of Zhao [8].

Theorem 2.5 *Let* R *be a square free positive integer, which is a product of* $r \geq 1$ *primes, all of which are* $\equiv 1 \, mod \, 8$. *Then, if* $L(E^{(2R)}, 1) \neq 0$, *we have*

$$ord_2(L^{(alg)}(E^{(2R)}, 1) \geq 2r.$$

When $M = p_0$, or equivalently $N/M = p_1...p_k$, Heegner was the first to point out that the theory of complex multiplication shows that $2u_M$ belongs to $E(\mathbb{Q}(\sqrt{M})^-$, whence, in this case, we deduce immediately from Zhao's theorem and (11) that

$$u_{M,N} \in 2^k E(\mathbb{Q}(\sqrt{M}))^- + E[2]. \tag{12}$$

Now suppose that $k(M) \geq 1$ but $M \neq N$. If $u_{N,M}$ is torsion, then it must belong to $E[2]$, and then (12) is clearly valid. If $u_{N,M}$ is not torsion, then, as remarked above, u_M is also not torsion and $E^{(M)}(\mathbb{Q})$ has rank 1. We can therefore use our induction hypothesis (10), combined with (11) and Zhao's theorem to conclude that (12) is again valid. Thus we can write $u_{M,N} = 2^k z_M + t_M$, with $z_M \in E(\mathbb{Q}(\sqrt{M})^-$ and $t_M \in E[2]$ for all $M \in \mathcal{D}_N$ with $M \neq N$. Hence, recalling that $u_N = u_{N,N}$, we conclude from Lemma 2.2 that

$$u_N = 2^k(v_N - \sum_{M \in \mathcal{D}_N, M \neq N} z_M) + t,$$

for some $t \in E[2]$. In particular, it follows that $2u_N$ belongs to $2^{k+1}E(F_N^+)$, and so the class of $2u_N$ lies in the kernel of the natural map from $E(\mathbb{Q}(\sqrt{N}))/2^{k+1}E(\mathbb{Q}(\sqrt{N}))$ to $E(F_N^+)/2^{k+1}E(F_N^+)$. But, by Kummer theory on the curve, this kernel is contained in $H^1(Gal(E(F_N^+/\mathbb{Q}(\sqrt{N})), E[2^{n+1}](F_N^+))$. This cohomology group is killed by 2, because any proper sub-extension of F_N^+ must be ramified at an odd prime, whereas only the prime 2 can ramify in the field $\mathbb{Q}(E[2^\infty])$, so that

$$E[2^{n+1}](F_N^+) = E[2]. \tag{13}$$

Therefore, we must have $4u_N \in 2^{k+1}E(\mathbb{Q}(\sqrt{N}))$. One deduces easily that $u_N = 2^{k-1}r_N \, mod \, E[2]$ for some $r_N \in E(\mathbb{Q}(\sqrt{N}))$, and that

$$r_N = 2v_N - \sum_{M \in \mathcal{D}_N, M \neq N} 2z_M + t_1 \tag{14}$$

for some $t_1 \in E[2]$. Let τ denote the Artin symbol of the unique prime of K_N above 2 for the extension H_N/K_N. It is easily seen that $\tau(\sqrt{M}) = -\sqrt{M}$ for all $M \in \mathcal{D}_N$. Hence $\tau(z_M) = -z_M$ for all $M \neq N$, and so (14) implies that

$$\tau(r_N) + r_N = 2(\tau(v_N) + v_N).$$

However, it is easily seen that

$$\tau(v_N) + v_N = \#(2C_N)(0,0), \tag{15}$$

and so $2(\tau(v_N) + v_N) = 0$. Hence r_N belongs to $E(\mathbb{Q}(\sqrt{N})^-$, and and so we have finally proven the assertion (9) by induction on k. This completes the proof of Theorem 2.1.

Tian's theorem in the Introduction for the residue class of 5 modulo 8 now follows by combining Theorem 2.1 with the following result.

Theorem 2.6 *If the ideal class group C_N of K_N has no element of order 4, then*

$$u_N \notin 2^k E(\mathbb{Q}(\sqrt{N}))^- + E[2].$$

Proof Suppose on the contrary that $u_N = 2^k z_N \, mod \, E[2]$ for some $z_N \in E(\mathbb{Q}(\sqrt{N}))^-$. It then follows from Lemma 2.2 that

$$2^k\theta \in E[2], \text{with } \theta = v_N - \sum_{M \in \mathcal{D}_N} z_M.$$

Since $\theta \in F_N^+$, it again follows from (13) that $\theta \in E[2]$, and so one sees that $\tau(v_N) + v_N = 0$. But, if C_N has no element of order 4, then $\#(2C_N)$ is odd, and so this last equation contradicts (15). This completes the proof. □

3 Generalization

It is natural to try and establish some analogue of these results for any elliptic curve E over $\mathbb{Q}$, in which one seeks to prove the existence of a suitably large infinite set $\Sigma(E)$ of square free integers R, which could be either positive or negative, such that the complex L-series $L(E^{(R)}, s)$ has a simple zero at $s = 1$ for all $R \in \Sigma(E)$. Here $L(E^{(R)}, s)$ now denotes the complex L-series of E, twisted by the quadratic extension $\mathbb{Q}(\sqrt{R})/\mathbb{Q}$. A first step in this direction is carried out in [2], where a similar induction argument to the one outlined above establishes the following result. Let E be the elliptic curve

$$y^2 + xy = x^3 - x^2 - 2x - 1.$$

It has complex multiplication by the ring of integers of the field $\mathbb{Q}(\sqrt{-7})$, and is well known to be isomorphic to the modular curve $X_0(49)$.

Theorem 3.1 *Let $N = p_0p_1...p_k$ be a square free positive integer, where $p_0 \neq 7$ is any prime which is $\equiv 3 \, mod \, 4$ and is a quadratic non-residue modulo 7, and where $p_1, \ldots, p_k$ are distinct primes which split completely in the field $\mathbb{Q}(E[4])$. Assume that the ideal class group of the field $\mathbb{Q}(\sqrt{-N})$ has no element of order 4. Then the complex L-series $L(E^{(-N)}, s)$ has a simple zero at $s = 1$, and consequently $E^{(-N)}(\mathbb{Q})$ has rank 1, and the Tate-Shafarevich group of $E^{(-N)}$ is finite.*

The field $\mathbb{Q}(E[4])$ for this curve is given explicitly by $\mathbb{Q}(\mu_4, \sqrt[4]{-7})$, where μ_4 denotes the group of 4-th roots of unity. Of course, by the Chebotarev theorem, there is a positive density of rational primes which split completely in $\mathbb{Q}(E[4])$, and the ones smaller than 1000 are

$$113, 149, 193, 197, 277, 317, 373, 421, 449, 457, 541, 557, 809, 821, 953.$$

We believe that, under the conditions of the theorem, the Tate-Shafarevich group of $E^{(-N)}$ should also have odd order, but we still have not proven it.

References

1. Coates, J.: The congruent number problem. Q. J. Pure Applied Math. **1**, 14-27 (2005)
2. Coates, J., Li, Y., Tian, Y., Zhai, S.: Quadratic twists of $X_0(49)$ with analytic rank 1. To appear
3. Heegner, K.: Diophantische analysis und modulfunktionen. Math. Z. **56**, 227-253 (1952)
4. Monsky, P.: Mock Heegner points and congruent numbers. Math. Z. **204**, 45-68 (1990)
5. Tian, Y.: Congruent numbers with many prime factors. Proc. Natl. Acad. Sc. USA **109**, 21256-21258 (2012)
6. Tian, Y.: Congruent numbers and Heegner points. To appear
7. Yuan, X., Zhang, S., Zhang, W.: The Gross-Zagier formula of Shimura curves. Annals of Mathematics Studies **184** (2012)
8. Zhao, C.: A criterion for elliptic curves with lowest 2-power order in $L(1)$ (II). Proc. Cambridge Phil. Soc. **134**, 407-420 (2003)

RECENT PROGRESS ON THE GROSS-PRASAD CONJECTURE

Wee Teck Gan

1 Introduction

The Gross-Prasad conjecture concerns a restriction or branching problem in the representation theory of real or p-adic Lie groups. It also has a global counterpart which is concerned with a family of period integrals of automorphic forms. The conjecture itself was proposed by Gross and Prasad in the context of special orthogonal groups in 2 papers [GP1,GP2] some twenty years ago. In a more recent paper [GGP1], the conjecture was extended to all classical groups, i.e. orthogonal, unitary, symplectic and metaplectic groups. Though the conjecture has the same flavour for these different groups, each of the cases have their own peculiarities which make a uniform exposition somewhat difficult. As such, for the purpose of this expository article, we shall focus only on the case of unitary groups.

A motivating example is the following classical branching problem in the theory of compact Lie groups. Let π be an irreducible finite dimensional representation of the compact unitary group $\mathrm{U}(n)$, and consider its restriction to the naturally embedded subgroup $\mathrm{U}(n-1)$. It is known that this restriction is multiplicity-free, but one may ask precisely which irreducible representations of $\mathrm{U}(n-1)$ occur in the restriction.

To give an answer to this question, we need to have names for the irreducible representations of $\mathrm{U}(n-1)$, so that we can say something like: "this one occurs but that one doesn't". Thus, it is useful to have a classification of the irreducible representations of $\mathrm{U}(n)$. Such a classification is provided by the Cartan-Weyl theory of highest weight, according to which the irreducible representations of $\mathrm{U}(n)$ are determined by their "highest weights" which are in natural bijection with sequences of integers

$$\underline{a} = (a_1 \leq a_2 \leq \cdots \leq a_n).$$

Now suppose that π has highest weight $\underline{a}$. Then a beautiful classical theorem says that: an irreducible representation τ of $\mathrm{U}(n-1)$ with highest weight $\underline{b}$ occurs in the restriction of π if and only if $\underline{a}$ and $\underline{b}$ are interlacing:

$$a_1 \leq b_1 \leq a_2 \leq b_2 \leq \cdots \leq b_{n-1} \leq a_n.$$

W.T. Gan
Department of Mathematics, National University of Singapore, 10 Lower Kent Ridge Road Singapore 119076
E-mail: matgwt@math.nus.edu.sg

The Gross-Prasad conjecture considers the analogous restriction problem for the non-compact Lie groups $\mathrm{U}(p,q)$ and their p-adic analogs. As an example, consider the case when $n=2$, where we have seen that the representation $\pi_{\underline{a}}$ of $\mathrm{U}(2)$ contains $\tau_{\underline{b}}$ precisely when $a_1 \leq b_1 \leq a_2$. Consider instead the non-compact $\mathrm{U}(1,1)$ which is closely related to the group $\mathrm{SL}_2(\mathbb{R})$. Indeed, one has an isomorphism of real Lie groups

$$\mathrm{U}(1,1) \cong (\mathrm{SL}_2(\mathbb{R})^{\pm} \times S^1)/\Delta\mu_2.$$

Now let π be an irreducible representation of $\mathrm{U}(1,1)$ (in an appropriate category); note that π is typically infinite-dimensional but since $\mathrm{U}(1)$ is compact, the restriction of π to $\mathrm{U}(1)$ is a direct sum of irreducible characters of $\mathrm{U}(1)$. It is known that this decomposition is multiplicity-free, so one is interested in determining precisely which characters of $\mathrm{U}(1)$ occur.

For this, it is again useful to have a classification of the irreducible representations of $\mathrm{U}(1,1)$. Such a classification has been known for a long time, and was the beginning of the systematic investigation of the infinite-dimensional representation theory of general reductive Lie groups, culminating in the work of Harish-Chandra, especially his construction and classification of the discrete series representations. These discrete series representations are the most fundamental representations, in the sense that every other irreducible representation can be built from them by a systematic procedure (parabolic induction and taking quotients).

For $\mathrm{U}(1,1)$, it turns out that the discrete series representations are classified by a pair of integers $\underline{a} = (a_1 \leq a_2)$. Then one can show that a irreducible representation $\tau_{\underline{b}}$ of $\mathrm{U}(1)$ occurs in the restriction of $\pi'_{\underline{a}}$ if and only if

$$b_1 \notin [a_1, a_2],$$

i.e. if and only if $\underline{a}$ and $\underline{b}$ do not interlace!

Let us draw some lessons from this simple example:

(a) to address the branching problem, it is useful, even necessary, to have some classification of the irreducible representations of a real or p-adic Lie group. A conjectural classification exists and is called the local Langlands conjecture.

(b) it is useful to group certain representations of different but closely related groups together. In the example above, we see that if one groups together the representations $\pi_{\underline{a}}$ of $\mathrm{U}(2)$ and $\pi'_{\underline{a}}$ of $\mathrm{U}(1,1)$, then the branching problem has a nice uniform answer:

$$\dim \mathrm{Hom}_{\mathrm{U}(1)}(\pi_{\underline{a}}, \tau_{\underline{b}}) + \dim \mathrm{Hom}_{\mathrm{U}(1)}(\pi'_{\underline{a}}, \tau_{\underline{b}}) = 1$$

for any $\underline{a}$ and $\underline{b}$. That the local Langlands conjecture can be expanded to allow for such a classification was first suggested by Vogan.

(c) there is a simple recipe for deciding which of the two spaces $\mathrm{Hom}_{\mathrm{U}(1)}(\pi_{\underline{a}}, \tau_{\underline{b}})$ or $\mathrm{Hom}_{\mathrm{U}(1)}(\pi'_{\underline{a}}, \tau_{\underline{b}})$ is nonzero, given by the interlacing or non-interlacing condition. In the general case, we would like a similar such recipe. However, it will turn out that this is a delicate issue and formulating the precise condition is the most subtle part of the Gross-Prasad conjecture.

Let us give a summary of the paper. In Section 2, we formulate the branching problem precisely, and recall some basic results, such as the multiplicity-freeness of the restriction. This multiplicity-freeness result is proved only surprisingly recently, by the work of Aizenbud-Gouretvitch-Rallis-Schiffman [AGRS], Waldspurger [W5], Sun-Zhu

[SZ] and Sun [S] . We shall introduce the local Langlands conjecture, in its refined form due to Vogan [V], in Section 3, where we use Vogan's notion of "pure inner forms". Then we shall state the local and global GP conjecture in Section 4. In the global setting, we also mention a refinement due to Ichino-Ikeda [II]. Finally, we describe some recent progress on the GP conjecture in the remaining sections, highlighting the work of Waldspurger [W1-4] and Beuzart-Plesis [BP] in the local case and the work of Jacquet-Rallis [JR], Wei Zhang [Zh1, Zh2], Yifeng Liu [L] and Hang Xue [X1, X2] in the global case. We conclude with listing some outstanding problems in this story in the last section.

2 The Restriction Problem

Let k be a field, not of characteristic 2. Let σ be a non-trivial involution of k having k_0 as the fixed field. Thus, k is a quadratic extension of k_0 and σ is the nontrivial element in the Galois group $\mathrm{Gal}(k/k_0)$. Let ω_{k/k_0} be the quadratic character of $k_0^\times$ associated to k/k_0 by local class field theory.

2.1 The spaces

Let V be a finite dimensional vector space over k. Let

$$\langle -,- \rangle : V \times V \to k$$

be a non-degenerate, σ-sesquilinear form on V, which is ϵ-symmetric (for $\epsilon = \pm 1$ in $k^\times$):

$$\begin{aligned} \langle \alpha v + \beta w, u \rangle &= \alpha \langle v, u \rangle + \beta \langle w, u \rangle \\ \langle u, v \rangle &= \epsilon \cdot \langle v, u \rangle^\sigma . \end{aligned}$$

2.2 The groups

Let $G(V) \subset \mathrm{GL}(V)$ be the algebraic subgroup of elements T in $\mathrm{GL}(V)$ which preserve the form $\langle -,- \rangle$:

$$\langle Tv, Tw \rangle = \langle v, w \rangle .$$

Then $G(V)$ is a unitary group, defined over the field k_0.

If one takes k to be the quadratic algebra $k_0 \times k_0$ with involution $\sigma(x,y) = (y,x)$ and V a free k-module, then a non-degenerate form $\langle -,- \rangle$ identifies the $k = k_0 \times k_0$ module V with the sum $V_0 + V_0^\vee$, where V_0 is a finite dimensional vector space over k_0 and $V_0^\vee$ is its dual. In this case $G(V)$ is isomorphic to the general linear group $\mathrm{GL}(V_0)$ over k_0. In our ensuing discussion, this split case can be handled concurrently, and is necessary for the global case.

2.3 Pairs of spaces

Now suppose that $W \subset V$ is a nondegenerate subspace, with $V = W \oplus W^{\perp}$, satisfying:

$$\dim W^{\perp} = \begin{cases} 1 \text{ if } \epsilon = 1; \\ 0 \text{ if } \epsilon = -1. \end{cases}$$

One thus has a natural embedding $G(W) \hookrightarrow G(V)$ with $G(W)$ acting trivially on $W^{\perp}$. We set

$$G = G(V) \times G(W) \quad \text{and} \quad H = \Delta G(W) \subset G.$$

2.4 Pure inner forms

We shall assume henceforth that k is a local field of characteristic 0. A pure inner form of $G(V)$ is a form of $G(V)$ constructed from an element of the Galois cohomology set $H^1(k_0, G(V))$. For the case at hand, the pure inner forms are easily described and are given by the groups $G(V')$ as V' ranges over similar type of spaces as V with $\dim V' = \dim V$.

More concretely, when k is p-adic, there are two Hermitian or skew-Hermitian spaces of a given dimension, so that $G(V)$ has a unique pure inner form $G(V')$ (other than itself). When $\dim V$ is odd, the groups $G(V)$ and $G(V')$ are quasi-split isomorphic even though the spaces V and V' are not. When $\dim V$ is even, we take the convention that $G(V)$ is quasi-split whereas $G(V')$ is not.

When $k = \mathbb{C}$ and $k_0 = \mathbb{R}$, the pure inner forms of $G(V)$ are precisely the groups $\mathrm{U}(p,q)$ with $p + q = \dim_k V$.

Now given a pair of spaces $W \subset V$, we have the notion of *relevant pure inner forms.* A pair $W' \subset V'$ is a relevant pure inner form of $W \subset V$ if W' and V' are pure inner forms of W and V respectively, and $V/W \cong V'/W'$. Then, in the p-adic case, $W \subset V$ has a unique relevant pure inner form (other than itself).

2.5 The restriction problem

Now we can state the restriction problem. Let $\pi = \pi_1 \boxtimes \pi_2$ be an irreducible smooth representation of $G(k_0) = G(V) \times G(W)$. When $\epsilon = 1$, we are interested in determining

$$\mathrm{Hom}_{H(k_0)}(\pi, \mathbb{C}) \cong \mathrm{Hom}_{G(W)}(\pi_1, \pi_2^{\vee}). \tag{2.1}$$

We shall call this the *Bessel case* of the local GP conjecture.

When $\epsilon = -1$, one needs an extra data to state the restriction problem. Since W is skew-Hermitian, the space $\mathrm{Res}_{k/k_0}(W)$ inherits the structure of a symplectic space, so that

$$\mathrm{U}(W) \subset \mathrm{Sp}(\mathrm{Res}_{k/k_0}(W)).$$

The metaplectic group $\mathrm{Mp}(\mathrm{Res}_{k/k_0}(W))$ (which is an S^1-extension of $\mathrm{Sp}(\mathrm{Res}_{k/k_0}(W))$) has a Weil representation ω_{ψ_0} associated to a nontrivial additive character ψ_0 of k_0. It is known that the metaplectic veering splits over the subgroup $\mathrm{U}(W)$ but the splitting is not unique since $\mathrm{U}(W)$ has nontrivial unitary characters. However, a splitting can be specified by a pair (ψ_0, χ) where χ is a character of $k^\times$ such that $\chi|_{k_0^\times} = \omega_{k/k_0}$. For such a splitting $i_{W,\psi_0,\chi}$, we obtain a representation

$$\omega_{W,\psi_0,\chi} := \omega_{\psi_0} \circ i_{W,\psi_0,\chi}$$

of $\mathrm{U}(W)$; we call this a Weil representation of $\mathrm{U}(W)$.

Then one is interested in determining

$$\mathrm{Hom}_{H(k_0)}(\pi, \omega_{W,\psi_0,\chi}). \tag{2.2}$$

We call this the *Fourier-Jacobi* case of the local GP conjecture. To unify notations in the two cases, we set

$$\nu = \nu_{\psi_0,\chi} = \begin{cases} \mathbb{C} \text{ if } \epsilon = 1; \\ \omega_{W,\psi_0,\chi} \text{ if } \epsilon = -1. \end{cases}$$

We note that [GGP1] considers pairs of spaces $W \subset V$ with arbitrary $\dim W^\perp$, and formulates a restriction problem in this general setting, whereas we have restricted ourselves to the case of $\dim W^\perp \leq 1$ in this article for simplicity.

2.6 Multiplicity-freeness

In a number of recent papers, beginning with [AGRS] and followed by [SZ] and [S], the following fundamental theorem was shown:

Theorem 2.3 *The space* $\mathrm{Hom}_{H(k_0)}(\pi, \nu)$ *is at most one-dimensional.*

Thus, the remaining question is whether this Hom space is 0 or 1-dimensional.

The case when $k = k_0 \times k_0$ is particularly simple. One has:

Proposition 2.4 *When $k = k_0 \times k_0$, the above Hom space is 1 -dimensional when π is generic, i.e. has a Whittaker model.*

The local GP conjecture gives a precise criterion for the Hom spaces above to be nonzero. However, to state the precise criterion requires substantial preparation and groundwork.

2.7 Periods

We now consider the global situation, where F is a number field with ring of addles $\mathbb{A}$ and E/F is a quadratic field extension. Hence the spaces $W \subset V$ are Hermitian or skew-Hermitian spaces over E and the associated groups G and H are defined over F. Let $\mathcal{A}_{cusp}(G)$ denote the space of cuspidal automorphic forms of $G(\mathbb{A})$. When $\epsilon = 1$, there is a natural $H(\mathbb{A})$-invariant linear functional on $\mathcal{A}_{cusp}(G)$ defined by

$$\mathcal{P}_H(f) = \int_{H(F)\backslash H(\mathbb{A})} f(h) \cdot dh.$$

This map is called the H-period integral.

When $\epsilon = -1$, the Weil representation $\omega_{W,\psi_0,\chi}$ admits an automorphic realization via the formation of theta series. Then one considers

$$\mathcal{P}_H : \mathcal{A}(G) \boxtimes \overline{\omega_{W,\psi_0,\chi}} \longrightarrow \mathbb{C}$$

given by

$$\mathcal{P}_H(f \otimes \overline{\theta_\varphi}) = \int_{H(F)\backslash H(\mathbb{A})} f(h) \cdot \overline{\theta_\varphi(h)}\, dh.$$

Now let

$$\pi = \pi_1 \boxtimes \pi_2 \subset \mathcal{A}_{cusp}(G)$$

be a cuspidal representation of $G(\mathbb{A})$. Then the restriction of $\mathcal{P}_H$ to π defines an element in

$$\mathrm{Hom}_{H(\mathbb{A})}(\pi \otimes \overline{\nu}, \mathbb{C}).$$

By Theorem 2.3, one knows that these adelic Hom spaces have dimension at most 1, and that the dimension is 1 precisely when the relevant local Hom spaces are nonzero for all places v of F. Moreover, it is clear that the nonvanishing of these Hom spaces is necessary for the nonvanishing of $\mathcal{P}_H$.

The global GP conjecture gives a precise criterion for the nonvanishing of the globally-defined linear functional $\mathcal{P}_H$.

3 Local Langlands Correspondence

To understand the restriction problem described in the previous section, it will be useful to have a classification of the irreducible representations of $G(k_0)$ in the local case, and a classification of the cuspidal representations of $G(\mathbb{A})$ in the global case. The Langlands program provides such a classification, known as the (local or global) Langlands correspondence. On one hand, the Langlands correspondence can be viewed as a generalisation of the Cartan-Weyl theory of highest weights which classifies the irreducible representations of a connected compact Lie group. On the other hand, it can be considered as a profound generalisation of class field theory, which classifies the abelian extensions of a local or number field. In this section, we briefly review the salient features of the Langlands correspondence.

3.1 Weil-Deligne group

We first introduce the parametrizing set. For a local field k, let W_k denote the Weil group of k. When k is a p-adic field, one has a commutative diagram of short exact sequences:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & I_k & \longrightarrow & \mathrm{Gal}(\overline{k}/k) & \longrightarrow & \widehat{\mathbb{Z}} & \longrightarrow & 1 \\ & & \| & & \uparrow & & \uparrow & & \\ 1 & \longrightarrow & I_k & \longrightarrow & W_k & \longrightarrow & \mathbb{Z} & \longrightarrow & 1 \end{array}$$

where I_k is the inertia group of $\mathrm{Gal}(\overline{k}/k)$, and $\widehat{\mathbb{Z}}$ is the absolute Galois group of the residue field of k, equipped with a canonical generator (the geometric Frobenius element Frob_k). This exhibits the Weil group W_k as a dense subgroup of the absolute Galois group of k. When k is archimedean, we have

$$W_k = \begin{cases} \mathbb{C}^\times \text{ if } k = \mathbb{C}; \\ \mathbb{C}^\times \cup \mathbb{C}^\times \cdot j, \text{ if } k = \mathbb{R}, \end{cases}$$

where $j^2 = -1 \in \mathbb{C}^\times$ and $j \cdot z \cdot j^{-1} = \overline{z}$ for $z \in \mathbb{C}^\times$. Set the Weil-Deligne group to be

$$WD_k = \begin{cases} W_k \text{ if } k \text{ is archimedean;} \\ W_k \times \mathrm{SL}_2(\mathbb{C}), \text{ if } k \text{ is p-adic.} \end{cases}$$

3.2 L-parameters

Now let ${}^L G(V)$ denote the L-group of $G(V)$ over k_0, so that

$${}^L G(V) = \mathrm{GL}_n(\mathbb{C}) \rtimes \mathrm{Gal}(k/k_0),$$

with $n = \dim_k V$ and σ acting on $\mathrm{GL}_n(\mathbb{C})$ as a pinned outer automorphism. By an L-parameter (or Langlands parameter) of $G(V)$, we mean a $\mathrm{GL}_n(\mathbb{C})$-conjugacy class of admissible homomorphisms

$$\phi : WD_{k_0} \longrightarrow {}^L G(V)$$

such that the composite of ϕ with the projection onto $\mathrm{Gal}(k/k_0)$ is the natural projection of WD_{k_0} to $\mathrm{Gal}(k/k_0)$.

The need to work with the semi-direct product $\mathrm{GL}_n(\mathbb{C}) \rtimes \mathrm{Gal}(k/k_0)$ is quite a nuisance, but the following useful result was shown in [GGP1]:

Proposition 3.1 *Restriction to W_k defines a bijection between the set of L-parameters for $G(V)$ and the set of equivalence classes of Frobenius semisimple, conjugate-self-dual representations*

$$\phi : WD_k \longrightarrow \mathrm{GL}_n(\mathbb{C})$$

of sign $(-1)^{n-1}$.

This means that L-parameters for unitary groups are essentially local Galois representations with some conjugate-self-duality property.

3.3 Component groups

A invariant that one can attach to an L-parameter ϕ is its component group

$$S_\phi = \pi_0(Z_{\mathrm{GL}_n(\mathbb{C})}(\phi)^{\mathrm{Gal}(k/k_0)})$$

where we regard ϕ as a map $WD_{k_0} \longrightarrow {}^L G(V)$ here. Thus, S_ϕ is a finite group which can be described more explicitly as follows. Regarding ϕ now as a representation of WD_k, let us decompose ϕ into its irreducible components:

$$\phi = \bigoplus_i n_i \cdot \phi_i.$$

Then S_ϕ is an elementary abelian 2-group, i.e. a vector space over $\mathbb{Z}/2\mathbb{Z}$, which is equipped with a canonical basis:

$$S_\phi = \prod_i \mathbb{Z}/2\mathbb{Z} \cdot e_i$$

where the product runs over all indices i such that ϕ_i is conjugate-self-dual of sign $(-1)^{n-1}$ (i.e. of same sign as ϕ).

3.4 Local Langlands conjecture

We can now formulate the local Langlands conjecture for the groups $G(V)$:

Local Langlands Conjecture (LLC)

There is a natural bijection

$$\bigsqcup_{V'} \mathrm{Irr}(G(V')) \longleftrightarrow \Phi(G(V))$$

where the union on the LHS runs over all pure inner forms V' of V and the set $\Phi(G(V))$ is the set of isomorphism classes of pairs (ϕ, η) where ϕ is an L-parameter of $G(V)$ and $\eta \in \mathrm{Irr}(S_\phi)$.

Given an L-parameter ϕ for $G(V)$, we let Π_ϕ be the finite set of irreducible representations of $G(V')$, with V' running over all pure inner forms of V, which corresponds to (ϕ, η) for some η. This is called the L-packet with L-parameter ϕ. So

$$\bigsqcup_{V'} \mathrm{Irr}(G(V')) = \bigsqcup_{\phi} \Pi_\phi.$$

and an irreducible representation π of $G(V)$ (or its pure inner form) is of the form

$$\pi = \pi(\phi, \eta)$$

for a unique pair (ϕ, η) as above. We shall frequently write $\pi = \pi(\eta)$ if ϕ is fixed and understood.

3.5 Status

The LLC has been established for the group $\mathrm{GL}(n)$ by Harris-Taylor [HT] and Henniart [He2]. For the unitary groups $G(V)$, the LLC was established in the recent paper of Mok [M] when $G(V)$ is quasi-split, following closely the book of Arthur [A] where the symplectic and orthogonal groups were treated. The results of [A] and [M] are at the moment conditional on the stabilisation of the twisted trace formula, but substantial efforts are currently being made towards this stabilisation, and one can be optimistic that in the coming months, the results will be unconditional. With the stabilisation at hand, one can also expect that the results for non-quasi-split unitary groups will also follow.

For the purpose of this article, we shall assume that the LLC has been established.

3.6 L-factors and ϵ-factors

Given an L-parameter ϕ of $G(V)$, one can associate some arithmetic invariants. More precisely, if

$$\rho : {}^L G(V) \longrightarrow \mathrm{GL}(U)$$

is a complex representation, then we may form the local Artin L-factor over k_0:

$$L_{k_0}(s, \rho \circ \phi) = \frac{1}{\det(1 - q^{-s}(\rho \circ \phi)(Frob_{k_0})|U^{I_{k_0}})},$$

where q is the cardinality of the residue field of k_0.

Similarly, if we are given

$$\rho : \mathrm{GL}_n(\mathbb{C}) \longrightarrow \mathrm{GL}(U),$$

then we have a composite $\rho \circ \phi : WD_k \longrightarrow \mathrm{GL}(U)$ and we can form the analogous local Artin L-factor $L_k(s, \rho \circ \phi)$ over k; here we have identified ϕ with its restriction to WD_k.

Further, one can associate a local epsilon factor $\epsilon(s, \rho \circ \phi, \psi)$ which is a nowhere zero entire function of s depending on ϕ, ρ and an additive character ψ of k. This local epsilon factor is quite a subtle invariant, which satisfies a list of properties. While it is not hard to show that this list of properties characterize the local epsilon factor, the issue of existence is not trivial at all. Indeed, the existence of this invariant is due independently to Deligne [De] and Langlands.

We shall mention only one key property of the local epsilon factors that we need. If $\rho \circ \phi$ is a conjugate symplectic representation of W_k, and ψ is a nontrivial additive character of k/k_0, then $\epsilon(1/2, \rho \circ \phi, \psi) = \pm 1$. Moreover, this sign depends only on the $Nk^{\times}$-orbit of ψ. Indeed, if $\dim \rho \circ \phi$ is even, then this sign is independent of the choice of ψ.

This sign will play an important role in the local GP conjecture.

3.7 Characterization by Whittaker datum

Perhaps some explanation is needed for the meaning of the adjective "natural" in the statement of the LLC. In what sense is the bijection in the LLC natural?

One possibility is that one could characterise the bijection postulated in the LLC by requiring that it preserves certain natural invariants that one can attach to both sides. This is the case for $\mathrm{GL}(n)$ where the local L-factors and local ϵ-factors of pairs are used to characterise the correspondence; such a characterisation is due to Henniart [He1].

For the unitary groups $G(V)$, the proof of the LLC given in [M] characterises the LLC in a different way: via a family of character identities arising in the theory of endoscopy. This elaborate theory requires one to normalize certain "transfer factors". By the work of Kottwitz-Shelstad [KS] and the recent work of Kaletha [K], one can fix a normalisation of the transfer factors by fixing a "Whittaker datum" for $G(V)$. Let us explain briefly what this means.

The group $G(V)$ being quasi-split, one can choose a Borel subgroup $B = T \cdot U$ defined over k_0, with unipotent radical U. A Whittaker datum on $G(V)$ is a character $\chi : U(k_0) \longrightarrow S^1$ which is in general position, i.e. whose $T(k_0)$-orbit is open, and two such characters are equivalent if they are in the same $T(k_0)$-orbit.

If $\dim V$ is odd, then any two generic characters of $U(k_0)$ are equivalent, so the LLC for $G(V)$ is quite canonical. On the other hand, when $\dim V$ is even, there are two equivalence classes of Whittaker data. In this case, we have:

Lemma 3.2 *Using the form $\langle -, - \rangle$ on V, one gets a natural identification*

$$\text{Whittaker datua for } G(V)$$

$$\updownarrow$$

$$\begin{cases} Nk^{\times}\text{-orbits on nontrivial } \psi : k/k_0 \longrightarrow S^1, & \text{if } V \text{ is Hermitian;} \\ Nk^{\times}\text{-orbits on nontrivial } \psi_0 : k_0 \longrightarrow S^1, & \text{if } V \text{ is skew-Hermitian.} \end{cases}$$

As an illustration of the difference between the Hermitian and skew-Hermitian case, consider the case when $\dim V = 2$. When V is split, we may choose a basis $\{e, f\}$ of V so that $\langle e, f \rangle = 1$. If V is Hermitian, then $U(k_0) \cong \{x \in k : Tr(x) = 0\}$, so that generic characters of $U(k_0)$ are identified with characters of the trace zero elements of k, which are simply characters of k/k_0. On the other hand, if V is skew-Hermitian, then $U(k_0) \cong k_0$ so that generic characters of $U(k_0)$ are simply characters of k_0.

3.8 Generic parameters

Having fixed a Whittaker datum (U, χ), the LLC is also fixed, and has the following property. We say that an L-parameter ϕ is *generic* if $L_{k_0}(s, Ad \circ \phi)$ is holomorphic at $s = 1$, where Ad denotes the adjoint representation of ${}^L G(V)$. For a generic ϕ, the corresponding L-packet Π_ϕ contains generic representations. Then, relative to the fixed Whittaker datum (U, χ), the trivial character of S_ϕ corresponds to a (U, χ)-generic representation; moreover, this is the unique (U, χ)-generic representation in Π_ϕ.

3.9 Global L-function

Now suppose we are in the global situation and $\pi = \otimes_v \pi_v$ is an automorphic representation of $G(V)$. Let ρ be a representation of ${}^L G(V)$ as above. Then under the LLC, each local representation π_v has an L-parameter ϕ_v and so one has the local L-factor $L_{F_v}(s, \pi_v, \rho) := L_{F_v}(s, \rho \circ \phi_v)$. Thus, one may form the global L-function

$$L_F(s, \pi, \rho) = \prod_v L_{F_v}(s, \pi_v, \rho)$$

where the product converges absolutely for $Re(s) >> 0$. This is an instance of automorphic L-functions. One of the basic conjectures of the Langlands program is that such automorphic L-functions admit a meromorphic continuation to $\mathbb{C}$ and satisfy a standard functional equation relating its value at s to its value at $1 - s$. One has an analogous L-function $L_E(s, \pi, \rho)$ over E if ρ is a representation of $\mathrm{GL}_n(\mathbb{C})$.

Now for the group $G = G(V) \times G(W)$, we note that the groups $\mathrm{GL}_n(\mathbb{C})$ and $\mathrm{GL}_{n-1}(\mathbb{C})$ come with a standard or tautological representation std. Thus, taking $\rho = \mathrm{std}_n \boxtimes \mathrm{std}_{n-1}$, we have the corresponding global L-function

$$L_E(s, \pi, \rho) =: L_E(s, \pi_1 \times \pi_2) \quad \text{if } \pi = \pi_1 \boxtimes \pi_2.$$

By the results of [A] and [M], and the theory of Rankin-Selberg L-functions on $\mathrm{GL}(n) \times \mathrm{GL}(n-1)$, one knows that $L(s, \pi_1 \times \pi_2)$ has a meromorphic continuation to $\mathbb{C}$ and satisfies the expected functional equation.

4 The Conjecture

After introducing the LLC in the last section, we are now ready to state the Gross-Prasad (GP) conjecture.

4.1 Multiplicity One

Let $\phi = \phi_1 \times \phi_2$ be a generic L-parameter for $G = G(V) \times G(W)$, and let Π_ϕ be the associated Vogan L-packet. A representation $\pi \in \Pi_\phi$ is thus a representation of $G(V') \times G(W')$ where V' and W' are pure inner forms of V and W respectively. We call π relevant if $W' \subset V'$ is relevant.

Local Gross-Prasad I

If ϕ is a generic L-parameter for G, then

$$\sum_{\text{relevant } \pi \in \Pi_\phi} \dim \mathrm{Hom}_H(\pi \otimes \overline{\nu}, \mathbb{C}) = 1.$$

Thus, one has multiplicity one in Vogan packets: this is the generalisation of the lesson (b) mentioned in the introduction. The next part of the conjecture pinpoints the unique relevant π for which the associated Hom space is nonzero. This is the most delicate part of the local GP conjecture, and we shall consider the Bessel and Fourier-Jacobi case separately.

4.2 Distinguished character

We shall define a distinguished character on S_ϕ.

- **Bessel case.** Suppose first that $\epsilon = 1$ so that $\dim W^\perp = 1$. We need to specify a character η of S_ϕ, which determines the distinguished representation in Π_ϕ. Suppose that $\phi = \phi_1 \times \phi_2$, with

$$S_{\phi_1} = \prod_i \mathbb{Z}/2\mathbb{Z} \cdot a_i \quad \text{and} \quad S_{\phi_2} = \prod_j \mathbb{Z}/2\mathbb{Z} \cdot b_j,$$

  so that $S_\phi = S_{\phi_1} \times S_{\phi_2}$. Then we need to specify $\eta(a_i) = \pm 1$ and $\eta(b_j) = \pm 1$.
  We fix a nontrivial character $\psi : k/k_0 \longrightarrow S^1$ which determines the LLC for the even unitary group in $G = G(V) \times G(W)$. If $\delta \in k_0^\times$ is the discriminant of the odd-dimensional space in the pair (V, W), we consider the character $\psi_{-2\delta}(x) = \psi(-2\delta x)$. Then we set

$$\eta(a_i) = \epsilon(1/2, \phi_{1,i} \otimes \phi_2, \psi_{-2\delta});$$

  and

$$\eta(b_j) = \epsilon(1/2, \phi_1 \otimes \phi_{2,j}, \psi_{-2\delta}).$$

- **Fourier-Jacobi case.** Suppose now that $\epsilon = -1$ so that $W^\perp = 0$. In this case, we need to fix a character $\psi_0 : k_0 \longrightarrow S^1$ and a character χ of $k^\times$ with $\chi|_{k_0^\times} = \omega_{k/k_0}$ to specify the representation $\nu_{W,\psi_0,\chi}$. The recipe for the distinguished character η of S_ϕ depends on the parity of $\dim V$.
  - If $\dim V$ is odd, let $e = \mathrm{disc} V \in k_{Tr=0}$, well-defined up to $Nk^\times$, and define an additive character of k/k_0 by $\psi(x) = \psi_0(Tr(ex))$. We set

$$\begin{cases} \eta(a_i) = \epsilon(1/2, \phi_{1,i} \otimes \phi_2 \otimes \chi^{-1}, \psi); \\ \eta(b_j) = \epsilon(1/2, \phi_1 \otimes \phi_{2,j} \otimes \chi^{-1}, \psi). \end{cases}$$

- If $\dim V$ is even, the fixed character ψ_0 is needed to fix the LLC for $G(V) = G(W)$. We set

$$\begin{cases} \eta(a_i) = \epsilon(1/2, \phi_{1,i} \otimes \phi_2 \otimes \chi^{-1}, \psi); \\ \eta(b_j) = \epsilon(1/2, \phi_1 \otimes \phi_{2,j} \otimes \chi^{-1}, \psi), \end{cases}$$

where the epsilon characters are defined using any nontrivial additive character of k/k_0 (the result is independent of this choice).

4.3 Local Gross-Prasad

Having defined a distinguished character η of S_ϕ, we obtain a representation $\pi(\eta) \in \Pi_\phi$. It is not difficult to check that $\pi(\eta)$ is a representation of a relevant pure inner form $G' = G(V') \times G(W')$. Now we have:

Local Gross-Prasad II

Let η be the distinguished character of S_ϕ defined above. Then

$$\mathrm{Hom}_H(\pi \otimes \overline{\nu}, \mathbb{C}) \neq 0.$$

4.4 Global conjecture

Suppose now that we are in the global situation, with $\pi = \pi_1 \boxtimes \pi_2$ a cuspidal representation of $G(\mathbb{A}) = \mathrm{U}(V)(\mathbb{A}) \times \mathrm{U}(W)(\mathbb{A})$. It follows by [M] that π occurs with multiplicity one in the cuspidal spectrum. The global conjecture says:

Global Gross-Prasad Conjecture

The following are equivalent:
(i) The period interval $\mathcal{P}_H$ is nonzero when restricted to π;
(ii) For all places v, the local Hom space $\mathrm{Hom}_{H(F_v)}(\pi_v, \nu_v) \neq 0$ and in addition,

$$L_E(1/2, \pi_1 \times \pi_2) \neq 0.$$

Indeed, after the local GP, there will be a unique abstract representation $\pi(\eta) = \otimes_v \pi_v(\eta_v)$ in the global L-packet of π which supports a nonzero abstract $H(\mathbb{A})$-invariant linear functional. This representation lives on a certain group $G'_{\mathbb{A}} = \prod_v G(V'_v)$. To even consider the period integral on $\pi(\eta)$, one must first ask whether the group $G'_{\mathbb{A}}$ arises from a space V' over E, or equivalently whether the collection of local spaces $\{V'_v\}$ is coherent. A necessary and sufficient condition for this is that $\epsilon_E(1/2, \pi_1 \times \pi_2) = 1$.

4.5 The refined conjecture of Ichino-Ikeda

Ichino and Ikeda [II] have formulated a refinement of the global Gross-Prasad conjecture for tempered cuspidal representations on orthogonal groups. This takes the form of a precise identity comparing the period integral $\mathcal{P}_H$ with a locally defined H-invariant functional on π, with the special L-value $L(1/2, \pi_1 \times \pi_2)$ appearing as a constant of proportionality. Their refinement was subsequently extended to the Hermitian case by N. Harris [Ha].

More precisely, suppose that $\pi = \otimes_v \pi_v$ is a tempered cuspidal representation. The Petersson inner product $\langle -, - \rangle_{\mathrm{Pet}}$ on π can be factored (non-canonically):

$$\langle -, - \rangle_{\mathrm{Pet}} = \prod_v \langle -, - \rangle_v.$$

In addition, the Tamagawa measures dg and dh on $G(\mathbb{A})$ and $H(\mathbb{A})$ admit decompositions $dg = \prod_v dg_v$ and $dh = \prod_v dh_v$. We fix such decompositions once and for all.

For each place v, we consider the functional on $\pi \otimes \overline{\pi}$ defined by

$$I_v^{\#}(f_1, f_2) = \int_{H(F_v)} \langle f_1, f_2 \rangle_v \, dh_v.$$

This integral converges when π_v is tempered, and defines an element of $\mathrm{Hom}_{H_v \times H_v}(\pi_v \otimes \overline{\pi_v}, \mathbb{C})$. This latter space is at most 1-dimensional, as we know, and it was shown by Waldspurger that

$$I_v^{\#} \neq 0 \Longleftrightarrow \mathrm{Hom}_{H_v}(\pi_v, \mathbb{C}) \neq 0.$$

We would like to take the product of the $I_v^{\#}$ over all v, but this Euler product would diverge. Indeed, for almost all v, where all the data involved are unramified, one can compute $I_v^{\#}(f_1, f_2)$ when f_i are the spherical vectors used in the restricted tensor product decomposition of π and $\overline{\pi}$. One gets:

$$I_v^{\#}(f_1, f_2) = \Delta_{G(V),v} \cdot \frac{L_{E_v}(1/2, \pi_1 \times \pi_2)}{L_{F_v}(1, \pi, Ad)},$$

where

$$\Delta_{G(V),v} = \prod_{k=1}^{\dim V} L(k, \omega_{E_v/F_v}^k).$$

Thus, though the Euler product diverges, it can be interpreted by meromorphic continuation of the L-functions which appear in this formula. Alternatively, one may normalise the local functionals by:

$$I_v(f_1, f_2) = \left(\Delta_{G(V),v} \cdot \frac{L_{E_v}(1/2, \pi_1 \times \pi_2)}{L_{F_v}(1, \pi, Ad)} \right)^{-1} \cdot I_v^{\#}(f_1, f_2).$$

Then we may set

$$I = \prod_v I_v \in \mathrm{Hom}_{H(\mathbb{A}) \times H(\mathbb{A})}(\pi \otimes \overline{\pi}, \mathbb{C}).$$

Since the period integral $\mathcal{P}_H \otimes \overline{\mathcal{P}_H}$ is another element in this Hom space, it must be a multiple of I.

Refined Gross-Prasad conjecture

One has:

$$\mathcal{P}_H \otimes \overline{\mathcal{P}_H} = \frac{1}{|S_\pi|} \cdot \Delta_{G(V)} \cdot \frac{L_E(1/2, \pi_1 \times \pi_2)}{L_F(1, \pi, Ad)} \cdot I,$$

where S_π denotes the "global component group" of π (which we have not really introduced before).

When V is skew-Hermitian (i.e. in the Fourier-Jacobi case)., an analogous refined conjecture was formulated in a recent preprint of Hang Xue [X2].

5 Recent Progress: Local Case

We now come to the more interesting part of the paper, namely the account of some definitive recent results concerning the above conjectures.

5.1 **Proof of local conjecture for Bessel model.**

In a stunning series of papers [W1-4], Waldspurger has proved the Local GP conjecture for tempered representations of orthogonal groups over p-adic fields; the case of generic nontempered representations is then deduced from this by Moeglin-Waldspurger [MW]. Shortly thereafter, his student Beuzart-Plessis adapted the arguments to the p-adic Hermitian case (for tempered representations) discussed in this paper. The results are proved under the same hypotheses needed to establish the LLC for $G(V)$ (i.e. stabilisation of the twisted trace formula, the LLC for inner forms etc). Thus, one might state the result as:

Theorem 5.1 *Assume that k is p-adic and the LLC (in the refined form due to Vogan) holds for $G = G(V) \times G(W)$. Then the local Gross-Prasad conjecture holds in the Bessel case.*

It would not be possible to give a respectable account of Waldspurger's proof here, but it is nonetheless useful to highlight the key idea. We will do so in a toy model which was studied in [GGP2, §5].

5.2 Toy model over finite fields

Imagine for a moment that the field k is a finite field, so that G and H are finite groups of Lie type. Then given an irreducible representation π of $G(k_0)$, one has:

$$\dim \mathrm{Hom}_{H(k_0)}(\pi, \mathbb{C}) = \langle \chi_\pi, 1 \rangle_H$$

where χ_π denotes the character of π and the inner product on the RHS denotes the inner product of class functions on $H(k_0)$. This gives a character-theoretic way of computing the LHS, but how does this help?

We need another idea: base change to GL_n. More precisely, consider the groups $G(V)(k_0) = \mathrm{U}_n(k_0)$ and $G(V)(k) = \mathrm{GL}_n(k)$. The Galois group $\mathrm{Gal}(k/k_0)$ acts on the latter, and one has the notion of σ-conjugacy classes on $G(V)(k)$, where σ is the nontrivial element in $\mathrm{Gal}(k/k_0)$. It turns out that there is a natural bijection between

$$\{\text{conjugacy classes of } G(V)(k_0)\} \longleftrightarrow \{\sigma\text{-conjugacy classes of } G(V)(k)\},$$

thus inducing an isomorphism of the space of class functions and the space of σ-class functions. This isomorphism is in fact an isometry for the natural inner products on these two spaces.

In view of the above, one expects a relation between the irreducible representations of $G(V)(k_0)$ and the irreducible representations of $G(V)(k)$ which are invariant under $\mathrm{Gal}(k/k_0)$. Such σ-invariant representations Π of $G(V)(k)$ can be extended to a representation $\tilde{\Pi}$ of $G(V)(k) \rtimes \mathrm{Gal}(k/k_0)$, and the restriction of $\chi_{\tilde{\Pi}}$ to the non-identity coset

$G(V)(k)\cdot\sigma$ is a σ-class function. Now, if π is a "sufficiently regular" representation of $G(V)(k_0)$, then there is a σ-invariant representation Π of $G(V)(k)$ such that

$$\chi_\pi = \chi_{\tilde{\Pi}}|_{G(V)(k)\cdot\sigma}.$$

Thus, one has

$$\dim \mathrm{Hom}_H(\pi,\mathbb{C}) = \langle \chi_\pi, 1\rangle_{H(k_0)} = \langle \chi_{\pi_1}, \chi_{\pi_2}^\vee\rangle_{G(W)(k_0)} = \langle \chi_{\tilde{\Pi}_1}, \chi_{\tilde{\Pi}_2^\vee}\rangle_{G(W)(k)\cdot\sigma}.$$

Hence, one obtains the interesting identity:

$$\dim \mathrm{Hom}_{H(k)\rtimes \mathrm{Gal}(k/k_0)}(\tilde{\Pi},\mathbb{C}) = \frac{1}{2}\cdot \dim \mathrm{Hom}_{H(k)}(\Pi,\mathbb{C}) + \frac{1}{2}\cdot \dim \mathrm{Hom}_{H(k_0)}(\pi,\mathbb{C}).$$

In particular, if one understands the restriction problem for GL_n well, one can infer information about the restriction problem for unitary groups. For example, if we know that $\dim \mathrm{Hom}_{H(k)}(\Pi,\mathbb{C}) = 1$, then we deduce immediately that $\dim \mathrm{Hom}_{H(k_0)}(\pi,\mathbb{C}) = 1$, since the LHS is equal to 0 or 1!

5.3 Work of Waldspurger and Beuzart-Plessis

We can now give an impressionistic sketch of the contributions of Waldspurger and Beuzart-Plessis. If π is an irreducible representation, its character distribution χ_π is a locally integrable function on the regular elliptic set of $G(k_0)$. One would like to integrate this character function over $H(k_0)$ and relate this integral to $\dim \mathrm{Hom}_{H(k_0)}(\pi,\mathbb{C})$.

The first innovation is thus to give a character-theoretic computation of $\dim \mathrm{Hom}_{H(k_0)}(\pi,\mathbb{C})$. Of course, the naive integral above does not make sense, and one has to discover the appropriate expression. The expression discovered by Waldspurger is a sum over certain elliptic (not necessarily maximal) tori of weighted integrals involving the character χ_π. One such torus is the trivial torus.

Now when one sums up the above expression for $\dim \mathrm{Hom}_H(\pi,\mathbb{C})$ over all relevant π's in Π_ϕ, one may exploit the character identities involved in the Jacquet-Langlands type transfer between a group and its inner forms. It turns out that the sum of terms corresponding to a fixed nontrivial torus cancels out and thus vanishes. Only the term corresponding to the trivial torus survive and this gives local GP I (multiplicity one in Vogan packet).

To obtain the more precise local GP II, one uses the character identities of twisted endoscopy to relate the problem to the analogous one on GL_n, which one understands completely; this is similar to what was done in the toy model over finite fields, but it is decidedly more involved. In particular, to be able to detect the local epsilon factors, Waldspurger and Beuzart-Plessis needed to express the local epsilon factors (on GL_n) in terms of certain character theoretic integrals. In this way, local GP conjecture II was deduced.

6 Rcent Progress: Global Case

In this section, we give an account of recent results on the global GP conjecture.

6.1 Work of Ginzburg-Jiang-Rallis

Even before the GP conjectures were extended to all classical groups in [GGP1], Ginzburg, Jiang and Rallis [GRS1-3] have studied the question of the nonvanishing of the relevant global periods. In particular, they showed one direction of the global GP conjecture:

Theorem 6.1 *Let $\pi = \pi_{\boxtimes}\pi_2$ be a cuspidal representation of $G(\mathbb{A})$ and assume that there is a weak lifting of π to an automorphic representation $\Pi = \Pi_1 \boxtimes \Pi_2$ of $G(\mathbb{A}_E)$ (as given for example in [M] when G is quasi-split over F). Then, in both the Bessel and Fourier-Jacobi cases, we have:*

$$\mathcal{P}_H \text{ nonzero on } \pi \Longrightarrow L_E(1/2, \pi_1 \times \pi_2) \neq 0.$$

It is not clear if their approach can be used to prove the converse direction.

6.2 Relative trace formula on unitary groups

The most general method of attack for the global GP conjecture in the Hermitian case (i.e. Bessel case) is a relative trace formula (RTF) which was developed by Jacquet-Rallis [JR], almost concurrently as [GGP1] was being written. We shall give a brief description of this influential approach.

Consider the action R of $C_c^\infty(G(\mathbb{A}))$ on $L^2(G(F)\backslash G(\mathbb{A}))$ by right translation. For $f \in C_c^\infty(G(\mathbb{A}))$, the operator $R(f)$ is given by a kernel function

$$K_f(x,y) = \sum_{\gamma \in G(F)} f(x^{-1}\gamma y)$$

on $G(\mathbb{A}) \times G(\mathbb{A})$. In the usual trace formula, one is interested in computing $\mathrm{Tr}R(f)$ and when G is anisotropic, this trace is computed as

$$\mathrm{Tr}R(f) = \int_{G(F)\backslash G(\mathbb{A})} K_f(x,x)\, dx.$$

When G is F-isotropic, one needs to regularise the RHS, and this is the elaborate theory of the invariant trace formula developed by Arthur. One then develops two expression for this integral: a spectral side involving the representations appearing in the automorphic discrete spectrum, and a geometric side involving weighted orbital integrals. Thus, in principle (and also in practice), the distribution $f \mapsto \mathrm{Tr}R(f)$ contains the full information of the automorphic discrete spectrum of G.

In the relative trace formula, one is interested in detecting automorphic representations on which certain period integrals are nonzero. In the global GP conjecture, we are interested in the nonvanishing of the period $\mathcal{P}_H$. Thus it is reasonable to consider the integral

$$I(f) = \int_{(H(F)\backslash H(\mathbb{A}))^2} K_f(x,y)\, dx\, dy. \tag{6.1}$$

When some local conditions are placed on the test function $f = \otimes_v f_v$, the integral above actually converges, and one can try to obtain a spectral expansion of the integral, as well as a geometric one. Such a theory has been carried out by Wei Zhang [Zh2], building upon the work of Jacquet-Rallis [JR].

6.3 Spectral and geometric expansions

More precisely, we assume that:

- at some finite place v_1, f_{v_1} is the matrix coefficient of a supercuspidal representation;
- for another finite place v_2, f_{v_2} is supported on the "regular semisimple" elements (where the notion of regular semisimple is relative to the action of $H \times H$ on G).

Assuming f is of this form, the spectral expansion of (6.1) takes the form

$$I(f) = \sum_{\pi} J_\pi(f)$$

where the sum runs over cuspidal representations of $G(\mathbb{A})$ and

$$J_\pi(f) = \sum_{\phi} \left(\int_{H(F)\backslash H(\mathbb{A})} (\pi(f)\phi)(x)\, dx \right) \cdot \left(\overline{\int_{H(F)\backslash H(\mathbb{A})} \phi(x)\, dx} \right)$$

is the Bessel distribution associated to π, with the sum running over an orthonormal basis of π.

On the other hand, one has a geometric expansion of (6.1) which is given by

$$I(f) = \sum_{\gamma} O(\gamma, f)$$

where the sum runs over "regular semisimple" $H(F) \times H(F)$-orbits on $G(F)$ and the orbital integral is given by

$$O(\gamma, f) = \int_{H(\mathbb{A}) \times H(\mathbb{A})} f(x^{-1}\gamma y)\, dx\, dy.$$

Hence one has an equality

$$\sum_{\pi} J_\pi(f) = \sum_{\gamma} O(\gamma, f).$$

Since $G(V)(\mathbb{A})$ has nontrivial centre, it is possible to incorporate a central character χ everywhere in the above discussion and consider the χ-part of K_f and $I(f)$; we omit the technical details here.

6.4 Relative trace formula on GL

As in the discussion over finite fields and local fields, we are seeking to relate the problem of periods on $G(V)(\mathbb{A})$ to the analogous problem on $G(V)(\mathbb{A}_E) = \mathrm{GL}_n(\mathbb{A}_E) \times \mathrm{GL}_{n-1}(\mathbb{A}_E)$. Thus, we need to set up an analogous relative trace formula on $G(V)(\mathbb{A}_E)$ and compare it to the one on $G(V)(\mathbb{A})$ which we have above.

Thus for $f' = \otimes_v f'_v \in C_c^\infty(G(V)(\mathbb{A}_E))$, we have the kernel function $K_{f'}$ as above. We then consider the following integral of the kernel function:

$$I'(f') = \int_{x \in H(E)\backslash H(\mathbb{A}_E)} \int_{y \in H'(F)\backslash H'(\mathbb{A})} K_{f'}(x, y) \cdot \mu(y)\, dx,\, dy \tag{6.2}$$

where $H' \cong \mathrm{GL}_n \times \mathrm{GL}_{n-1}$ over F and μ is an automorphic character of H' given by

$$\mu = (\omega_{E/F}^{n-1} \circ \det) \otimes (\omega_{E/F}^{n} \circ \det).$$

Once again, we should have fixed a central character χ' and then consider the χ'-part of $K_{f'}$ and $I'(f')$, but we will ignore this technical issue in this paper.

The period over $H(\mathbb{A}_E)$ is the analog of the Gross-Prasad period for general linear groups. On the other hand, the period over $H'(\mathbb{A})$ conjecturally detects those automorphic representations of $G(\mathbb{A}_E)$ which lies in the image of the base change from $G(\mathbb{A})$; this is apparently a conjecture of Flicker-Rallis. Thus, the distribution I' is designed to capture automorphic representations of $G(\mathbb{A}_E)$ which comes from the unitary group $G(\mathbb{A})$ and at the same time supports a Gross-Prasad period.

In any case, assuming that f' satisfies the analogous condition as f, one has an equality

$$\sum_{\Pi} I_{\Pi}(f') = \sum_{\gamma'} O_{\gamma'}(f')$$

where the sum on the LHS runs over all cuspidal representations of $G(\mathbb{A}_E)$ and

$$I_{\Pi}(f') = \sum_{\phi} \left(\int_{H(E)\backslash H(\mathbb{A}_E)} (\Pi(f')\phi)(x)\, dx \right) \cdot \left(\overline{\int_{H'(F)\backslash H'(\mathbb{A})} \phi(x) \cdot \mu(x)\, dx} \right)$$

is the analogous Bessel distribution, and the sum on the RHS runs over all "regular semisimple" elements for the action of $H(E) \times H'(F)$ on $G(E)$ and

$$O(\gamma', f') = \int_{H(\mathbb{A}_E)} f(x^{-1}\gamma' y) \cdot \mu(y)\, dx\, dy$$

is the analogous orbital integral.

6.5 Transfer

Now one would like to compare the distribution I and I' by matching the geometric expansion of I and I'. As in the theory of the usual trace formula, one first considers whether the "regular semisimple" orbits occurring in the geometric expansion of I and I' are in natural bijection. This matching of "regular semisimple" orbits was done by Jacquet-Rallis [JR]. Thus, one is led to consider the comparison of orbital integrals on both sides. For this, one needs to introduce a transfer factor

$$\Delta_v : G(F_v)_{rs} \times G(E_v)_{rs} \longrightarrow \mathbb{C}$$

for each place v which is nonzero only on matching pairs (γ, γ') of regular semisimple elements. We may thus regard Δ_v as a function on $G(E_v)_{rs}$. Then the main properties of Δ_v are:

- (equivariance) For $(x, y) \in H(E_v) \times H'(F_v)$,

$$\Delta_v(xg'y) = \mu(y) \cdot \Delta_v(g').$$

- (product formula) For $\gamma' \in G(E)$,

$$\prod_v \Delta_v(\gamma') = 1.$$

It turns out that it is not hard to write down such a function. This allows us to make the following basic definition:

Definition We say that f and f' are matching test functions if

$$\Delta_v(\gamma') \cdot O(\gamma', f') = O(\gamma, f)$$

for every pair (γ, γ') of matching regular semisimple orbits.

One is thus led to the following local problems:

- (Fundamental lemma): For almost all places v of F, let f_0 and f_0' be the unit element in the spherical Hecke algebra of $G(F_v)$ and $G(E_v)$ respectively. Then f_0 and f_0' are matching functions.
- (Transfer) For every f, one can find a matching f'. Conversely, for every f', one can find a matching f.

These two statements are straightforward when v is a place of F which splits in E. Thus, the main issue is in the inert case.

6.6 Work of Zhiwei Yun

In [Y], Zhiwei Yun has verified the fundamental lemma over local function fields, using geometric techniques analogous to those used by B.C. Ngo in his thesis (which verified the Jacquet-Ye fundamental lemma for another relative trace formula). Having this, J. Gordon has shown (in the appendix to [Y]) that this implies the desired fundamental lemma over p-adic fields, when p is sufficiently large. Thus, we know that the fundamental lemma for unit elements hold.

6.7 Work of Wei Zhang

In a recent breakthrough paper [Zh1], Wei Zhang has shown that (Transfer) also holds at any p-adic place. Let us give an impressionistic sketch of the main steps of the proof.

- Using a Cayley transform argument, one reduces the existence of transfer on the group level to the existence of an analogous transfer on the level of Lie algebras;
- By a local characterisation of orbital integrals, one is reduced to showing the existence of transfer in a neighbourhood of each semisimple (not necessarily regular) element;
- Using the theory of generalised Harish-Chandra descent, as developed by Aizenbud-Gourevitch [AG], one is reduced to showing the existence of local transfer in a neighbourhood of the zero element (of another space);
- One shows that, modulo test functions which are killed by the infinitesimal version of I, the space of test functions is generated by those functions supported off the null cone and their various partial Fourier transforms. This requires an uncertainty principle type result of Aizenbud-Gourevitch on the nonexistence of certain equivariant distributions whose various Fourier transforms (including the trivial transform) are supported in the null cone.
- For test functions supported off the null cone, transfer is shown inductively.

- Finally, one shows that if f and f' are matching functions, then their partial Fourier transforms are also matching.

As a consequence of the fundamental lemma and the existence of transfer, one can compare the two RTF's and obtain an identity:

$$\sum_{\pi} J_{\pi}(f) = \sum_{\Pi} I_{\Pi}(f')$$

for matching test functions f and f' satisfying some local conditions as specified above. Using this, Wei Zhang showed in [Zh1] the global GP conjecture subject to certain local conditions. This is the strongest result on the global GP conjecture known to date.

Theorem 6.4 *Let $\pi = \pi_1 \boxtimes \pi_2$ be a cuspidal representation of $G = G(V) \times G(W)$. Suppose that*

- *all infinite places of F are split in E;*
- *for two finite places of F split in E, π_v is supercuspidal.*

Then the following are equivalent:
(i) $L_E(1/2, \pi_1 \times \pi_2) \neq 0$.
(ii) *For some relevant pure inner form $W' \subset V'$, and some cuspidal representation π' of $G' = G(V') \times G(W')$ which is nearly equivalent to π (i.e. belongs to the same global L-packet as π), the period integral $\mathcal{P}_H$ is nonzero on π'.*

Indeed, in view of the local GP, the pure inner form G' and its representation π' is uniquely determined, since there is a unique relevant representation in each local L-packet which could support a nonzero abstract H-invariant functional.

In a sequel [Zh2], Wei Zhang was able to refine the argument to derive the refined GP conjecture, as formulated by N. Harris [Ha].

Theorem 6.5 *Let $\pi = \pi_1 \boxtimes \pi_2$ be a tempered cuspidal representation of $G = G(V) \times G(W)$. Suppose that*

- *all infinite places of F are split in E;*
- *for at least one finite place of F split in E, π_v is supercuspidal;*
- *for every place v of F inert in E such that π_v is not unramified, either $H(F_v)$ is compact or π_v is supercuspidal.*

Then the refined GP conjecture holds for π, with $|S_{\pi}| = 4$.

6.8 Work of Yifeng Liu and Hang Xue

We have devoted considerable attention to the Bessel case of the global GP. Let us consider the Fourier-Jacobi case now. Following the work of Jacquet-Rallis [JR], Yifeng Liu [L] has developed an analogous relative trace formula in the skew-Hermitian case, which compares the Fourier-Jacobi period on unitary groups with the analogous period on general linear groups. He also showed that the fundamental lemma in this case can be reduced to that in the Bessel case.

Building upon this, Hang Xue [X1] has recently achieved in his thesis work the analog of [Zh1] for the Fourier-Jacobi case, thus establishing the global GP conjecture with some local conditions in the skew-Hermitian case. In a recent preprint [X2], Xue has formulated the refined GP conjecture in the Fourier-Jacobi case and then verified it subject again to some local conditions, analogous to what was done in [Zh2].

7 Outstanding Questions

After this lengthy discussion of recent progresses, it is perhaps time to take stock of the remaining questions concerning the GP conjecture. Here are some problems which come to mind:

(i) **Archimedean case:** Almost nothing is known about the local GP conjecture in the archimedean case. The methods developed by Waldspurger is character theoretic in nature and should apply in the archimedean case too, at least in principle. One naturally expects that there may be greater analytic difficulties in the archimedean case, but these should be regarded as more of a technical, rather than a fundamental, nature. Thus, it is reasonable to expect that someone with a strong analytic background could work through the proof of Waldspurger and Beuzart-Plessis and adapt the arguments to the archimedean case.

(ii) **Fourier-Jacobi case.** What about the Fourier-Jacobi case of the local GP conjecture in the p-adic case? As explained in [G], the Bessel and Fourier-Jacobi periods are connected by the local theta correspondence, and if one knows enough about the local theta correspondence, the Fourier-Jacobi case of the local GP will also follow. A pending work of the author with A. Ichino, which hopefully will appear soon, should establish this link conclusively and thus complete the proof of the Fourier-Jacobi case of local GP.

(iii) **Global conjecture for orthogonal groups.** The Jacquet-Rallis trace formula has been very successful for the Hermitian and skew-Hermitian case, but for the orthogonal groups, which is the original case of the GP conjecture, one still does not have a strategy which works in all cases. It will be very interesting to have a new approach. The main difficulty in formulating a relative trace formula which compares the orthogonal groups and the general linear groups is that there is no convenient characterisation, in terms of periods, of automorphic representations of $\mathrm{GL}(n)$ which are lifted from orthogonal groups.

(iv) **Arithmetic Case.** If $\epsilon_E(1/2, \pi_1 \times \pi_2) = -1$, the unique representation in the global L-packet of π which supports an abstract $H(\mathbb{A})$-invariant period lives on a group $\prod_v G(V_v')$ which is incoherent, i.e. there is no Hermitian space over E whose localisation agrees with $\{V_v'\}$. In this case, $L_E(1/2, \pi_1 \times \pi_2) = 0$ and the period integral $\mathcal{P}_H$ is automatically zero on each automorphic π' in the global L-packet of π.

It turns out that something even more interesting happens in this case. Namely, under some conditions, one may construct another "period integral" on π coming from a height pairing on a certain Shimura variety associated to G, and the nonvanishing of this arithmetic period integral is then governed by the nonvanishing of the central derivative $L_E'(1/2, \pi_1 \times \pi_2)$. The precise conjecture is given in [GGP1, §27].

Now as in the usual GP conjecture, one expects a refinement of this arithmetic GP conjecture in the form of an exact formula relating the arithmetic period integral to the locally defined functional I in the refined GP. Such a refinement has been proposed by Wei Zhang. For the group $\mathrm{U}(2) \times \mathrm{U}(1)$, it specializes to the generalised Gross–Zagier formula, which was shown in the recent book [YZZ] of Yuan-Zhang-Zhang in the parallel weight two case.

More amazingly, Wei Zhang has suggested that the relative trace formula of Jacquet-Rallis could be used to attack this refined arithmetic GP conjecture. This new application of the relative trace formula is an extremely exciting development, and it remains to be seen if it can be carried out in this case, and perhaps even in other scenarios.

Acknowledgements This paper is a written account of my lecture delivered at the annual meeting of the Vietnam Institute for Advanced Study in Mathematics held on July 20-21, 2013. I thank Bao Chau Ngo for his invitation to deliver a lecture and the local organisers for their warm hospitality and travel support.

References

[A] J. Arthur, *The endoscopic classification of representations: orthogonal and symplectic groups*, to appear as a Colloquium Publication of the American Mathematical Society (2013).

[AG] A. Aizenbud and D. Gourevitch, *Generalized Harish-Chandra descent, Gelfand pairs, and an Archimedean analog of Jacquet-Rallis's theorem*, with an appendix by the authors and Eitan Sayag, Duke Math. J. 149 (2009), no. 3, 509-567.

[AGRS] A. Aizenbud, D. Gourevitch, S. Rallis and G. Schiffmann, *Multiplicity One Theorems*, Ann. of Math. (2) 172 (2010), no. 2, 1407- 1434.

[BP] R. Beuzart-Plessis, *La conjecture locale de Gross-Prasad pour les représentations tempérées des groupes unitaires*, preprint (2012).

[De] P. Deligne, *Les constantes des équations fonctionnelles des fonctions* L, Modular Functions on one variable, II, pp 501-597, Lecture Notes in Mathematics, vol. 349, Springer, Berlin 1973.

[G] W. T. Gan, *Representations of metaplectic groups*, Fifth International Congress of Chinese Mathematicians. Part 1, 2, 155-170, AMS/IP Stud. Adv. Math., 51, pt. 1, 2, Amer. Math. Soc., Providence, RI, 2012.

[GGP1] W.T. Gan, B. H. Gross and D. Prasad, *Symplectic local root numbers, central critical L-values, and restriction problems in the representation theory of classical groups*, Asterisque 346 (2012), 1-110.

[GGP2] W.T. Gan, B. H. Gross and D. Prasad, *Restrictions of representations of classical groups: examples*, Asterisque 346 (2012), 111-170.

[GJR1] D. Ginzburg, D. Jiang, and S. Rallis, *On the nonvanishing of the central value of the Rankin-Selberg L-functions*, J. Amer. Math. Soc. 17 (2004), no. 3, 679–722.

[GJR2] D. Ginzburg, D. Jiang, and S. Rallis, *On the nonvanishing of the central value of the Rankin-Selberg L-functions, II*, Automorphic Representations, L-functions and Applications: Progress and Prospects, 157-191, Ohio State Univ. Math. Res. Inst. Publ. 11, de Gruyter, Berlin 2005.

[GJR3] D. Ginzburg, D. Jiang, and S. Rallis, *Models for certain residual representations of unitary groups*, in *Automorphic forms and L-functions I. Global aspects*, 125-146, Contemp. Math., 488, Amer. Math. Soc., Providence, RI, 2009.

[GP1] B. Gross and D. Prasad, *On the decomposition of a representation of* SO_n *when restricted to* SO_{n-1}. Canad. J. Math. 44 (1992), no. 5, 974–1002.

[GP2] B. Gross and D. Prasad, *On irreducible representations of* $\mathrm{SO}_{2n+1} \times \mathrm{SO}_{2m}$. Canad. J. Math. 46 (1994), no. 5, 930–950.

[Ha] N. Harris, *The Refined Gross-Prasad Conjecture for Unitary Groups*, to appear in IMRN, available at http://arxiv.org/abs/1201.0518

[HT] M. Harris and R. Taylor, *The geometry and cohomology of some simple Shimura varieties*, With an appendix by Vladimir G. Berkovich, Annals of Mathematics Studies, 151. Princeton University Press, Princeton, NJ, 2001.

[He1] G. Henniart, *Caractérisation de la correspondance de Langlands locale par les facteurs ? de paires*, Invent. Math. 113 (1993), no. 2, 339350.

[He2] G. Henniart, *Une preuve simple des conjectures de Langlands pour* $\mathrm{GL}(n)$ *sur un corps p-adique*, Invent. Math. 139 (2000), no. 2, 439–455.

[II] A. Ichino and T. Ikeda, *On the periods of automorphic forms on special orthogonal groups and the Gross-Prasad conjecture*, Geometric and Functional Analysis Vol. 19 (2010), No. 5, 1378-1425.

[JR] H. Jacquet and S. Rallis, *On the Gross-Prasad conjecture for unitary groups*, in *On certain L-functions*, 205-264, Clay Math. Proc., 13, Amer. Math. Soc., Providence, RI, 2011.

[K] T. Kaletha, *Rigid inner forms of real and p-adic groups*, preprint, available at http://arxiv.org/abs/1304.3292.

[KS] R. Kottwitz and D. Shelstad, *Foundations of twisted endoscopy.* Astérisque 255 (1999).

[L] Y. F. Liu, *Relative trace formulae toward Bessel and Fourier-Jacobi periods of unitary groups*, preprint (2010), available at http://math.mit.edu/ liuyf/Bessel.pdf.

[M] C. P. Mok, *Endoscopic classification of representations of quasi-split unitary groups*, to appear in Memoirs of AMS.

[MW] C. Moeglin and J.-L.Waldspurger, *La conjecture locale de Gross-Prasad pour les groupes spéciaux orthogonaux: le cas général*, Asterisuqe 347 (2012).
[S] Binyong Sun, *Multiplicity one theorems for Fourier-Jacobi models*, Amer. J. Math. 134 (2012), no. 6, 1655-1678.
[SZ] Binyong Sun, Chen-Bo Zhu, *Multiplicity one theorems: the Archimdean case*, Ann. of Math. (2) 175 (2012), no. 1, 23-44.
[Vo] D. Vogan, *The local Langlands conjecture*, Contemporary Maths, AMS, 145, 305-379 (1993).
[W1] J.-L.Waldspurger, *Une Formule intégrale reliée à la conjecture locale de Gross-Prasad,* Compositio Math. 146 (2010), 1180-1290.
[W2] J.-L. Waldspurger, *Une Formule intégrale reliée à la conjecture locale de Gross-Prasad, 2ème partie: extension aux représentations tempérées*, Asterisque 346, 171-311.
[W3] J.-L.Waldspurger, *Calcul d'une valeur d'un facteur epsilon par une formule intégrale*, Asterisque 347 (2012).
[W4] J.-L.Waldspurger, *La conjecture locale de Gross-Prasad pour les représentations tempérées des groupes spciaux orthogonaux*, Asterisque 347 (2012).
[W5] J.-L. Waldspurger, *Une variante d'un résultat de Aizenbud, Gourevitch, Rallis et Schiffmann*, Asterisuqe 346 (2012), 313-318.
[X1] H. Xue, *The Gan–Gross–Prasad conjecture for* $\mathrm{U}(n) \times \mathrm{U}(n)$, preprint (2012), available at http://math.columbia.edu/ xuehang/ggp.pdf.
[X2] H. Xue, *Fourier–Jacobi periods and the central value of Rankin–Selberg L-function*, preprint (2013).
[Y] Z. W. Yun, *The fundamental lemma of Jacquet and Rallis*, with an appendix by Julia Gordon, Duke Math. J. 156 (2011), no. 2, 167-27.
[YZZ] X. Yuan, S. W. Zhang and W. Zhang, *Heights of CM points I: Gross-Zagier formula*, preprint, available at *http://www.math.columbia.edu/ szhang/papers/HCMI.pdf.*
[Zh1] W. Zhang, *Fourier transform and the global Gan-Gross-Prasad conjecture for unitary groups*, preprint, available at http://www.math.columbia.edu/ wzhang/math/online/transfer.pdf.
[Zh2] W. Zhang, *Automorphic period and the central value of Rankin-Selberg L-function*, available at http://arxiv.org/abs/1208.6280.

PARABOLIC FLOWS IN COMPLEX GEOMETRY

Duong H. Phong

Abstract An informal introduction is provided to some current problems on canonical metrics in complex geometry, from the point of view of non-linear partial differential equations. The emphasis is on the parabolic flow approach and open problems.

1 Introduction

One of the most fundamental results in mathematics is the Uniformization Theorem, which says that a compact complex curve can be characterized by a Hermitian metric of constant scalar curvature and given area. This basic correspondence between complex/algebraic geometry on one hand and differential geometry on the other hand requires in turn on a third mathematical theory, namely the theory of non-linear partial differential equations. Indeed, fix a metric ds_0^2 on the complex curve and look for the desired metric under the form $ds^2 = e^{2u(x)}ds_0^2$. Since the scalar curvatures $R_0(x)$ and $R(x)$ of the metrics ds_0^2 and ds^2 are related by the transformation $R(x) = e^{-2u}(-2\Delta_0 u + R_0(x))$, where Δ_0 is the Laplacian with respect to ds_0^2, the problem reduces to solving the equation

$$2\Delta_0 u - R_0(z) + ce^{2u} = 0, \qquad c = \text{constant}. \tag{1.1}$$

The picture that has emerged starting from the works of Yau [64], Uhlenbeck-Donaldson-Yau [62] and others in the last 40 years is that the same relation between complex/algebraic geometry, differential geometry, and partial differential equations can be expected to hold in higher dimensions as well, albeit in a more sophisticated form: a "canonical metric" characterizing the underlying complex/algebraic structure should still exist, but it will have in general singularities. The singularities reflect the underlying global geometry, and will not occur if and only if the complex/algebraic structure is "stable" in a suitable algebraic geometric sense.

Work supported in part by the National Science Foundation under Grant DMS-12-66033. Contribution to the Annual Meeting of the Vietnam Institute for Advanced Study in Mathematics, 2013.

D.H. Phong
Columbia University, USA
E-mail: phong@math.columbia.edu

The partial differential equations describing canonical metrics are inherently geometric. But more precisely, their origin as an optimality condition on the curvature on a given space links them directly to the equations of theoretical physics, where all four fundamental forces in nature - electromagnetism, weak interactions, strong interactions, gravitation - manifest themselves through curvature. While the metrics can differ in signature - usually Euclidian in geometry and Lorentz in theoretical physics - the equations can still be often related by analytic continuation processes such as tunnelling and instanton effects, or serve as sources of inspiration for each other.

The goal of the present lecture is to describe several of the above geometric partial differential equations in complex geometry, including the Yang-Mills equation, the Hermitian-Einstein equation, the Kähler-Einstein equation, and the equation for metrics of constant scalar curvature in a given Kähler class. Given the breadth of the subject, and the many methods developed over the years, we have chosen to focus on one common approach, namely that of parabolic flows. In this approach, the equation at hand is viewed as a fixed point of a dynamical system. The main questions are then to determine the maximum time intervals of existence of the flow, the possible emergence of singularities, the continuation of solutions beyond singularities whenever possible, and the asymptotic behavior for large time. In geometry, the parabolic flow approach started with the harmonic maps flow of Eells and Sampson [26] and the Ricci flow of Hamilton [28]. It has been the vehicle for some of the most striking advances in geometry and physics ever since (see e.g. [30, 28, 35]), and this is likely to continue far in the future.

2 Examples of Geometric Flows

Given a differential equation, there will be in general many natural ways to represent it as the fixed point of a dynamical system. For example, for the equation (1.1), one may wish to consider the flow $u(z,t)$,

$$\dot{u} = 2\Delta_0 u - R_0(z) + ce^{2u}, \qquad u(x,0) = 0 \tag{2.1}$$

2.1 The Yamabe flow

One possible drawback of the flow (2.1) is that the geometric origin of the problem is obscured, depriving us of geometric intuition. A more geometric flow is

$$\dot{g}_{ij}(x,t) = -(R-c)g_{ij}(x,t), \qquad g_{ij}(x,0) = g^0_{ij}(x) \tag{2.2}$$

Here $ds^2(t) = g_{ij}(x,t)dx^i dx^j$ is a metric evolving with the time t, $R = R(x,t)$ is the scalar curvature of the evolving metric, and c is a constant. Clearly, the fixed points of this flow are also metrics of constant scalar curvature. Since the deformation $\dot{g}_{ij}$ is proportional to g_{ij}, the metric $g_{ij}(x,t)$ is conformally equivalent at all times to the original metric $g^0_{ij}(x)$, and we can set $g_{ij}(x,t) = e^{2u(x,t)}g^0_{ij}(x)$ for some scalar function $u(x,t)$. Using the same relation between scalar curvatures of conformally equivalent metrics that we used earlier in setting up (6.2), we can rewrite the flow (2.2) as

$$\dot{u} = e^{-2u}(2\Delta_0 u - R_0 + ce^{2u}). \tag{2.3}$$

Not surprisingly, the two flows (2.1) and (2.2) have the same fixed points. But from the point of view of partial differential equations, they are very different: the Laplacian

Δ_0 occurs in (2.2) with a factor e^{-2u}, which can diverge or go to 0 as t evolves. Thus the right hand side of (2.2) is not uniformly elliptic. However, its geometric formulation puts at our disposal all the tools and concepts of differential geometry.

Although our original motivation is the constant curvature problem for complex curves, the flow (2.2) is well-defined for Riemannian manifolds of arbitrary dimensions. It is called the Yamabe flow, and it is clearly of particular interest in the problem of finding a metric of constant scalar curvature in a given conformal class.

2.2 The Ricci flow

In two real dimensions, the scalar curvature determines the whole curvature tensor. In particular, the Ricci curvature can be expressed as $R_{ij} = \frac{1}{2}Rg_{ij}$, and the flow (2.2) can be rewritten as

$$\dot{g}_{ij} = -(R_{ij} - \mu g_{ij}), \qquad g_{ij}(x,0) = g^0_{ij}(x) \tag{2.4}$$

where $\mu = c/2$ is a constant. The flow (2.4) is Hamilton's Ricci flow [27]. Clearly, it differs significantly from the Yamabe flow in higher dimensions: in general it will not preserve the conformal class of the initial metric, and its fixed points are Einstein metrics, i.e., metrics satisfying the equation $R_{ij} = \mu g_{ij}$. The Ricci flow is one of the main topics in this lecture, and we shall write it down more explicitly later in the Kähler case.

2.3 Mean curvature flows

In extrinsic geometry, there is considerable interest in finding submanifolds, say of $\mathbf{R}^N$, of constant mean curvature, of which minimal surfaces would be a prime example. For this, we can look for the fixed points of the mean curvature flow, defined in the case of hypersurfaces by

$$\dot{P}(x,t) = H\boldsymbol{N}, \qquad P(x,0) \in S(0) \tag{2.5}$$

where $P(x,t)$ are the points of an evolving hypersurface $S(t)$, H is the mean curvature of $S(t)$, and N is its unit normal. In terms of a parametrization $x^i \to y^\alpha(x,t)$, $1 \le i \le n$, $1 \le \alpha \le n+1 \equiv N$ of the hypersurface $S(t)$, the second fundamental form $\mathit{\Pi}_{ij}$ is given by $\mathit{\Pi}^\alpha_{ij} = \nabla_i \nabla_j y^\alpha$. It turns out that it is normal to $S(t)$ for any i, j. Thus $H\boldsymbol{N}^\alpha = \Delta y^\alpha$, where Δ is the Laplacian on $S(t)$ with respect to the induced metric, and the mean curvature flow can be written explicitly as

$$\dot{y}^\alpha(x,t) = \Delta y^\alpha. \tag{2.6}$$

Formally, it resembles the usual heat equation, but it is non-linear, since the Laplacian Δ depends on $y^\alpha(x,t)$.

Similar equations can be written down for the case of submanifolds of higher codimension, as well as for flows by other extrinsic notions of curvature, such as the Gaussian curvature or the inverse mean curvature, which is of considerable interest in general relativity.

2.4 The Yang-Mills flow

In this lecture, we shall however concentrate on flows of intrinsic structures, which generalize in some sense the Ricci flow to higher dimensions. A prototype is the Yang-Mills flow, which can be described as follows.

Let $E \to X$ be a smooth vector bundle of rank r over a compact smooth manifold X of dimension n, equipped with a metric $g_{ij}(x)$. Let $\Lambda^p(X)$ be the bundle of p-forms over X. A connection ∇ on E is a linear map $\nabla : C^\infty(X,E) \to C^\infty(X, E\otimes \Lambda^1)$ which can be expressed locally as follows. If $x = (x^\alpha)$ are local coordinates for X and sections φ of E expressed by vectors φ^α in a local trivialization of E, then $\nabla\varphi = \nabla_j \varphi^\alpha dx^j$, with

$$\nabla_j \varphi^\alpha(x) = \partial_j \varphi^\alpha(x) + A^\alpha_{j\beta}(x)\varphi^\alpha(x). \tag{2.7}$$

The curvature ${F_{ij}}^\alpha{}_\beta$ of the connection ∇ is defined by the commutation laws

$$[\nabla_i, \nabla_j]\varphi^\alpha = -{F_{ij}}^\alpha{}_\beta \varphi^\beta. \tag{2.8}$$

It follows readily from the definition that $F(A) \equiv {F_{ij}}^\alpha{}_\beta dx^j \wedge dx^i$ is a genuine 2-form, valued in the bundle $End(E)$ of endomorphisms of E, i.e. $F(A) \in C^\infty(X, End(E)\otimes \Lambda^2)$. When a metric $H_{\alpha\beta}$ is given on the bundle E, we can also consider connections which are unitary, in the sense that

$$\partial_j \langle \varphi, \psi\rangle = \langle \nabla_j \varphi, \psi\rangle + \langle \varphi, \nabla_j \psi\rangle \tag{2.9}$$

for arbitrary sections φ and ψ of E, and the inner product is taken with respect to H. If we let $e_a = ({e^\alpha}_a)$ be a local frame for E, i.e. $\{e_a\}$ is an orthonormal set of smooth sections of E and $\{e^b\} = ({e^b}_\beta)$ is the dual frame, ${e^b}_\alpha {e^\alpha}_a = {\delta^b}_a$, then sections φ^α of E can be expressed in terms of frames by $(\varphi^\alpha) \leftrightarrow (\varphi^a)$, with $\varphi^\alpha = \varphi^a {e^\alpha}_a$. The same applies for sections of E^* and $End(E)$. Expressed in this manner, the connection and curvature tensor are anti-symmetric, $A^a_{jb} = -A^b_{ja}$ and ${F_{jk}}^a{}_b = -{F_{jk}}^b{}_a$.

Let g_{ij} be a fixed metric on the manifold X, and $H_{\alpha\beta}$ a fixed metric on the vector bundle E. The Yang-Mills equation is the following equation for the connection $A = A^\alpha_{j\beta}dx^j$,

$$\nabla^j {F_{jk}}^\alpha{}_\beta = 0. \tag{2.10}$$

It can be viewed as the Euler-Lagrange equation of the Yang-Mills functional,

$$I_{YM}(A) = \int_X |F|^2 \sqrt{g} \tag{2.11}$$

where $|F|^2$ is more explicitly given by $|F|^2 = g^{jp}g^{kq}{F_{jk}}^\alpha{}_\beta {F_{pq}}^\gamma{}_\delta H_{\alpha\gamma}H^{\beta\delta}$, with g^{jp} and $H^{\beta\delta}$ the inverses of the metrics g_{ij} and $H_{\alpha\beta}$ respectively. Thus the Yang-Mills equation is, intuitively speaking, the equation for minimizing the total curvature of the vector bundle E, measured in the L^2 norm with respect to the metrics g_{ij} on X and $H_{\alpha\beta}$ on E.

The Yang-Mills flow is the gradient flow for the Yang-Mills functional. Thus we look for a time-evolution of connections $t \to A^\alpha_{k\beta}(t,x)$ satisfying the equation

$$\dot A^\alpha_{k\beta} = -\nabla^j {F_{jk}}^\alpha{}_\beta, \qquad A^\alpha_{k\beta}(0,x) = (A_0)^\alpha_{k\beta}(x) \tag{2.12}$$

where $(A_0)^\alpha_{k\beta}$ is a given initial connection, and the dot indicates time derivatives. Clearly the flow decreases the Yang-Mills functional, and stationary points of the flow will satisfy the Yang-Mills equation.

It is known that the Yang-Mills flow will exist and converge to a Yang-Mills connection when X is a manifold of dimension $n \leq 3$. At the other end, when the dimension of X is $n \geq 5$, it is known that the flow will develop singularities in finite time. It is still not known at the present time if the Yang-Mills flow will develop singularities in finite time when $n = 4$. The dimension $n = 4$ is special, since the Yang-Mills equation acquire then an additional symmetry which is conformal invariance.

The Ricci flow can be viewed as an even more non-linear version of the Yang-Mills flow. Indeed, if X is a Riemannian manifold, and we consider $E = T(X)$, then the Levi-Civita connection A evolves under the Ricci flow by the same equation (2.12). Note however that in the case of the Yang-Mills flow, the metric g_{ij} on the base is fixed independently of the evolving connection A, while in the Ricci flow the two are related by the requirement that A be the Levi-Civita connection of g_{ij}.

2.5 The Calabi flow

All the flows that we have described so far are second-order flows, in the sense that they can be written as partial differential equations of order 2. We conclude this discussion of examples with a fourth-order flow, namely the Calabi flow defined by

$$\dot{g}_{ij} = \nabla_i \nabla_j R, \qquad g_{ij}(0) = g^0_{ij}. \tag{2.13}$$

This flow is of considerable interest in Kähler geometry, and we shall return to it later in this context.

3 The Yang-Mills flow and complex structures

In general, without some additional information on the geometry of the base manifold X or of the bundle E, we have only scant information on the existence, regularity, and general behavior of the geometric flow. Here we discuss some remarkable properties which arise in the Yang-Mills flow when the underlying manifold X is complex. We also assume that it is equipped with a Kähler metric, although this assumption can be relaxed.

To motivate them, we start from a simple, but striking observation which sparked great developments in geometry and physics in the late 1970's: it is that a class of solutions to the 4-dimensional Yang-Mills equation is provided by the so-called instantons.

For this, we need an alternative formulation of the Yang-Mills functional which exhibits more clearly its topological aspects. Note first that since F_{jk} is an anti-symmetric matrix, its norm can be expressed by the Killing form. More concretely,

$$F_{jk}{}^\alpha{}_\beta F_{pq}{}^\gamma{}_\delta H_{\alpha\gamma} H^{\beta\delta} = F_{jk}{}^a{}_b F_{pq}{}^a{}_b = -F^a_{jkb} F^b_{pqa} = -\mathrm{Tr}\,(F_{jk} F_{pq}). \tag{3.1}$$

Next, let $\star$ be the Hodge $\star$-operator, defined on p-forms on X by the requirement that $A \wedge \star B = \langle A, B \rangle \sqrt{g}\, dx^1 \wedge \cdots \wedge dx^n$, for $A, B \in \Lambda^p(X)$. We can now rewrite the Yang-Mills functional as

$$I_{YM}(A) = -\int_X \mathrm{Tr}(F \wedge \star F) \tag{3.2}$$

Now, in 4-dimensions, the Hodge $\star$-operator is an endomorphism of $\Lambda^2(X)$ with eigenvalues ± 1 and a corresponding orthogonal eigenspace decomposition

$$\Lambda^2 = \Lambda^2_+ \oplus \Lambda^2_+. \tag{3.3}$$

Decomposing accordingly the curvature F as $F = F_+ + F_-$, where $F_\pm$ are forms in $\Lambda^2_\pm$ with values in $End(E)$, the Yang-Mills functional takes the form

$$I_{YM}(A) = \int_X |F_+|^2 + \int_X |F_-|^2. \tag{3.4}$$

On the other hand, the Chern-Weil theory says that the de Rham cohomology classes $c_k(E)$ of the $2k$-forms $\mathrm{Tr}(\wedge^k F)$ are independent of A and define topological invariants of the bundle. In particular,

$$\langle c_2(E), X\rangle = \int_X \mathrm{Tr}(F\wedge F) = \int_X |F_-|^2 - \int_X |F_+|^2. \tag{3.5}$$

and we obtain the lower bound

$$I_{YM}(A) \geq |\langle c_2(E), X\rangle|. \tag{3.6}$$

Thus the minimum of $I_{YM}(A)$ is achieved when either F_+ or F_- vanishes, depending on the sign of $\langle c_2(E), X\rangle$. This condition is only of first order in A, and such connections are called instantons. A first explicit instanton solution was written down by Belavin-Polyakov-Schwarz-Tyuptin [4] for $SU(2)$ bundles over S^4.

In presence of a Kähler structure on X, the instanton equation admits a remarkable interpretation. In this case, the eigenspaces $\Lambda^2_\pm$ can be explicitly identified as

$$\Lambda^2_+ = \Lambda^{2,0} \oplus [\omega] \oplus \Lambda^{0,2}, \quad \Lambda^2_- = \Lambda^{1,1} \ominus [\omega]. \tag{3.7}$$

Thus the condition that a connection A be an instanton, say $F_+(A) = 0$, is equivalent to

$$F_{\bar j \bar k} = 0, \quad F_{jk} = 0, \quad g^{j\bar k} F_{\bar k j} = 0. \tag{3.8}$$

The first two conditions simply mean that the connection A is equipping the vector bundle $E \to X$ with the structure of a holomorphic vector bundle. Indeed, a section s of E can be defined to be holomorphic if $\nabla_{\bar j} s = 0$, and the condition that $F_{\bar j \bar k} = 0$ just means that this structure is integrable, i.e., there are enough linearly independent holomorphic sections for the bundle to be locally holomorphically trivial. The last condition $g^{j\bar k} F_{\bar k j} = 0$ is an analogue of the vacuum Einstein equation $R_{jk} = 0$ in general relativity. Because of this, it is called the Hermitian-Einstein condition. Topologically, it requires that the degree of E, defined by $\deg(E) = \langle X, c_1(E) \wedge [\omega^{n-1}]\rangle$ be 0. Thus we can generalize the instanton equation to arbitrary Hermitian bundles $E \to (X, \omega)$ by taking the equation (3.8), with the third condition replaced by

$$\Lambda F - \kappa I = 0. \tag{3.9}$$

where Λ is the Hodge contraction operator defined by $\Lambda \Psi_{\bar k j} \equiv g^{j\bar k} \Psi_{\bar k j}$, and κ is a constant. Integrating the equation, we find that κ is determined by the Chern class of E and the Kähler class of ω,

$$\kappa = \frac{1}{V} \frac{\int_X c_1(E) \wedge \omega^{n-1}}{\mathrm{rank}\, E} \tag{3.10}$$

where $V = \langle \omega^n, X\rangle$ is the volume of X.

A very important property of the Yang-Mills flow is that the flow stays in the complex orbit of the initial connection. What this means is that the complex structures defined by $\bar\partial^{A(t)} \equiv \partial_{\bar j} + A_{\bar j}(t)$ are all equivalent for any t. Indeed, if g is an endomorphism of E, we can define the gauge group action of g on a connection A by

$$\begin{aligned}(g\cdot A)_{\bar k} &= gA_{\bar k}g^{-1} - (\partial_{\bar k} g)g^{-1}\\ (g\cdot A)_k &= \tilde g A_{\bar k}\tilde g^{-1} - (\partial_{\bar k}\tilde g)\tilde g^{-1}\end{aligned} \tag{3.11}$$

where $\tilde g \equiv (g^*)^{-1}$, and g^* is the adjoint of g with respect to the fixed metric H. We have then $\bar\partial^{g\cdot A} = g\bar\partial^A g^{-1}$, so that the complex structures defined by A and $g\cdot A$ are equivalent. The space of all such transformations is the complex gauge group $\mathcal{G}^c$, and the space of all $\{g\cdot A\}$ is the orbit of A. It is not difficult to check that the Yang-Mills flow is always tangent to the orbit of $A(t)$ for any t, and hence our earlier assertion. The group $\mathcal{G}^c$ is in a sense the complexification of the gauge $\mathcal{G}$ defined by $\tilde g = g$, which is the group of gauge transformations preserving both the complex and the metric structure of E (see e.g. [24]).

So far, we have discussed the Yang-Mills flow as a flow for a connection A, with a Hermitian metric $H \equiv H_0$ on E fixed, and the complex structure defined by A evolving. There is an equivalent viewpoint, by which we fix the complex structure on E, and evolve the metric $H_{\bar k j}$ according to the following flow, called the Donaldson heat flow,

$$H^{-1}\dot H = -(\Lambda F - \mu I). \tag{3.12}$$

In view of the fact that the evolution of the connection A by the Yang-Mills flow preserves the complex orbit of A, we can write $A(t) = g(t)\cdot A_0$ for $g(t)$ a complex-valued gauge transformation. If we define the endomorphism $h(t)$ by $h(t) = g(t)^* g(t)$, and $H(t) = H_0 h(t)$, then the metric $H(t)$ satisfies the Donaldson heat flow. Conversely, if $H(t) = h(t)H_0$ satisfies the Donaldson heat flow, then the connection $A(t) = h(t)^{\frac12} A$ is gauge equivalent to a solution of the Yang-Mills flow. In the complex setting, we view then the flows (2.12) and (3.12) as equivalent.

Analytically, a first major advantage in presence of a complex structure is that the Yang-Mills flow exists for all time. So the real issue is the asymptotic behavior of the flow as $t \to +\infty$. The convergence of the flow would imply the existence of a fixed point, and hence of a solution to the Hermitian-Einstein equation (3.9). It is a remarkable fact that this turned out to be equivalent to a deep, algebraic-geometric property of the bundle E.

Let $E \to X$ be an irreducible holomorphic vector bundle. For each subsheaf $\mathcal{F}$ of $\mathcal{O}(E)$, define the slope $\mu(\mathcal{F})$ by

$$\mu(\mathcal{F}) = \frac{1}{\mathrm{rank}(F)}\int_X c_1(\mathcal{F})\wedge\omega^{n-1}. \tag{3.13}$$

The bundle E is said to be stable in the sense of Mumford-Takemoto if for any subsheaf $\mathcal{F}$ of E, we have $\mu(\mathcal{F}) < \mu(E)$ if $\mathcal{F} \neq \mathcal{O}(E)$. Then the theorem of Donaldson-Uhlenbeck-Yau [20, 62] asserts that the holomorphic irreducible vector bundle $E \to X$ admits a Hermitian-Einstein metric if and only if it is stable in the sense of Mumford-Takemoto.

The direct link that this theorem provides between PDE and algebraic geometry can actually be made much more precise: in the paradigm of Uhlenbeck-Yau [62], the solvability of the PDE is equivalent to an a priori C^0 estimate, and it is the failure of this a priori estimate which gives rise to a destabilizing sheaf $\mathcal{F}$ of E. More concretely, consider the following family of equations:

$$\Lambda F_\varepsilon - \mu I = -\varepsilon \log h_\varepsilon \tag{3.14}$$

for a positive endomorphism h of E, and where F_ε is the curvature of the metric $H_\epsilon \equiv H_0 h_\varepsilon$. It is not difficult to show that for each fixed $\varepsilon > 0$, the equation admits a solution h_ε. The question is, is the set of solutions H_ε precompact in the C^2 norm, so that a subsequence can be extracted that converges in C^2 norm to an endomorphism h_∞ and hence a metric H_ε satisfying the Hermitian-Einstein equation? A first step in [62] is to show that a C^0 bound on h_ε would imply a C^k bound on h_ε for any $k \geq 1$. By the Arzela-Ascoli theorem, we see now that the solvability of the equation is reduced to the existence of a C^0 estimate for H_ε.

Uhlenbeck-Yau [62] show that, if $\sup_X \mathrm{Tr}\, h_\varepsilon \to \infty$, then a subsequence of the rescaled endomorphisms $\tilde{h}_\varepsilon \equiv (\sup_X \mathrm{Tr}\, h_\varepsilon)^{-1} h_\varepsilon$ tends in the Sobolev norm $\| \cdot \|_{(1)}$ to an endomorphism h_∞, and that the projection π defined by

$$\pi = \lim_{\varepsilon \to 0} (I - h_\infty^\sigma) \tag{3.15}$$

is an orthogonal projection, in $H_{(1)}(X, End(E))$, and satisfies the key condition

$$(1 - \pi)\bar{\partial}\pi = 0 \quad a.e. \tag{3.16}$$

They establish another deep theorem which shows that for such a projection π, there exists a subvariety $S \subset X$ away from which π is smooth, and a coherent analytic subsheaf $\mathcal{F} \subset \mathcal{O}(E)$ which is holomorphic away from S and agrees there with $\pi(E_{|X\setminus S})$. In the particular case when π arises from the above construction with rescaled endomorphisms $\tilde{h}_\varepsilon$, they can then show that $\mathcal{F}$ is destabilizing.

We return now to the discussion of the Yang-Mills flow. In the original proof of Donaldson [20], it was shown that when E is stable in the sense of Mumford-Takemoto, then a subsequence of metrics $H(t)$ converges weakly to a Hermitian-Einstein metric. Although it does not appear that the detailed arguments are available in the literature, there is little doubt that the full convergence of the flow when E is table can be established by either a refinement of the arguments of [20] or an adaptation of the methods of [62].

The main problem is then to determine the behavior of the flow for general holomorphic vector bundles. In view of the Donaldson-Uhlenbeck-Yau theorem, the limiting behavior is necessarily singular. Let $t_j \to \infty$ be a sequence of times. Define the set Z_{an} of analytic singularities by

$$Z_{an} = \cap_{r>0} \{x \in X; \mathrm{liminf}_{j\to\infty} r^{4-2n} \int_{B_r(x)} |F(A(t_j)|^2 \omega^n \geq \varepsilon\} \tag{3.17}$$

for a small $\varepsilon > 0$ (it turns out that there is a lower bound on ε depending only on X). Thus the set of analytic singularities is the set of points where the curvature concentrates. It is known that it has real codimension at least 4, and that the Yang-Mills flow will converge away from Z_{an} [61,29], and define a holomorphic vector bundle E_∞ on $X \setminus Z_{an}$. In 2 complex dimensions, where the singular set Z_{an} is a set of isolated points, it is a classical theorem of Uhlenbeck [61] the bundle E_∞ can be extended to a bundle $\hat{E}_\infty$ on the whole of X, whose topology, a fortiori whose complex structure, may be different from that of E. The conjecture of Bando-Siu [3] goes further: it asserts that the extension $\hat{E}_\infty$ actually coincides with a canonical bundle determined by the complex structure of E, namely the double dual $Gr_\omega^{hns}(E)^{**}$ of its Harder-Narasimhan-Seshadri filtration. The Bando-Siu conjecture was proved in complex dimension $n = 2$ by G. Daskalopoulos-Wentworth [18,19]. They also found in this case a striking interpretation of the analytic set Z_{an} in algebraic-geometric terms: define the set Z_{alg} of algebraic singularities by

$$Z_{alg} = \{x \in X;\ Gr_\omega^{hns}(E)_x \text{ is not free}\} \tag{3.18}$$

Then in dimension $n = 2$, they showed that

$$Z_{an} = Z_{alg}. \tag{3.19}$$

In particular, the set Z_{an} is independent of the sequence of times $\{t_j\}$. It is natural to conjecture that this holds in all dimensions. Very recently, the Bando-Siu conjecture has been proved in all dimensions, and there has also been considerable progress towards proving this last conjecture on the equality of the analytic and algebraic sets of singularities [43,31,32,16].

4 The Kähler-Ricci flow

We consider next the Kähler-Ricci. This is the Ricci flow when the underlying manifold (X, ω) is Kähler,

$$\dot g_{\bar kj} = -R_{\bar kj} + \mu g_{\bar kj}, \qquad g_{\bar kj}(z,0) = g^0_{\bar kj}(z), \tag{4.1}$$

with μ a constant which can be normalized to be 0, 1 or -1. Recall that a Hermitian metric $g_{\bar kj}$ is Kähler if the corresponding $(1,1)$-form $\omega = \frac{i}{2} g_{\bar jk} dz^k \wedge d\bar z^j$ is closed. We denote the cohomology class of ω by $[\omega]$. The fixed points of the flow are Kähler-Einstein metrics, defined to be Kähler metrics satisfying the equation

$$R_{\bar kj} = \mu g_{\bar kj}. \tag{4.2}$$

Clearly, this is a Kähler version of the Einstein vacuum equation, as well as a non-linear version of the Hermitian-Einstein equation defined in (3.8).

The Kähler-Ricci flow has many advantages over the more general Ricci flow on smooth manifolds, due to the cohomological properties of the Ricci curvature for Kähler metrics. For Kähler metrics, the Levi-Civita connection is consistent with the complex structure, and given by the Chern unitary connection,

$$\nabla_{\bar j} V^p = \partial_{\bar j} V^p, \qquad \nabla_k V^p = g^{p\bar m} \partial_k (g_{\bar mq} V^q) \tag{4.3}$$

The Riemann curvature tensor is obtained by the same commutation relations in (2.8), $[\nabla_k, \nabla_{\bar j}] V^p = R_{\bar jk}{}^p{}_q V^q$, and hence

$$R_{\bar jk}{}^p{}_q = -\partial_{\bar j}(g^{p\bar m} \partial_k g_{\bar mq}). \tag{4.4}$$

It satisfies the identity $R_{\bar jk\bar mq} = R_{\bar mk\bar jq} = R_{\bar mq\bar jk}$. Thus the Ricci curvature $R_{\bar jk}$ is given by $R_{\bar jk} = R_{\bar jp}{}^p{}_k = R_{\bar jk}{}^p{}_p$, and we find, in view of (4.4),

$$R_{\bar jk} = -\partial_k \partial_{\bar j} \log \det (g_{\bar pq}). \tag{4.5}$$

Since $\det(g_{\bar{p}q})$ is a metric on the anti-canonical bundle K_X^{-1}, it follows that the Ricci form $Ric(\omega) = \frac{i}{2} R_{\bar{j}k} dz^k \wedge d\bar{z}^j = -\frac{i}{2}\partial\bar{\partial}\log\det(g_{\bar{p}q})$ is a closed form in the cohomology class $c_1(K_X^{-1}) \equiv c_1(X)$. A $(1,1)$ de Rham cohomology class is said to be positive (respectively negative) if it admits a smooth representative which is strictly positive (respectively) as a current. Thus a necessary condition for the existence of a Kähler-Einstein metric is that the class $c_1(X) = 0$ when $\mu = 0$, $c_1(X) > 0$ when $\mu = 1$, and $c_1(X) < 0$ when $\mu = -1$. We say in these cases that $c_1(X)$ is definite.

Since the right hand side of (4.1) is closed, it follows that the Kähler property of $g_{\bar{k}j}$ is preserved along the flow. When $c_1(X)$ is definite, we can make sure by an appropriate choice of original cohomology class $[\omega_0]$ that the right hand side of (4.1) has vanishing cohomology class. When $c_1(X) = 0$, this is true for ω_0 in any Kähler class. When $c_1(X)$ is strictly positive or negative, it suffices to choose ω_0 in $\pm c_1(X)$. With this choice, $\omega(t)$ will remain in the same Kähler class for all time. This allows to express the Kähler-Ricci flow, which is a flow of metrics, as a flow of scalar functions. Indeed, by the $\partial\bar{\partial}$-lemma, we can express the Kähler form $\omega(t)$ corresponding to the evolving metric $g_{\bar{j}k}(t)$ as

$$\omega(t) = \omega_0 + \frac{i}{2}\partial\bar{\partial}\varphi \tag{4.6}$$

for some potential $\varphi(z,t)$ satisfying the plurisubharmonicity condition $\omega_0 + \frac{i}{2}\partial\bar{\partial}\varphi > 0$. The right hand side of the Kähler-Ricci flow can be expressed as

$$Ric(\omega) - \mu\omega = -\frac{i}{2}\partial\bar{\partial}\log\frac{\omega^n}{\omega_0^n} + Ric(\omega_0) - \mu\omega_0 - \frac{i}{2}\mu\partial\bar{\partial}\varphi. \tag{4.7}$$

By the definiteness condition on $c_1(X)$, the choice of the inital Kähler form ω_0 and again the $\partial\bar{\partial}$-lemma, we can write

$$Ric(\omega_0) - \mu\omega_0 = \frac{i}{2}\partial\bar{\partial}f, \tag{4.8}$$

for some smooth scalar function f. Comparing with the left hand side of the Kähler-Ricci flow, we can integrate, and arrive at the following parabolic Monge-Ampère equation for the Kähler potential φ,

$$\dot{\varphi} = \log\frac{\omega^n}{\omega_0^n} - f + \mu\varphi + c(t), \qquad \varphi(z,0) = c_0 \tag{4.9}$$

where $c(t), c_0$ are independent of z. The choice of constants $c(t), c_0$ is immaterial in the end, but a convenient choice is $c(t) = 0$ for all t, and c_0 a suitable value (see e.g. [39]).

The existence of Kähler-Einstein metrics was proved in the 1970's by Yau [64] and Aubin [1] when $c_1(X) < 0$, and by Yau [64] when $c_1(X) = 0$. This last case is also known as the Calabi conjecture, and it plays a very important role in superstring theory [6]. In the $c_1(X) > 0$ case, Kähler-Einstein metrics do not always exist. Surfaces which are $\mathbf{CP}^2$ with one or two points blown-up are examples. A well-known conjecture is the Yau-Tian-Donaldson conjecture [65, 55, 22] which says that the existence of a Kähler-Einstein metric with positive $c_1(X)$ should be equivalent to the K-stability of the manifold X. Very recently, a solution has been announced in the two series of independent works [10–12] and [56]. The works [9–12] rely in an essential manner on many advances almost as recent, including the uniform lower bounds for the Bergman kernel established in [25], the simplification of the notion of K-stability due to [34], and particularly the deep analogues of Brunn-Minkowski inequalities in complex geometry obtained in [5]. The special case of toric varieties was settled earlier in [63].

When $c_1(X)$ is definite, the Kähler-Ricci flow exists for all time [7]. So the main problem is to determine its asymptotic behavior for $t = \infty$. The situation here is somewhat similar now to the situation for the Yang-Mills flow for holomorphic bundles: algebraic-geometric conditions characterizing the existence of fixed points are known, and the task is to show the convergence of the flow when the conditions are satisfied, or to describe the resulting singularities when they are not. When $c_1(X) = 0$ or $c_1(X) < 0$, the convergence of the Kähler-Ricci flow was also established in [7]. But for $c_1(X) > 0$, the convergence of the Kähler-Ricci flow has been established only under the assumption that a Kähler-Einstein metric exists [36, 57, 58, 38, 17]. It is still not known how to establish directly its convergence from stability conditions, unlike in the case of the Yang-Mills flow. The general behavior of the flow under no stability conditions is even less understood. We explain some of these issues below (see e.g. [37, 42, 59] and references therein).

It is not difficult to adapt the estimates of Yau [64] for the Monge-Ampère equation to the parabolic case, and obtain the following criterion for the convergence of the Kähler-Ricci flow: if there exists a constant C so that

$$\sup_X |\varphi| \leq C \tag{4.10}$$

for all time $t > 0$, then the Kähler-Ricci flow converges (in fact, exponentially [41]). This explains why the case $c_1(X) < 0$ is easy: in this case, at a maximum point (z_0, t_0) for φ on any range $X \times [0, T]$ for finite T, the derivative $\dot{\varphi}(z_0, t_0)$ must be non-negative and the ratio of determinants ω^n / ω_0^n must be less than 1. Since

$$\varphi = \log \frac{\omega^n}{\omega_0^n} - f - \dot{\varphi} \tag{4.11}$$

it follows that $\sup_{X\times[0,T]}\varphi = \varphi(z_0, t_0) \leq \|f\|_{C^0(X)}$. A lower bound for $\inf_{X\times[0,T]}\varphi$ can be obtained in the same way, and since both bounds are independent of T, the desired estimate (4.10) follows. In the case $c_1(X) = 0$, it is readily seen that $\dot{\varphi}$ satisfies the heat equation $(\partial_t - \Delta)\dot{\varphi} = 0$, with initial value $\dot{\varphi}(z, 0) = -f(z)$. The same maximum principle argument shows that $\|\dot{\varphi}\|_{C^0(X\times[0,T])} \leq \|f\|_{C^0(X)}$ for any T. Incorporating $\dot{\varphi}$ into f, the parabolic Monge-Ampère equation can be treated as a standard Monge-Ampère equation, and the desired C^0 estimate follows by Moser iteration, as in [64].

To investigate the difficult case $c_1(X) > 0$, it is useful to refine the criteria for convergence as the following seemingly weaker estimate,

$$\frac{1}{V}\int_X \varphi\, \omega_0^n \leq C \tag{4.12}$$

for all time t, where it is convenient to include the volume V (fixed) of X with respect to ω_0. To control the left hand side, we introduce several useful functionals in Kähler geometry,

$$K_{\omega_0}(\varphi) = \frac{1}{V}\Big\{\int_X (\log \frac{\omega^n}{\omega_0^n})\omega^n - \varphi \sum_{j=0}^{n-1} Ric(\omega)\omega^j \omega_0^{n-1-j} + \mu \sum_{j=0}^{n-1} \varphi \omega^j \omega_0^{n-j}\Big\}$$

$$F_{\omega_0}(\varphi) = J_{\omega_0}(\varphi) - \frac{1}{V}\int_X \varphi \omega_0^n \tag{4.13}$$

where $J_{\omega_0}(\varphi)$ is defined in turn by

$$J_{\omega_0}(\varphi) = \frac{1}{2V}\sum_{j=0}^{n-1} \frac{n-j}{n+1}\int_x i\partial\varphi \wedge \bar{\partial}\varphi \wedge \omega^{n-1-j}\omega_0^j \tag{4.14}$$

The functional $K_{\omega_0}(\varphi)$ is called the Mabuchi K-functional, and the functional $F_{\omega_0}(\varphi)$ is called the Aubin-Yau functional. They are of particular interest because their variations give respectively the constant scalar curvature and the Monge-Ampère equations,

$$\delta K_{\omega_0}(\varphi) = -\frac{1}{V}\int_X \delta\varphi\,(R-\bar{R})\omega_\varphi^n$$
$$\delta F_{\omega_0}(\varphi)^0 = -\frac{1}{V}\int_X \delta\varphi\,\omega_\varphi^n. \tag{4.15}$$

Here $\bar{R}$ is the average of $R(z)$, which is a cohomological invariant. In particular, the following functional

$$F_{\omega_0}(\varphi) = F_{\omega_0}(\varphi)^0 - \log\big(\frac{1}{V}\int_X e^{f-\varphi}\omega_0^n\big) \tag{4.16}$$

is the functional whose Euler-Lagrange equation corresponds to the fixed points of the Kähler-Ricci flow (4.9). Returning to the Kähler-Ricci flow proper, one can show that the functionals $K_{\omega_0}(\varphi)$ and $F_{\omega_0}(\varphi)$ decrease along the flow, and that, also along the flow, they are all equivalent in sizes,

$$|F^0_{\omega_0}(\varphi) - F_{\omega_0}(\varphi)| + |F_{\omega_0}(\varphi) - K_{\omega_0}(\varphi)| \leq C. \tag{4.17}$$

Furthermore, the quantities $\frac{1}{V}\int_X \varphi\omega_0^n$, $J_{\omega_0}(\varphi)$ and $K_{\omega_0}(\varphi)$ satisfy the following inequalities,

$$\frac{1}{nV}\int_X(-\varphi)\omega_0^n \leq J_{\omega_0}(\varphi) \leq \frac{1}{V}\int_X \varphi\omega_0^n$$
$$\frac{1}{V}\int_X \varphi\omega_0^n \leq n\frac{1}{V}\int_X(-\varphi)\omega_0^n - (n+1)K_{\omega_0}(\varphi) + C \tag{4.18}$$

for some constants C (c.f. [38], Lemma 9).

With these preparations, it is now easy to prove the convergence of the Kähler-Ricci flow, if a Kähler metric is known to exist, and if X admits no holomorphic vector fields. Indeed, in this case, the following Moser-Trudinger inequality holds [55],

$$F_{\omega_0}(\varphi) \geq A_\gamma\, J_{\omega_0}(\varphi)^\gamma - B_\gamma, \tag{4.19}$$

for some strictly positive power γ, some positive constant A_γ, and all φ which are ω_0-plurisubharmonic (as shown in [40], γ can actually be taken to be 1). But then the fact that $F_{\omega_0}(\varphi)$ decreases along the Kähler-Ricci flow shows that it is bounded from above. By the Moser-Trudinger inequality, both $F_{\omega_0}(\varphi)$ and $J_{\omega_0}(\varphi)$ are bounded. By the inequalities (4.17), $K_{\omega_0}(\varphi)$ is bounded. And by the inequalities (4.18), both $\int_X(-\varphi)\omega_0^n$ and $\int_X \varphi\omega_0^n$ are bounded. This establishes the convergence of the Kähler-Ricci flow in this case, a result first announced by Perelman [36] and proved in [57]. Our discussion here follows [38].

The convergence of the Käler-Ricci flow under the assumption that a Kähler-Einstein metric exists, but allowing for the presence of non-holomorphic vector fields on X is much more delicate. The reason is that no suitable version of the Moser-Trudinger inequality (4.19) is available in this case. Some important ideas have been introduced in [58], and a full proof has been obtained only rather recently [17].

Thus, despite the above progress, the original problem remains of a direct proof of the convergence of the Kähler-Ricci flow under an algebraic-geometric stability condition, or of describing the singularities of the flow in general. As we had stressed earlier, the situation here is much worse than for the Yang-Mills flow, since we do not have any analogue of the Uhlenbeck-Yau method for producing a destabilizing object from the failure of an a priori C^0 estimate. We discuss now some attempts in this direction, exploiting the new results of Perelman for the Kähler-Ricci flow.

In his monumental work [35], Perelman introduced an array of powerful tools for the investigation of the Ricci flow. Among these is the entropy functional

$$\mu(g_{ij}) = \inf{}_f \{\mathcal{W}(g,f); \frac{1}{V}\int_X e^{-f}\omega^n = 1\},$$
$$\mathcal{W}(g,f) = \frac{1}{V}\int_X (R + |\nabla f|^2 + f - n)e^{-f}\omega^n \tag{4.20}$$

Perelman showed that the entropy functional is monotone increasing along the Ricci flow. For the Kähler-Ricci flow, there is a particular important scalar function u, namely the Ricci potential, defined by

$$R_{\bar{k}j} - \mu g_{\bar{k}j} = \partial_j\partial_{\bar{k}}u, \qquad \frac{1}{V}\int_X e^{-u}\omega^n = 1. \tag{4.21}$$

Using the resulting lower bound for $\mu(g_{ij})$, and the choice of $f = u$, he was able to deduce that

$$\|u\|_{C^0} + \|\nabla u\|_{C^0} + \|\Delta u\|_{C^0} \le C \tag{4.22}$$

where C is a constant independent of time. Furthermore, the diameter of X is uniformly bounded, and we have the following no-collapse statement,

$$\mathrm{Vol}(B_r(x)) \ge c\, r^{2n} \tag{4.23}$$

for another strictly positive constant $c > 0$. Since then, further development of Perelman's methods have also led to a uniform bound for the Sobolev constant, and a uniform upper bound for the growth of $\mathrm{Vol}(B_r(x))$ [66,67].

We can now link the convergence of the Kähler-Ricci flow to a stability condition in the following manner. First, using Perelman's results, we can obtain another criterion for the convergence of the Kähler-Ricci flow; namely, the flow converges if and only if there exists constants $C, \mu > 0$ so that [42]

$$Y(t) \equiv \int_X |\nabla u|^2\omega^n \le C\, e^{-\mu t}. \tag{4.24}$$

Next, we apply a differential inequality which always holds for the Kähler-Ricci flow [37]

$$\dot{Y}(t) \le -\lambda_t Y(t) - 2\lambda_t \mathrm{Fut}(\pi_t(\nabla u)) - \int_X |\nabla u|^2 (R-u)\omega^n$$
$$- \int_X \nabla^j u \nabla^{\bar{k}} u (R_{\bar{k}j} - g_{\bar{k}j})\omega^n \tag{4.25}$$

Here π is the orthogonal projection onto the space of holomorphic vector fields, $\mathrm{Fut}(W) = \int_X W u\, \omega^n$ is the Futaki invariant for the vector field W. The key ingredient in the formula is the lowest eigenvalue λ_t of the Laplacian $\bar{\partial}^\dagger\bar{\partial}$ on vector fields, which is defined by

$$\lambda_t = \inf{}_{W \perp H^0(X,T^{1,0}(X))} \frac{\|\bar{\partial}W\|^2}{\|W\|^2} \tag{4.26}$$

Without the three additional terms on the right hand side of (4.25), it would be clear that a strictly positive uniform lower bound for the eigenvalue λ_t

$$\inf_{t>0} \lambda_t \geq c > 0. \tag{4.27}$$

would imply the desired exponential decay for $Y(t)$. Nevertheless, this exponential decay can still be established under the condition (4.27) combined with another condition, namely that the K-energy $K_{\omega_0}(\varphi)$ is bounded from below [41]. This additional condition has been weakened since to the condition that the Futaki invariant be 0 [68].

The condition (4.27) is a stability condition. For example, if it fails, and if there exists a family of diffeomorphisms F_t so that $F_t^*(g(t))$ converges, then the pull-backs $F_t^*(J)$ of the complex structure would converge to a complex structure J_∞ which is in the closure of the orbit of J under the group of diffeomorphisms, but which is not equivalent to J. Thus, with such unstable points, the moduli space of complex structures on X would not be Hausdorff.

The eigenvalue λ_t provides one possibility for relating the convergence of the Kähler-Ricci flow to stability. It would be very interesting to relate the convergence even more directly to the properties of the orbits of complex structures under the group of diffeomorphisms, or to other notions of stability such as K-stability. As in the case of the Yang-Mills flow, a complementary direction is to investigate the behavior of the flow in general. A conjecture, often referred to as the Hamilton-Tian conjecture, states that, with suitable reparametrizations, the Kähler-Ricci flow should always converge to a Kähler-Ricci soliton, away from a subvariety of real codimension at least 4. This would be an analogue of the theorem of Uhlenbeck for instantons in dimension 4 [61], with the additional complication that the analogues of the gauge transformations for connections are here diffeomorphisms.

5 The Kähler-Ricci flow on manifolds of general type

In general, when the first Chern class $c_1(X)$ of a complex manifold is not definite, we cannot expect the existence of a smooth Kähler-Einstein metric. Nevertheless, because the Kähler-Ricci flow is arguably the most natural geometric flow on a Kähler manifold, it should provide a particularly deep probe of the underlying geometry of X. This is a very active research area at the present time, and we shall restrict ourselves to a few brief remarks, and defer the reader to the papers [44–48], and particularly the survey [49].

A first remark is that the Kähler-Ricci flow can still be reduced to a parabolic Monge-Ampère equation, albeit with an evolving background metric. More precisely, given a Kähler manifold (X, ω_0), and consider the flow

$$\dot{\omega}(t) = -Ric(\omega), \qquad \omega(0) = \omega_0. \tag{5.1}$$

Since the cohomology class of $Ric(\omega)$ is $c_1(X)$, it is clear that the cohomology class of $\omega(t)$ is $[\omega_0] - tc_1(X)$, so the flow can only exist as long as $[\omega_0] - tc_1(X)$ remains positive. Under this condition, we can choose a background Kähler form $\hat{\omega}(t)$ with $[\hat{\omega}(t)] = [\omega_0] - tc_1(X)$, and a volume form Ω with $\frac{i}{2}\partial\bar{\partial}\Omega = \partial_t \hat{\omega}(t)$. Setting $\omega(t) = \hat{\omega}(t) + \frac{i}{2}\partial\bar{\partial}\varphi$, we can rewrite the Kähler-Ricci flow as

$$\dot{\varphi} = \log \frac{(\hat{\omega}(t) + \frac{i}{2}\partial\bar{\partial}\varphi)^n}{\Omega}. \tag{5.2}$$

The above formulation of the flow shows that its behavior will depend very much on the nature of the limiting Kähler class as $t \to T \equiv \sup\{t; [\omega_0] - tc_1(X) > 0\}$. A very instructive example has been worked out in [46] for the flow on Hirzebruch surfaces, which are projectivizations of the direct sum of the trivial line bundle on $\mathbf{CP}^1$ with powers of the hyperplane bundle. Depending on the surface and on the initial Kähler class, the flow can exhibit very different behaviors: the fiber can collapse, or the manifold can shrink to a point, or the exceptional divisor can be contracted.

In general, one may hope that, as $t \to T$, the flow $(X, \omega(t))$ will converge in some sense to a possibly different manifold $(X_T, \omega(T))$, on which the flow can be continued. In this way, one would obtain a notion of weak solution allowing the continuation of the flow through singularities. This would be an analogue of the "Ricci flow with surgeries" initiated by Hamilton [28] and Perelman [35] in their approach to the Poincaré and the Geometrization conjectures. A detailed program has in fact been laid out by Song and Tian for the Kähler-Ricci flow on a projective algebraic variety [45]. According to this program, the Kähler-Ricci flow will either collapse in finite time, or deform the variety X to its minimal model through a finite number of divisorial contractions and flips, and finally converge to a generalized Kähler-Einstein metric on the canonical model of X. Thus the Kähler-Ricci flow would provide an analytic implementation of Mori's Minimal Model Program in algebraic geometry. The Song-Tian program is ambitious and challenging, but some important progress has already been made. For these progresses, many earlier estimates for complex Monge-Ampère equations had to be refined, and in particular, full use had to be made of the new powerful C^0 estimates due to Kolodziej [33] for Monge-Ampère equations with L^p right hand side, $p > 1$.

6 The Calabi flow

In this section, we consider the flows associated with a problem which is even more general than the preceding problem of existence of Kähler-Einstein metrics, namely: let $[\omega_0]$ be a positive Kähler class on a compact complex manifold X. When does there exist a Kähler form $\omega \in [\omega_0]$ with constant scalar curvature?

We note that the constant value $\bar{R}$ of the scalar curvature is again determined by the cohomology class. When $c_1(X)$ is definite, we can recover the Kähler-Einstein problem as follows. As in the Kähler-Ricci flow, let ω_0 be in any Kähler class if $c_1(X) = 0$, or $\omega_0 \in \pm c_1(X)$ id $c_1(X)$ is positive or negative. Then in all three cases, $Ric(\omega) - \mu\omega$ is cohomologically trivial, for $\mu = 0, 1, -1$ respectively. We can write

$$Ric(\omega) - \mu\omega = \frac{i}{2}\partial\bar{\partial} f \tag{6.1}$$

for some smooth scalar function f. This implies $R(\omega) - \mu n = \Delta f$. If $R(\omega)$ is constant, then Δf is constant. Since the range of the Laplacian is orthogonal to constants, the constant must be 0, and f itself is constant. But this implies that $Ric(\omega) - \mu\omega = 0$, which means that ω is actually Kähler-Einstein.

In general, it is difficult to prescribe a metric by curvature conditions, because of possible mismatches between the numbers of degrees of freedom: the metric $g_{\bar{j}}$ is a symmetric two-tensor, while e.g. the full curvature tensor is a four-tensor with very specific symmetries. The above constant scalar curvature problem in Kähler geometry is then particularly natural, since the requirement that the Kähler form ω be in a given class restricts it to a single degree of freedom, $\omega = \omega_0 + \frac{i}{2}\partial\bar{\partial}\varphi$. In view of the expression

(2.4) for the Ricci curvature, the equation for constant scalar curvature can be written explicitly as

$$-g^{j\bar{k}}\partial_j\partial_{\bar{k}} \log \det((g_0)_{\bar{q}p} + \partial_p\partial_{\bar{q}}\varphi) = \bar{R} \tag{6.2}$$

We note that this is a 4th-order non-linear partial differential equation, which combines the complex Monge-Ampère equation with its linearization. The Kähler-Einstein equation happens to be a simpler equation, because the operator $\partial\bar{\partial}$ can be integrated out in that case, resulting in a simpler 2nd-order equation. The same conjecture of Yau-Tian-Donaldson for Kähler-Einstein metrics is expected to hold for constant scalar curvature metrics in a given Kähler class: such a metric should exist if and only if the class $[\omega_0]$ is K-stable [22]. At the present time, it is known that the K-stability condition is necessary [50] (at least when $\mathrm{Aut}(X, L)$ is discrete), but its sufficiency has only been proved in the case of two-dimensional toric varieties [23].

Metrics of constant scalar curvature are the fixed points of the following flow, called the Calabi flow,

$$\dot{g}_{\bar{k}j} = \partial_j\partial_{\bar{k}}R, \qquad g_{\bar{k}j}(0) = g^0_{\bar{k}j}. \tag{6.3}$$

In terms of Kähler potentials, the flow is equivalent to

$$\dot{\varphi} = R - \bar{R}. \tag{6.4}$$

which shows that it is the gradient flow of the Mabuchi K-energy (see (4.14, 4.15)). In particular,

$$\frac{d}{dt}K_{\omega_0}(\varphi) = -\int_X |R - \bar{R}|^2\omega^n \equiv -C_{\omega_0}(\varphi). \tag{6.5}$$

The right hand side is known as (the negative of) the Calabi energy. A direct calculation shows that it also decreases along the Calabi flow,

$$\frac{d}{dt}C_{\omega_0}(\varphi) = -2\int_X |\nabla_{\bar{k}}\nabla_{\bar{j}}R|^2\omega^n \leq 0. \tag{6.6}$$

This formula shows that, besides metrics of constant scalar curvature which are its minima, the Calabi energy may admit more critical points. They are defined by the equation

$$\nabla_{\bar{k}}\nabla_{\bar{j}}R = 0, \tag{6.7}$$

which is equivalent to the condition that the gradient $\nabla^p R$ be a holomorphic vector field. Such metrics are called "extremal metrics".

Perhaps not surprisingly, much less is known about the Calabi flow than about any of the other flows. We discuss briefly some of what is known.

In one complex dimension, the Calabi flow is known to exist for all time and to converge to a metric of constant curvature [8, 51, 15]. Following Struwe [51], one can see this in the following way. In one complex dimension, exploiting the fact that any metric g_{ij} is Weyl equivalent to a metric of constant curvature $g^0_{ij}(x)$, we can write $g_{ij}(x) = e^{2u(x)}g^0_{ij}(x)$. Expressed in terms of $u(x)$, the Calabi flow becomes

$$\dot{u} = \frac{1}{2}\Delta R = \frac{1}{2}e^{-2u}\Delta_0(-\Delta_0 u + R_0 e^{-2u}). \tag{6.8}$$

The key step is the following "concentration/compactness" general statement about any sequence g_ℓ of metrics on a given compact surface X:

Let $g_\ell = e^{2u_\ell(x)}g_0$ be a sequence on X with fixed volume and uniformly bounded Calabi energy $C(g_\ell)$. Then

(a) either $\|u_\ell\|_{(2)} \leq C$, where $\|\cdot\|_{(2)}$ denotes the Sobolev norm of order 2 with respect to the fixed metric g^0;

(b) or there exists points $x_1, \cdots, x_M$ and a subsequence, still denoted by $\{u_\ell\}$, such that for any $r > 0$ and any m, $1 \leq m \leq M$, we have

$$\mathrm{liminf}_{\ell\to+\infty} \int_{B_r(x_m)} |R_\ell|\omega_\ell \geq 2\pi. \tag{6.9}$$

The number M can be bounded in terms of the maximum of the Calabi energy of g_ℓ. Furthermore, away from the points $x_1, \cdots, x_M$, either $u_\ell \to -\infty$, or they are uniformly bounded in the Sobolev norm of order 2.

Returning to the Calabi flow, we observe now that the Liouville energy [8]

$$L(u) = \frac{1}{2}\int_X (|\nabla u|_0^2 + 2R_0 u)\omega_0. \tag{6.10}$$

decreases along the flow, and thus it is bounded from above. We can now sketch the proof of the easier case, but geometrically more interesting, when the genus of the surface is ≥ 1, and $R_0 \leq 0$. First note that the area of ω is constant along the flow. This is because

$$\frac{d}{dt}(\int_X \omega) = \int_X \varDelta R\,\omega = 0. \tag{6.11}$$

Normalizing ω so that the area is 1, we obtain

$$2\int_X u\omega_0 \leq \log(\int_X e^{2u}\omega_0) = 0. \tag{6.12}$$

The upper bound for the Liouville energy implies now that $\|\nabla u\|_{L^2(\omega_0)}$ is bounded from above. The Trudinger inequality implies now

$$\int_X e^{4|u|}\omega_0 \leq C. \tag{6.13}$$

for some constant C. The proof when X is topologically a sphere is more delicate, because of the presence of holomorphic vector fields. It requires both a transformation by suitable Moebius transformations and a refinement of the Moser-Trudinger inequality due to Aubin [2].

The estimate for the L^1 norm of $e^{4|u|}$ combined with the uniform upper bound for the Calabi energy suffices to rule out concentration. Indeed,

$$\int_{B_r(x)} |R|\omega \leq (\int_{B_r(x)} |R|^2\omega)^{\frac{1}{2}}(\int_{B_r(x)} \omega)^{\frac{1}{2}} \tag{6.14}$$

and the second factor can be estimated in turn by

$$\int_{B_r(x)} \omega = \int_{B_r(x)} e^{2u}\omega_0 \leq (\int_{B_r(x)} e^{4u}\omega_0)^{\frac{1}{2}}(\int_{B_r(x)} \omega_0)^{\frac{1}{2}} \leq C(\int_{B_r(x)} \omega_0)^{\frac{1}{2}}. \tag{6.15}$$

The right hand side goes to 0 as $r \to 0$, ruling out concentration.

Since concentration has been ruled out, we can conclude from the above concentration/compactness dichotomy that the norms $\|u_\ell\|_{(2)}$ are uniformly bounded, for any sequence $u_{\ll} = u(t_\ell)$, $t_\ell \to +\infty$. Since the real dimension of the manifold is 2, the Sobolev imbedding theorem implies that the norms $\|u_\ell\|_{C^0}$ are uniformly bounded. Thus the metrics $g_\ell = e^{2u_\ell} g_0$ are all uniformly equivalent. This was the main difficulty. From here, it is not difficult to deduce the existence of a converging subsequence, from which the whole sequence converges, since it is a gradient flow. Working harder, one can show that the Calabi energy decays exponentially, and the convergence is exponential.

In higher dimensions, no general existence or convergence result for the Calabi flow is available at the present time. In fact, even the preceding one-dimensional result, as important as it is in itself, sheds only a limited light on what the higher dimensional case may be like. This is because it relied on a conformal representation of the evolving metrics instead of a representation as metrics in the same Kähler class. It also made essential use of the fact that a metric of constant scalar curvature already exists. There are on the other hand a wealth of existence and convergence results in higher dimensions, but under various additional assumptions. It is not possible to even try and cite them all here, but here are some illustrative few: existence of the flow under boundedness assumptions on the Ricci or Riemannian curvatures [13,52]; convergence of the flow under assumptions including smallness of the Calabi energy [14,60]; and behavior of the flow on specific classes of Kähler manifolds [53,54]. The work in [53] is of particular interest, as it exhibits in the case of ruled surfaces the behavior that one would expect for the Calabi flow: if no extremal metric exists in the given class, then the Calabi flow would either collapse, or break up the manifold into pieces which admit complete extremal metrics.

References

1. Aubin, T.: Equations du type Monge-Ampère sur les variétés Kähleriennes compacts. Bull. Sc. Math. **102**, 119-121 (1976)
2. Aubin, T.: Meilleures constantes dans le theoreme d'inclusion de Sobolev et un theoreme de Fredholm non-lineaires pour la transformation conforme de la courbure scalaire. J. Functinal Anal. **32**, 148-174 (1972)
3. Bando, S., Siu, Y.T.: Stable sheaves and Einstein-Hermitian metrics. In Geometry and Analysis on Complex Manifolds. World Scientific, 39-50 (1994)
4. Belavin, A., Polyakov, A., Schwarz, A., Tyuptin, Y.: Pseudoparticle solutions of the Yang-Mills equations. Phys. Lett. B **59**, 85-88 (1975)
5. Berndtsson, B.: A Brunn-Minkowski type inequality for Fano manifolds and some uniqueness theorems in Kähler geometry. arXiv:1303.4975
6. Candelas, P., Horowitz, G., Strominger, A., Witten, E.: Vacuum configurations for superstrings. Nucl. Phys. **B 258**, 46-74 (1985)
7. Cao, H.D.: Deformation of Kähler metrics to Kähler-Einstein metrics on compact Kähler manifolds. Invent. Math. **81**(2), 359-372 (1985)
8. Chen, X.X.: Calabi flow in Riemann surfaces revisited: a new point of view. Internat. Math. Res. Notices. No. 6, 275-297 (2001)
9. Chen, X.X., Donaldson, S. Sun, S.: Kähler-Einstein metrics and stability. arXiv:1210.7494
10. Chen, X.X., Donaldson, S. Sun, S.: Kähler-Einstein metrics on Fano manifolds, I: approximation of metrics with cone singularities. arXiv:1211.4566
11. Chen, X.X., Donaldson, S. Sun, S.: Kähler-Einstein metrics on Fano manifolds, II: limits with cone angle less than 2 π. arXiv:1212.4714
12. Chen, X.X., Donaldson, S., Sun, S.: Kähler-Einstein metrics on Fano manifolds, III: limits as cone angle approaches 2 and completion of the main proof. arXiv:1302.0282
13. Chen, X.X., He, W.: On the Calabi flow. Amer. J. Math. **130**(2), 539-570 (2008)
14. Chen, X.X., He, W.: The Calabi flow on toric Fano surfaces. Math. Res. Lett. **17**(2), 231-241 (2010)

15. Chrusciel, P.: Semi-global existence and convergence of the Robinson-Trautman (2-dimensional Calabi) equation. Commun. Math. Phys. **137**, 289-313 (1991)
16. Collins, T., Jacob, A.: On the bubbling set of the Yang-Mills flow on a compact Kähler manifold. arXiv:1206.6790
17. Collins, T., Szekelyhidi, G.: The twisted Kähler-Ricci flow. arXiv:1207.5441
18. Daskalopoulos, G., Wentworth, R.: Convergence properties of the Yang-Mills flow on Kähler surfaces. J. Reine Angew. Math. **575**, 69-99 (2004)
19. Daskalopoulos, G., Wentworth, R.: On the blow-up set of the Yang-Mills flow on Kähler surfaces. Math. Z. **256**(2), 301-310 (2007)
20. Donaldson, S.K.: Infinite deteminants, stable bundles, and curvature. Duke Math. J. **54**, 231-247 (1987)
21. Donaldson, S.K.: Anti self dual Yang-Mills connections over complex algebraic surfaces and stable vector bundles. Proc. London Math. Soc. **50**,1-26 (1985)
22. Donaldson, S.K.: Scalar curvature and stability of toric varieties. J. Differential Geom. **59**, 289-349 (2002)
23. Donaldson, S.K.: Constant scalar curvature metrics on toric surfaces. arXiv: 0805.0128
24. Donaldson, S.K., Kronheimer, P.: The Geometry of Four-Manifolds. Oxford University Press (1990)
25. Donaldson, S.K., Sun, S.: Gromov-Hausdorff limits of Kähler manifolds and algebraic geometry. arXiv:1206.2609
26. Eells, J., Sampson, G.: Harmonic mappings into Riemannian manifolds. Amer. J. Math. **86**, 109-160 (1964)
27. Hamilton, R.S.: Three manifolds with positive Ricci curvature. J. Differential Geom. **17**, 255-306 (1982)
28. Hamilton, R.S.: Formation of singularities in the Ricci flow. Surveys in Differential Geom. **II**, 7-136 (1995)
29. Hong, M.C., Tian, G.: Asymptotical behavior of the Yang-Mills flow and singular Yang-Mills connections. Math. Ann. **330**(3), 441-472 (2004)
30. Ilmanen, T., Huisken, G.: The inverse mean curvature flow and the Riemannian Penrose inequality. J. Differential Geom. (2001)
31. Jacob, A.: The limit of the Yang-Mills flow on semi-stable bundles. arXiv:1104.4767
32. Jacob, A.: The Yang-Mills flow and the Atiyah-Bott formula for compact Kähler manifolds. arXiv:1109.1550.
33. Kolodziej, S.: The complex Monge-Ampère equation. Acta Math. **180**, 69-117 (1998)
34. Li, Q., Xu, C.: Special test configurations and K-stability of Fano varieties. arXiv:1111.5398
35. Perelman, G.: The entropy formula for the Ricci flow and its geometric applications. arXiv: math.DG/0211159
36. Perelman, G., unpublished announcement.
37. Phong, D.H., Sturm, J.: On stability and the convergence of the Kähler-Ricci flow. J. Differential Geom. **72**(1), 149-168 (2006)
38. Phong, D.H., Sturm, J.: Lectures on Stability and Constant Scalar Curvature. In Handbook of Geometric Analysis Vol III, ALM 14, Higher Education Press and International Press, Beijing-Boston (2009)
39. Phong, D.H., Sesum, N., Sturm, J.: Multiplier ideal sheaves and the Kähler-Ricci flow. Commun. Anal. Geom. **15**(3), 613-632 (2007)
40. Phong, D.H., Song, J., Sturm, J., Weinkove, B.: The Moser-Trudinger inequality on Kähler-Einstein manifolds. Amer. J. Math. **130**(4), 1067-1085 (2008)
41. Phong, D.H., Song, J., Sturm, J., Weinkove, B.: The Kähler-Ricci flow and the $\bar\partial$ operator on vector fields. J. Differential Geom. **81**(3), 631-647 (2009)
42. Phong, D.H., Song, J. Sturm, J., Weinkove, B.: On the convergence of the modified Kähler-Ricci flow and solitons. Comment. Math. Helv. **86**(1), 91-112 (2011)
43. Sibley, B.: Asymptotics of the Yang-Mills flow for holomorphic vector bundles over Kähler manifolds: the canonical structure of the limit. arXiv: 1206.5491
44. Song, J., Tian, G.: The Kähler-Ricci flow on surfaces of positive Kodaira dimension. Inventiones Math. **170**(3), 609-653
45. Song, J., Tian, G.: Canonical measures and Kähler-Ricci flow. Preprint, arXiv:0802.2570
46. Song, J., Weinkove, B.: The Kähler-Ricci flow on Hirzebruch surfaces. J. reine Ange. Math. arXiv: 0903.1900, to appear
47. Song, J., Weinkove, B.: Contracting exceptional divisors by the Kähler-Ricci flow. arXiv:1003.0718
48. Song, J., Weinkove, B.: Contracting exceptional divisors by the Kähler-Ricci flow II. arXiv:1102.1759

49. Song, J., Weinkove, B.: Lecture notes on the Kähler-Ricci flow. arXiv:1212.3653
50. Stoppa, J.: K-stability of constant scalar curvature Kähler manifolds. Advances in Mathematics **221**(4), 1397-1408 (2009)
51. Struwe, M.: Curvature flows on surfaces. Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **1**(2), 247-274 (2002)
52. Szekelyhidi, G.: Remark on the Calabi flow with bounded curvature. arXiv:1209.2649
53. Szekelyhidi, G.: The Calabi functional on a ruled surface. Ann. Sci. Ec. Norm. Sup?er. (4) **42**(5), 837-856 (2009)
54. Szekelyhidi, G.: Optimal test-configurations for toric varieties J. Differential Geom. **80**, 501-523 (2008)
55. Tian, G.: Kähler-Einstein metrics with positive scalar curvature. Invent. Math. **130**(1), 1- 37 (1997)
56. Tian, G.: K-stability and Kähler-Einstein metrics. arXiv:1211.4669
57. Tian, G., Zhu, X.H.: Convergence of Kähler-Ricci flow. J. Amer. Math. Soc. **20**(3), 675-699 (2007)
58. Tian, G., Zhu, X.H.: Convergence of the K?ahler-Ricci flow on Fano manifolds, II. arXiv:1102.4798
59. Tosatti, V.: Kähler-Ricci flow on stable manifols. J. Reine Angew. Math. **640**, 67-84 (2010)
60. Tosatti, V., Weinkove, .: The Calabi flow with small initial energy. Math. Res. Lett. **14**(6), 1033-1039 (2007)
61. Uhlenbeck, K.: Removable singularities in Yang-Mills fields. Comm. Math. Phys. **83**, 11-29 (1982)
62. Uhlenbeck, K., Yau, S.T.: On the existence of Hermitian Yang-Mills connections on stable vector bundles. Commun. Pure Appl. Math. **39**, 257-293 (1986)
63. Wang, X.J., Zhu, X.H.: Kähler-Ricci solitons on toric manifolds with positive first Chern class. Advances Math. **188**, 87-103 (2004)
64. Yau, S.T.: On the Ricci curvature of a compact Kähler manifold and the complex Monge-Amp'ere equation, I. Comm. Pure Appl. Math. **31**, 339 - 411 (1978)
65. Yau, S.T.: Open problems in geometry. Proc. Symposia Pure Math. **54**, 1-28 (1993)
66. Zhang, Q.: A uniform Sobolev Inequality under the Ricci flow. Inter. Math. Res. Not. (2007)
67. Zhang, Q.: Bounds on volume growth of geodesic balls under Ricci flow. Math. Res. Lett. **19** (2012) arXiv: 1107.4262
68. Zhang, Z.: Kähler-Ricci flow on Fano manifolds with vanished Futaki invariants. Math. Res. Lett. **18** (2011) arXiv:1010.5959, 2010

PERFECTOID SPACES AND THE WEIGHT-MONODROMY CONJECTURE, AFTER PETER SCHOLZE

Takeshi Saito

1 Weight-monodromy conjecture

The Galois representation associated to the étale cohomology of a variety defined over a number field is a central subject of study in number theory. The Galois action on the Tate module of an elliptic curve and the Galois representation associated to a modular form that played the central role in the proof of Fermat's last theorem by Wiles and Taylor are typical examples.

To simplify the notation, we assume that a proper smooth variety X is defined over the rational number field $\mathbf{Q}$. We fix a prime number ℓ and consider the representation of the absolute Galois group $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ acting on the ℓ-adic étale cohomology $H^q(X_{\bar{\mathbf{Q}}}, \mathbf{Q}_\ell)$. As is seen in the definition of the Hasse-Weil L-function, a standard method in the study of a Galois representation is to investigate it locally at each prime.

Let p be a prime number different from ℓ. If X has good reduction at p, the Galois representation $H^q(X_{\bar{\mathbf{Q}}}, \mathbf{Q}_\ell)$ at p is almost completely understood thanks to the Weil conjecture proved by Deligne [1]. Namely its restriction to the decomposition group $\mathrm{Gal}(\bar{\mathbf{Q}}_p/\mathbf{Q}_p)$ at p is unramified and the characteristic polynomial $\det(1 - F_p t : H^q(X_{\bar{\mathbf{Q}}}, \mathbf{Q}_\ell))$ of the geometric Frobenius is determined by counting the number of points of the reduction of X modulo p defined over $\mathbf{F}_{p^n}$ for every $n \geqq 1$.

However, for a prime of bad reduction, an important piece, called the weight-monodromy conjecture, is still missing. To state it, let us recall briefly the structure of the absolute Galois group $\mathrm{Gal}(\bar{\mathbf{Q}}_p/\mathbf{Q}_p)$. Corresponding to the maximal unramified extension and the maximal tamely ramified extension $\mathbf{Q}_p \subset \mathbf{Q}_p^{\mathrm{ur}} = \mathbf{Q}_p(\zeta_m; p \nmid m) \subset \mathbf{Q}_p^{\mathrm{tr}} = \mathbf{Q}_p^{\mathrm{ur}}(p^{1/m}; p \nmid m) \subset \bar{\mathbf{Q}}_p$, the inertia subgroup and its wild part $\mathrm{Gal}(\bar{\mathbf{Q}}_p/\mathbf{Q}_p) \supset I \supset P \supset 1$ are defined. The quotient $\mathrm{Gal}(\bar{\mathbf{Q}}_p/\mathbf{Q}_p)/I$ is canonically identified with $\mathrm{Gal}(\bar{\mathbf{F}}_p/\mathbf{F}_p)$ and is topologically generated by the geometric Frobenius F_p. The quotient I/P by the pro-p Sylow subgroup P is non-canonically identified with the product of $\mathbf{Z}_{p'}$ for $p' \neq p$.

By the monodromy theorem of Grothendieck, there exists a nilpotent operator N on $V = H^q(X_{\bar{\mathbf{Q}}}, \mathbf{Q}_\ell)$ such that the restriction to an open subgroup of I is given by $\exp(t_\ell(\sigma)N)$ where $t_\ell \colon I \to \mathbf{Z}_\ell$ is a surjection. By elementary linear algebra, the nilpo-

T. Saito
University of Tokyo, Japan
E-mail: t-saito@ms.u-tokyo.ac.jp

tent operator defines a unique finite increasing filtration W on V characterized by the property that $N(W_iV) \subset W_{i-2}V$ for every integer i and N^i induces an isomorphism $\mathrm{Gr}_i^W V = W_iV/W_{i-1}V \to \mathrm{Gr}_{-i}^W V$ for every $i \geqq 0$. Then the weight-monodromy conjecture is stated as follows.

Conjecture 1.1 Let $F \in \mathrm{Gal}(\bar{\mathbf{Q}}_p/\mathbf{Q}_p)$ be a lifting of the geometric Frobenius F_p. Then the eigenvalues of F acting on $\mathrm{Gr}_i^W H^q(X_{\bar{\mathbf{Q}}}, \mathbf{Q}_\ell)$ are algebraic integers and the complex absolute values of the conjugates are $p^{(q+i)/2}$.

Conjecture was proved for $q = 1$ by Grothendieck by studying the Néron model of an abelian variety. It is proved for $q = 2$ by Rapoport-Zink using the weight spectral sequence in a semi-stable case and in general by using alteration by de Jong. For an arbitrary q, the Weil conjecture and alteration by de Jong imply that the eigenvalues of F acting on $\mathrm{Gr}_i^W V$ are algebraic integers and the complex absolute values of the conjugates are $p^{n/2}$ for an integer n independent of the conjugates. The function field analogue is proved by Deligne in the course of the proof of the Weil conjecture in [1].

P. Scholze introduced a new method in [7] to study the weight-monodromy conjecture and proved it for smooth complete intersections in smooth toric varieties, stated later as Theorem 6.2. The method is to construct and study the following diagram and reduce it to the function field case already proved by Deligne. In the diagram, k, k', K and $K^\flat$ denote finite extensions of $\mathbf{Q}_p, \mathbf{F}_p((t)), \mathbf{Q}_p(\zeta_{p^\infty})$ and of $\mathbf{F}_p((t))(t^{1/p^\infty})$ respectively.

$$\begin{array}{ccc} (\text{algebraic varieties}/k) & & \left(\text{algebraic varieties}/k'\right) \\ {\scriptstyle ?}\downarrow & & \downarrow{\scriptstyle \text{perfection}} \\ (\text{perfectoid spaces}/K) & \xrightarrow{\simeq} & \left(\text{perfectoid spaces}/K^\flat\right). \end{array} \tag{1.1}$$

The upper categories stay in the realm of algebraic geometry while the lower ones are in that of rigid geometry in the sense of R. Huber [6]. A point is that we have a canonical equivalence of categories on the lower line. Another point is that we have a canonical isomorphism

$$\mathrm{Gal}(\bar{K}/K) \leftarrow \mathrm{Gal}(\bar{K}^\flat/K^\flat). \tag{1.2}$$

Since some important special cases were first introduced by Fontaine-Wintenberger [4], the isomorphism (1.2) has been used effectively in the study of p-adic Galois representations of p-adic fields. It allows us to expect to deduce results in characteristic 0 from the corresponding results in characteristic $p > 0$.

A problem is that we do not have a natural functor for the left vertical arrow in (1.1) with ?, in general. Another problem is that we may lose some information by going down. The morphisms $\mathrm{Gal}(\bar{k}/k) \leftarrow \mathrm{Gal}(\bar{K}/K)$ and $\mathrm{Gal}(\bar{k}'/k') \leftarrow \mathrm{Gal}(\bar{K}^\flat/K^\flat)$ defined by inclusions $k \subset K$ and $k' \subset K^\flat$ induce isomorphisms on the quotients by the wild inertia subgroups. Since the weight-monodromy conjecture can be regarded as a statement on the actions of these quotients, we do not lose too much information by going down.

The purpose of the lecture is to explain the diagram and sketch the proof of the weight-monodromy conjecture for smooth complete intersections in toric varieties. For these varieties, one can construct the left vertical arrow with ? and recover enough information to prove the weight-monodromy conjecture.

The contents of the notes are summarized as follows. In Section 2, we introduce a perfectoid field and the tilting functor associating to a perfectoid field of characteristic

0 a perfectoid field of characteristic $p > 0$. The canonical isomorphism (1.2) is obtained as a special case of the almost purity theorem discussed in Section 5.

We also introduce perfectoid algebras in Section 2. The perfectoid spaces are defined by patching their associated adic spaces in Section 4, after recalling briefly some foundations on adic spaces. The key fact that the categories of perfectoid algebras over a perfectoid field K and its tilt $K^\flat$ are equivalent to each other is proved using the language of almost commutative algebra recalled in Section 3. The idea of the proof of the equivalence will be sketched at the end of the section.

In Section 5, we recall a crucial generalization of the almost purity theorem of Faltings and compare the étale topology in characteristic $p > 0$ and characteristic 0. Finally, we state the main result on the weight-monodromy conjecture and sketch the proof in Section 6.

For the full detail of proof, we refer to the original article [7]. A survey [3] is written by Fontaine.

2 Perfectoid fields and perfectoid algebras

We begin with recalling the definition of perfectoid fields and the tilting construction. We also state an isomorphism of Galois groups of a perfectoid field of characteristic 0 and its associated perfectoid field of characteristic $p > 0$, whose sketch of proof is postponed to Section 5.

Let K be a field. A mapping $v\colon K \to \mathbf{R} \cup \{\infty\} = (-\infty, \infty]$ is called a(n additive) valuation of K (of height 1) if $v(a+b) \geqq \min(v(a), v(b))$ and $v(ab) = v(a) + v(b)$ for $a, b \in K$, if $v(a) = \infty$ is equivalent to $a = 0$ and if $v(K) \supsetneqq \{0, 1\}$. A field K equipped with a valuation of height 1 will be called a valuation field. The subring $\mathcal{O}_K = \{a \in K \mid v(a) \geqq 0\}$ is called the valuation ring of $\mathcal{O}_K$ and $\mathfrak{m} = \{a \in K \mid v(a) > 0\}$ is the maximal ideal of $\mathcal{O}_K$. Choosing a real number $0 < a < 1$, we define a metric $d(x, y) = a^{v(x-y)}$ on K. The topology is independent of the choice of a.

A valuation field K of characteristic 0 is said to be of mixed characteristic $(0, p)$ if the residue field $\mathcal{O}_K/\mathfrak{m}$ is of characteristic $p > 0$.

Definition 2.1 Let K be a valuation field such that the restriction $v|_{K^\times} : K^\times \to \mathbf{R}$ has dense image and assume that K is either of characteristic $p > 0$ or of mixed characteristic $(0, p)$. Then K is called a *perfectoid field* if K is complete and if the Frobenius endomorphism $\mathcal{O}_K/p\mathcal{O}_K \to \mathcal{O}_K/p\mathcal{O}_K : x \mapsto x^p$ is a surjection.

A perfectoid field K of characteristic $p > 0$ is a perfect field and its valuation ring $\mathcal{O}_K$ is a perfect ring. Let K be a perfectoid field of mixed characteristic $(0, p)$. Then the inverse limit $\varprojlim \mathcal{O}_K/p\mathcal{O}_K$ with respect to the Frobenius endomorphism is a complete valuation ring and its fraction field $K^\flat$ is a perfectoid field of characteristic p, called the *tilt* of K. For $a = (a_n) \in \mathcal{O}_{K^\flat}$, the limit $a^\sharp = \lim_{n\to\infty} (\tilde{a}_n)^{p^n} \in \mathcal{O}_K$, defined by taking liftings, is independent of the choices and induces a mapping $K^\flat \to K$ compatible with the multiplications and the valuations.

Example 2.2 The p-adic completion K of $\mathbf{Q}_p(\zeta_{p^\infty})$ is a perfectoid field of mixed characteristic $(0, p)$. Its tilt $K^\flat$ is isomorphic to the t-adic completion of $\mathbf{F}_p((t))\ (t^{1/p^\infty})$.

The following fact is fundamental.

Proposition 2.3 [7, Theorem 3.7] *Let K be a perfectoid field of mixed characteristic* $(0, p)$. *Then, a finite extension L of K is also a perfectoid field and its tilt $L^\flat$ is a finite (separable) extension of $K^\flat$. Further, the functor*

$$\text{(finite separable extensions of } K) \to \text{(finite separable extensions of } K^\flat) \tag{2.1}$$

sending L to $L^\flat$ is an equivalence of categories.

Proposition 2.3 is proved as a special case of Theorem 5.1 and makes an important step of the proof. The equivalence of categories (2.1) induces a canonical isomorphism (1.2) of absolute Galois groups. It gives a generalization of the theory of fields of norms by Fontaine-Wintenberger [4].

We define perfectoid algebras over a perfectoid field. We fix a perfectoid field K of characteristic $p > 0$ or of mixed characteristic $(0, p)$. Let $\mathcal{O}_K$ denote its valuation ring and we also fix a non-zero element ϖ satisfying $0 < v(\varpi) \leqq v(p)$. We take a real number $0 < a < 1$ and define a norm $|x| = a^{v(x)}$ on K. A K-algebra R is said to be a K-Banach algebra if it is complete and separated with respect to a norm $|\ |$ on the K-vector space R compatible with multiplication. An element x of R is said to be power-bounded if the subset $\{x^n \mid n \geqq 0\}$ is bounded with respect to a norm. Let A denote the subring of R consisting of power-bounded elements.

Definition 2.4 Let K be a perfectoid field of characteristic $p > 0$ or of mixed characteristic $(0, p)$. We say that a Banach K-algebra R is a *perfectoid K-algebra* if the subring $A \subset R$ consisting of power-bounded elements is bounded and if the Frobenius endomorphism $A/\varpi \to A/\varpi \colon x \mapsto x^p$ is surjective. A morphism of perfectoid K-algebras is a continuous morphism of K-algebras.

Similarly as for perfectoid fields, we define tilting construction. Let K be a perfectoid field, R be a perfectoid K-algebra and $A \subset R$ be the subring as above. Then, we set

$$A^\flat = \varprojlim A/\varpi A \tag{2.2}$$

with respect to the Frobenius endomorphism. It is naturally an $\mathcal{O}_{K^\flat}$-algebra. Thus $R^\flat = A^\flat \otimes_{\mathcal{O}_{K^\flat}} K^\flat$ is defined as a Banach $K^\flat$-algebra, called the *tilt* of R.

Lemma 2.5 [7, Proposition 5.9] *Let K be a perfectoid field of* characteristic $p > 0$ *and R be a Banach K-algebra such that the subring A consisting of power-bounded elements is open and bounded. Then, the following conditions are equivalent:*

(1) *R is a perfectoid K-algebra.*

(2) *R is perfect.*

(3) *A is perfect.*

Example 2.6 1. If k is a complete discrete valuation field of characteristic $p > 0$ with a uniformizer t, the fraction field K of the t-adic completion of the perfection $\mathcal{O}_k^{1/p^\infty}$ is a perfectoid field. If A is a flat $\mathcal{O}_k$-algebra of finite type, the t-adic completion of the perfection A^{1/p^∞} tensored K is a perfectoid algebra over K.

2. The ϖ-adic completion R of $K[T_1^{1/p^\infty}, \ldots, T_n^{1/p^\infty}]$ is a perfectoid K-algebra, denoted $K\langle T_1^{1/p^\infty}, \ldots, T_n^{1/p^\infty}\rangle$. Its tilt $R^\flat$ is $K^\flat\langle T_1^{1/p^\infty}, \ldots, T_n^{1/p^\infty}\rangle$.

Now we state a key result.

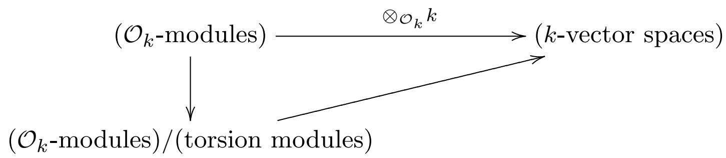
Theorem 2.7 [7, Theorem 5.2] *Let K be a perfectoid field of mixed characteristic* $(0, p)$. *Then, for a perfectoid K-algebra R, its tilt $R^\flat$ is a perfectoid $K^\flat$-algebra. Further the functor*

$$\text{(perfectoid } K\text{-algebras)} \to \text{(perfectoid } K^\flat\text{-algebras)} \tag{2.3}$$

sending R to $R^\flat$ is an equivalence of categories.

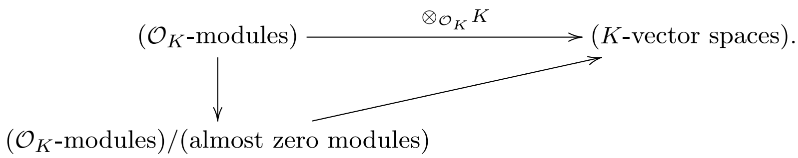
3 Almost commutative algebra

A basic idea on almost commutative algebra in the context of perfectoid extension of a complete discrete valuation ring is the following. Let k be a complete discrete valuation field and $\mathcal{O}_k$ be the valuation ring. We consider the factorization



of the scalar extension functor by the quotient category. Then the slant arrow is an equivalence of category.

Now, let K denote a perfectoid field and $\mathcal{O}_K$ be the valuation ring. Similarly, we consider the factorization



Here an $\mathcal{O}_K$-module is said to be almost zero if every element is annihilated by the maximal ideal $\mathfrak{m}_K$, satisfying $\mathfrak{m}_K = \mathfrak{m}_K^2$. This time, the vertical arrow is very close to an equivalence of categories. This is analogous to ignoring infinitesimal in classical calculus. Moreover, one can develop a theory of commutative algebra in the lower tensor abelian category, which is called *almost* commutative algebra.

We work in a more general setting. Let A be a commutative ring and let $\mathfrak{m}$ be a flat ideal of A satisfying $\mathfrak{m}^2 = \mathfrak{m}$. We say that an A-module M is *almost zero* if $\mathfrak{m}M = 0$. The essential image of the natural fully faithful functor

$$(A/\mathfrak{m}\text{-modules}) \to (A\text{-modules})$$

is the subcategory consisting of almost zero modules.

Lemma 3.1 *For an A-module M, the following conditions are equivalent.*
(1) *M is almost zero.*
(2) $Tor_q^A(\mathfrak{m}, M) = 0$ *for every* $q \geqq 0$.
(3) $Ext_A^q(\mathfrak{m}, M) = 0$ *for every* $q \geqq 0$.

Proof Since $\mathfrak{m}M$ is the image of the composition $\mathfrak{m} \otimes_A M \to \mathfrak{m} \otimes_A Hom_A(\mathfrak{m}, M) \to M$, either of the vanishings $\mathfrak{m} \otimes_A M = 0$ and $Hom_A(\mathfrak{m}, M) = 0$ implies $\mathfrak{m}M = 0$.

A free resolution of the A-module $\mathfrak{m}$ defines spectral sequences

$$E^2_{p,q} = Tor_p^{A/\mathfrak{m}}(Tor_q^A(\mathfrak{m}, A/\mathfrak{m}), M) \Rightarrow Tor_{p+q}^A(\mathfrak{m}, A/\mathfrak{m} \otimes_{A/\mathfrak{m}} M),$$

$$E_2^{p,q} = Ext^p_{A/\mathfrak{m}}(Tor^A_{-q}(\mathfrak{m}, A/\mathfrak{m}), M) \Rightarrow Ext_A^{p+q}(\mathfrak{m}, Hom_{A/\mathfrak{m}}(A/\mathfrak{m}, M))$$

for an $A/\mathfrak{m}$-module M. Since the A-module $\mathfrak{m}$ is assumed flat, they induce isomorphisms $Tor_p^{A/\mathfrak{m}}(\mathfrak{m} \otimes_A (A/\mathfrak{m}), M) \to Tor_p^A(\mathfrak{m}, M)$ and $Ext^p_{A/\mathfrak{m}}(\mathfrak{m} \otimes_A (A/\mathfrak{m}), M) \to Ext^p_A(\mathfrak{m}, M)$. Since $\mathfrak{m} \otimes_A (A/\mathfrak{m}) = \mathfrak{m}/\mathfrak{m}^2 = 0$, the assertion follows. □

Either of the conditions (2) and (3) implies that the subcategory consisting of almost zero modules is closed under extensions.

We say that a morphism of A-modules $f\colon M \to N$ is an *almost isomorphism* if the kernel and the cokernel of f are almost zero. For morphisms $f\colon L \to M,\ g\colon M \to N$ of A-modules, if two of f, g and $g \circ f$ are almost isomorphisms, so is the third.

Lemma 3.2 1. *For an A-module M, the canonical morphisms $\mathfrak{m}\otimes_A M \to M \to Hom_A(\mathfrak{m}, M)$ are almost isomorphisms.*

2. *For a morphism of A-modules $f\colon M \to N$, the following conditions are equivalent.*

(1) *f is an almost isomorphism.*

(2) *$f_*\colon \mathfrak{m} \otimes_A M \to \mathfrak{m} \otimes_A N$ is an isomorphism.*

(3) *$f_*\colon Hom_A(\mathfrak{m}, M) \to Hom_A(\mathfrak{m}, N)$ is an isomorphism.*

Proof 1. The exact sequence $0 \to \mathfrak{m} \to A \to A/\mathfrak{m} \to 0$ induces exact sequences $Tor_1^A(A/\mathfrak{m}, M) \to \mathfrak{m} \otimes_A M \to M \to (A/\mathfrak{m}) \otimes_A M \to 0$ and $Hom_A(A/\mathfrak{m}, M) \to M \to Hom_A(\mathfrak{m}, M) \to Ext^1_A(A/\mathfrak{m}, M)$ and the assertion follows.

2. Since (1) in Lemma 3.1 implies (2) and (3) in Lemma 3.1 respectively, (1) implies (2) and (3) respectively. Conversely, the left (resp. right) square of the commutative diagram

$$\begin{array}{ccccc} \mathfrak{m} \otimes_A M & \longrightarrow & M & \longrightarrow & Hom_A(\mathfrak{m}, M) \\ \downarrow & & \downarrow & & \downarrow \\ \mathfrak{m} \otimes_A N & \longrightarrow & N & \longrightarrow & Hom_A(\mathfrak{m}, N) \end{array}$$

shows that (2) (resp. (3)) implies (1) by 1. and the remark preceding Lemma. □

We define the category of *almost A-modules* to be the quotient category of that of A-modules by the subcategory consisting of almost zero A-modules. We have a canonical functor

$$(A\text{-modules}) \to (\text{almost } A\text{-modules}) \tag{3.1}$$

sending an A-module M to the associated almost A-module M^a. The category (almost A-modules) is an abelian category and inherits tensor products and internal Hom from the category (A-modules). Consequently, we can do linear algebra as well as multi-linear algebra and commutative algebra in the category.

For A-modules M and N, we call a morphism in the category of almost A-modules *an almost morphism* and let $Hom_{A^a}(M^a, N^a)$ denote the A-module of almost morphisms. We set ${}_*M = \mathfrak{m} \otimes_A M$ and $N_* = Hom_A(\mathfrak{m}, N)$.

Lemma 3.3 1. *The functors $M \mapsto {}_*M$ and $N \mapsto N_*$ are adjoint to each other. The canonical morphisms ${}_*M \to M$ and $N \to N_*$ induce isomorphisms of functors ${}_*({}_*M) \to {}_*M$ and $N_* \to (N_*)_*$.*

2. *The canonical functor* (3.1) *admits left and right adjoint functors*

$$(\text{almost } A\text{-modules}) \to (A\text{-modules}), \tag{3.2}$$

induced by the functors $M \mapsto {}_*M$ and $N \mapsto N_*$ respectively. For A-modules M and N, we have a canonical isomorphism $Hom_{A^a}(M^a, N^a) \to Hom_A(M, N)_*$.

3. *The functors* (3.2) *are fully faithful and the essential images consist of A-modules M such that the canonical morphism $M \to {}_*M$ is an isomorphism and N such that $N \to N_*$ is an isomorphism respectively.*

Proof 1. The canonical isomorphisms $Hom_A(\mathfrak{m} \otimes_A M, N) \to Hom_A(M, Hom_A\ (\mathfrak{m}, N))$ define an adjunction. Since the multiplication induces an isomorphism $\mathfrak{m} \otimes_A \mathfrak{m} \to \mathfrak{m}$, we obtain an isomorphism ${}_*({}_*M) = \mathfrak{m} \otimes_A \mathfrak{m} \otimes_A M \to\ \mathfrak{m} \otimes_A M =\ {}_*M$. The isomorphism $N_* \to (N_*)_*$ follows from this and the adjunction.

2. By Lemma 3.2, the functors $M \mapsto {}_*M$ and $N \mapsto N_*$ induce functors (3.2). By the definition of the quotient category and Lemma 3.2, the almost isomorphisms ${}_*M \to M$ and $N \to N_*$ induce an isomorphism $Hom_A({}_*M, N_*) \to Hom_{A^a}(M^a, N^a)$. By 1., we obtain canonical isomorphisms $Hom_A({}_*M, N_*) \to Hom_A({}_*({}_*M), N) \to Hom_A({}_*M, N)$ and $Hom_A({}_*M, N_*) \to Hom_A(M, (N_*)_*) \to Hom_A(M, N_*)$.

Further, we have a canonical isomorphism $Hom_A({}_*M, N) = Hom_A(\mathfrak{m} \otimes_A M, N) \to Hom_A(\mathfrak{m}, Hom_A(M, N)) = Hom_A(M, N)_*$.

3. Since $M \to M_*$ is an almost isomorphism and $N \mapsto N_*$ induces the right adjoint, we obtain canonical isomorphisms $Hom_{A^a}(M^a, N^a) \leftarrow Hom_{A^a}(M_*^a, N^a) \to Hom_A(M_*, N_*)$. The description of the essential image follows from the isomorphism $N \to (N_*)_*$.

The assertion for the functor ${}_*M$ is proved similarly. □

Let R be an A-algebra. We say that an R-module M is *almost locally free of finite rank* if the canonical map $M \otimes_R Hom_R(M, R) \to End_R(M)$ is an almost isomorphism. For an almost locally free R-module M of finite rank, the *trace map* is defined to be the composition

$$End_R(M) \to End_R(M)_* \overset{\sim}{\leftarrow} (M \otimes_R Hom_R(M, R))_* \to R_*,$$

where the last map is induced by the evaluation map $x \otimes f \mapsto f(x)$.

We say that a surjection $R \to S$ of commutative A-algebras *defines an almost open immersion* if S is an almost locally free R-module of finite rank. We say that a morphism $R \to S$ of commutative A-algebras is *almost finite étale* if S is an almost locally free R-module of finite rank and if the surjection $S \otimes_R S \to S$ defines an almost open immersion.

A commutative ring R over A is an object of the category of (A-modules) equipped with the multiplication $R \otimes_A R \to R$ satisfying a certain set of axioms. Similarly, one defines an almost commutative ring R over A as an object of the category of (almost A-modules) equipped with an almost multiplication $R \otimes_A R \to R$ satisfying the corresponding set of axioms. It is established in [5] that one obtains a completely parallel theory.

We will sketch the proof of the equivalence of categories (2.3). The functor $\otimes_{\mathcal{O}_K} K$: ($\mathcal{O}_K$-modules) $\to$ (K-vector spaces) induces a functor $\otimes_{\mathcal{O}_K^a} K$: (almost $\mathcal{O}_K$-modules) $\to$ (K-vector spaces). Let (Perf/K) and (Perf/$K^\flat$) denote the categories of perfectoid algebras. Theorem 2.7 is proved by constructing a diagram

$$(\mathrm{Perf}/K) \leftarrow (\mathrm{Perf}/\mathcal{O}_K^a) \to (\mathrm{Perf}/(\mathcal{O}_K/\varpi\mathcal{O}_K)^a) \leftarrow (\mathrm{Perf}/\mathcal{O}_K^{\flat a}) \to (\mathrm{Perf}/K^\flat) \qquad (3.3)$$

of equivalences of categories. First, we define the categories in the diagram (3.3). By abuse of notation, for an $\mathcal{O}_K^a$-algebra A and for a non-zero element $\varpi \in \mathfrak{m}_K$, let $A/\varpi^{1/p}A$ denote the quotient by the principal ideal generated by an element of valuation $v(\varpi)/p$. An almost $\mathcal{O}_K$-module M is said to be complete if the canonical morphism $M \to (\varprojlim_n M_*/\varpi^n M_*)^a$ is an isomorphism.

Definition 3.4 Let K be a perfectoid field of characteristic $p > 0$ or of mixed characteristic $(0, p)$. Let $\mathcal{O}_K$ be the valuation ring and ϖ be a non-zero element satisfying $0 < v(\varpi) \leqq v(p)$.

1. We say that a ϖ-adically complete flat $\mathcal{O}_K^a$-algebra A is a *perfectoid* $\mathcal{O}_K^a$*-algebra* if the morphism $A/\varpi^{1/p}A \to A/\varpi A\colon x \mapsto x^p$ is an isomorphism. A morphism of perfectoid $\mathcal{O}_K^a$-algebras is a morphism of $\mathcal{O}_K^a$-algebras.

2. We say that a flat $(\mathcal{O}_K/\varpi\mathcal{O}_K)^a$-algebra $\bar{A}$ is a *perfectoid* $(\mathcal{O}_K/\varpi\mathcal{O}_K)^a$*-algebra* if the morphism $\bar{A}/\varpi^{1/p}\bar{A} \to \bar{A}\colon x \mapsto x^p$ is an isomorphism. A morphism of perfectoid $(\mathcal{O}_K/\varpi\mathcal{O}_K)^a$-algebras is a morphism of $(\mathcal{O}_K/\varpi\mathcal{O}_K)^a$-algebras.

The following Lemma defines the arrows in (3.3) and shows that the first and the last ones are equivalences of categories.

Lemma 3.5 *Let K be a perfectoid field and $\mathcal{O}_K$ be the valuation ring.*

1.([7, Proposition 5.5]) *Let R be a perfectoid algebra over K and A be the subring consisting of power-bounded elements. Then, the Frobenius morphism induces an isomorphism $A/\varpi^{1/p}A \to A/\varpi A$ and A^a is a perfectoid algebra over $\mathcal{O}_K^a$.*

2.([7, Lemma 5.6]) *Let A be a perfectoid algebra over $\mathcal{O}_K^a$ and equip the K-algebra $R = A \otimes_{\mathcal{O}_K^a} K$ a Banach K-algebra structure such that $A_* \subset R$ is open and bounded. Then, R is a perfectoid algebra over K and $A_* \subset R$ is the subring consisting of power-bounded elements.*

To prove that the middle arrows are equivalences of categories, we need to find liftings of perfectoid $(\mathcal{O}_K/\varpi\mathcal{O}_K)^a$-algebras and their morphisms. This is done by using the theory of cotangent complexes adjusted to the context of almost commutative algebras developed in [5]. It is eventually reduced to that the cotangent complex of a perfect $\mathbf{F}_p$-algebra vanishes. One also needs to check that the composition of (3.3) is actually given by the construction (2.2).

4 Perfectoid spaces

So far, we have studied only local pieces. We make a global construction using the language of adic spaces in the sense of R. Huber [6]. A building block of an adic space is a locally ringed space $\mathrm{Spa}(R, R^+)$ defined for an affinoid k-algebra (R, R^+).

Let k denote a complete valuation field. We call a topological k-algebra R a *Tate k-algebra* if there exists a subring R_0 (over $\mathcal{O}_k$) such that aR_0 for $a \in k^\times$ form a basis of open neighborhoods of 0. For a Tate k-algebra R, let A denote the subring consisting of power-bounded elements. A pair (R, R^+) of a Tate k-algebra R and an open and integrally closed subring $R^+ \subset A$ (over $\mathcal{O}_k$) is called an *affinoid k-algebra*.

For an affinoid k-algebra (R, R^+), the underlying set of $\mathrm{Spa}(R, R^+)$ is defined as the set of (equivalence classes of) continuous valuations v satisfying $v(f) \geqq 0$ for $f \in R^+$. For a ring R and a totally ordered additive group Γ, a mapping $v\colon R \to \Gamma \cup \{\infty\}$ is called a(n additive) *valuation* if the following conditions are satisfied; $v(xy) = v(x) + v(y)$ and $v(x+y) \geqq \min(v(x), v(y))$ for $x, y \in R$, $v(0) = \infty$ and $v(1) = 0$. If R is a topological ring, a valuation v is said to be *continuous* if $v^{-1}((g, \infty]) \subset R$ is open for every $g \in \Gamma$ such that $g = v(x)$ for some $x \in R$.

For a valuation v, define the value group Γ_v to be the subgroup of Γ generated by $\{g \in \Gamma \mid g = v(x) \text{ for some } x \in R\}$ and the support as a prime ideal of R by $\mathfrak{p}_v = \{x \in R \mid v(x) = \infty\}$. Then, the valuation v induces a valuation of the fraction field of $R/\mathfrak{p}_v$. Valuations v and v' are said to be equivalent, if there exists an isomorphism of totally ordered groups $\Gamma_v \to \Gamma_{v'}$ compatible with v and v'.

For an affinoid k-algebra (R, R^+), we define the set $X = \mathrm{Spa}(R, R^+)$ to be the equivalence classes of continuous valuations v of R such that $v(R^+) \subset [0, \infty]$. We equip

X a topology with a basis consisting of *rational subsets* $U\Big(\dfrac{f_1,\dots,f_n}{g}\Big)=\{v\in X\mid v(f_i)\geqq v(g)$ for $i=1,\dots,n\}$ defined for $f_1,\dots,f_n,g\in R$ satisfying $(f_1,\dots,f_n)=R$.

We define the structure presheaf $\mathcal{O}_X$ on X. Let $f_1,\dots,f_n,g\in R$ be elements satisfying $(f_1,\dots,f_n)=R$. We define a topology on the ring $R\Big[\dfrac{f_1}{g},\dots,\dfrac{f_n}{g}\Big]=R\Big[\dfrac{1}{g}\Big]$ by a basis of open neighborhood of 0 consisting of $aR_0\Big[\dfrac{f_1}{g},\dots,\dfrac{f_n}{g}\Big]$ for $a\in k^\times$. Let $R\Big\langle\dfrac{f_1}{g},\dots,\dfrac{f_n}{g}\Big\rangle$ be the completion of $R\Big[\dfrac{f_1}{g},\dots,\dfrac{f_n}{g}\Big]$ and let $R^+\Big\langle\dfrac{f_1}{g},\dots,\dfrac{f_n}{g}\Big\rangle$ denote abusively the completion of the integral closure of $R^+\Big[\dfrac{f_1}{g},\dots,\dfrac{f_n}{g}\Big]$ in $R\Big[\dfrac{f_1}{g},\dots,\dfrac{f_n}{g}\Big]$. Then, the morphism $R\to R\Big\langle\dfrac{f_1}{g},\dots,\dfrac{f_n}{g}\Big\rangle$ induces a homeomorphism

$$\mathrm{Spa}\left(R\Big\langle\frac{f_1}{g},\dots,\frac{f_n}{g}\Big\rangle,R^+\Big\langle\frac{f_1}{g},\dots,\frac{f_n}{g}\Big\rangle\right)\to U\Big(\frac{f_1,\dots,f_n}{g}\Big)\subset X=\mathrm{Spa}(R,R^+).$$

Further, the topological rings $R\Big\langle\dfrac{f_1}{g},\dots,\dfrac{f_n}{g}\Big\rangle$ and $R^+\Big\langle\dfrac{f_1}{g},\dots,\dfrac{f_n}{g}\Big\rangle$ depend only on the rational subset $U\Big(\dfrac{f_1,\dots,f_n}{g}\Big)\subset X=\mathrm{Spa}(R,R^+)$.

We define a presheaf $\mathcal{O}_X$ on X by requiring $\mathcal{O}_X(U)=R\Big\langle\dfrac{f_1}{g},\dots,\dfrac{f_n}{g}\Big\rangle$ for rational subsets $U=U\Big(\dfrac{f_1,\dots,f_n}{g}\Big)$. For each point x of $X=\mathrm{Spa}(R,R^+)$, the (equivalence class of) continuous valuation of R induces a(n equivalence class of) continuous valuation of the local ring $\mathcal{O}_{X,x}$. We regard $X=\mathrm{Spa}(R,R^+)$ as a topological space equipped with the presheaf $\mathcal{O}_X$ of topological rings together with the (equivalence class of) continuous valuation of the local ring at each point.

If the presheaf $\mathcal{O}_X$ is a sheaf, we call $X=\mathrm{Spa}(R,R^+)$ equipped with these structures an *affinoid adic space*. An *adic space* X is defined to be a topological space equipped with a sheaf $\mathcal{O}_X$ of topological rings together with a (equivalence class of) continuous valuation of the local ring at each point, that is locally isomorphic to an affinoid adic space.

For a perfectoid affinoid algebra (R,R^+), the presheaf $\mathcal{O}_X$ on $X=\mathrm{Spa}(R,R^+)$ is a sheaf. Let K be a perfectoid field. We say that an affinoid K-algebra (R,R^+) is a *perfectoid* affinoid K-algebra if R is a perfectoid K-algebra. The tilting functor (2.3) induces an equivalence of categories

$$\text{(perfectoid affinoid } K\text{-algebras)}\to\text{(perfectoid affinoid } K^\flat\text{-algebras)}\tag{4.1}$$

sending (R,R^+) to $(R^\flat,R^{\flat+})$. Here $R^{\flat+}$ denotes the open and integrally closed subalgebra satisfying $\mathfrak{m}R^{\flat\circ}\subset R^{\flat+}\subset R^{\flat\circ}$ and corresponding to $\mathfrak{m}R^\circ\subset R^+\subset R^\circ$.

Theorem 4.1 [7, Theorem 6.3] *Let K be a perfectoid field, (R,R^+) be a perfectoid affinoid K-algebra and let $X=\mathrm{Spa}(R,R^+)$ be the associated topological space equipped with the structures as above.*

1. *The presheaf $\mathcal{O}_X$ is a sheaf.*
2. *Assume K is of mixed characteristic $(0,p)$ and let $K^\flat$ and $(R^\flat,R^{\flat+})$ be the tilts of K and (R,R^+). Then, there exists a unique homeomorphism $\flat\colon X\to X^\flat=\mathrm{Spa}(R^\flat,R^{\flat+})$ compatible with the construction of rational subsets U and isomorphisms $\mathcal{O}_X(U)^\flat\to\mathcal{O}_{X^\flat}(U^\flat)$.*

Theorem 4.1.1 is a part of an analogue for perfectoid algebras of Tate's acyclicity theorem in rigid geometry. Theorem 4.1.2 is proved by using the equivalence of categories established in Theorem 2.7 and some approximation property. Theorem 4.1.1 is reduced to the case of characteristic $p > 0$ by Theorem 4.1.2. In the latter case, it is proved by reducing eventually to Tate's acyclicity theorem in the classical case.

Theorem 4.1 enables us to define perfectoid spaces.

Definition 4.2 An adic space over a perfectoid field K is called a *perfectoid space* if it is locally isomorphic to $\mathrm{Spa}(R, R^+)$ for an perfectoid affinoid K-algebra (R, R^+).

Theorem 2.7 and Theorem 4.1 immediately imply an equivalence of categories

$$(\text{perfectoid spaces over } K) \to (\text{perfectoid spaces over } K^\flat) \tag{4.2}$$

attaching to a perfectoid space X over K its tilt $X^\flat$ over $K^\flat$. Further Theorem 4.1 implies a homeomorphism $X \to X^\flat$ compatible with tilting construction on the structure sheaves.

To prove (cases of) the weight-monodromy conjecture, we still need to understand the compatibility of étale topology with the equivalence of categories (4.2).

5 Almost purity theorem and étale topology

The most important point in the theory is the following generalization of the almost purity theorem of Faltings [2].

Theorem 5.1 [7, Theorem 7.9] *Let K be a perfectoid field, R be a perfectoid K-algebra and A be the perfectoid $\mathcal{O}_K^a$-algebra associated to the subring of R consisting of power-bounded elements. Then, a finite étale R-algebra S is a perfectoid K-algebra and the $\mathcal{O}_K^a$-algebra B associated to the subring of S consisting of power-bounded elements is a perfectoid $\mathcal{O}_K^a$-algebra and is almost finite étale over A.*

We sketch the proof. Let R be a perfectoid K-algebra and let A be the perfectoid $\mathcal{O}_K^a$-algebra associated to the subring consisting of power-bounded elements. Setting $\bar{A} = A/\varpi A$ and we consider the following diagram of categories

$$(\text{FÉt}/R) \leftarrow (\text{FÉt}/A) \to (\text{FÉt}/\bar{A}) \leftarrow (\text{FÉt}/A^\flat) \to (\text{FÉt}/R^\flat) \tag{5.1}$$

consisting of (almost) finite étale algebras, similar to the diagram (3.3). Using the theory of almost commutative algebra [5], one checks ([7, Theorem 4.17, Proposition 5.22]) that the middle two arrows are equivalences of categories, that a finite étale A-algebra is a perfectoid $\mathcal{O}_K^a$-algebra and that a finite étale $\bar{A}$-algebra is a perfectoid $(\mathcal{O}_K/\varpi\mathcal{O}_K)^a$-algebra, respectively. This implies that the middle three categories in (5.1) are subcategories of the corresponding categories in (3.3) and that the middle two arrows are compatible to each other.

Using Lemma 2.5 in characteristic $p > 0$, it is proved rather directly ([7, Proposition 5.23]) that the last arrow in (5.1) is an equivalence of categories and is compatible with that in (3.3). By what is already proven and by Theorem 2.7, the proof of Theorem 5.1 is reduced to showing that the composite functor from the right end to the left end of (5.1) is essentially surjective. This is proved first in the case where R is a field. Then, applying this to the residue field of each point and using that the local rings are henselian, we show that we obtain an etale covering locally. Then, by using the sheaf property, Theorem 4.1.1, we conclude the proof.

An isomorphism of étale sites follows directly from Theorem 5.1. First, we define the étale site. A morphism $X \to Y$ of adic spaces is said to be étale if locally on Y, there exists an open covering by affinoids $V = \mathrm{Spa}(R)$ and an almost finite étale R-algebra S such that $X \times_Y V \to V$ factors through an open immersion $X \times_Y V \to U = \mathrm{Spa}(S)$ over V.

Definition 5.2 Let X be a perfectoid space. Then, the underlying category of the étale site $X_{\text{ét}}$ consists of perfectoid spaces étale over X. A family of morphisms in $X_{\text{ét}}$ is a covering if the family of underlying continuous mappings is a covering.

Theorem 5.1 and Theorem 4.1.2 imply the following.

Corollary 5.3 [7, Theorem 7.12] *The tilting induces an isomorphism $X_{\text{ét}} \to X^{\flat}_{\text{ét}}$ of the étale sites.*

This completes the construction of the bottom arrow in the diagram (1.1).

6 Complete intersections in toric varieties

We recall the definition of toric varieties. Let P be a free abelian group of finite rank and $N = Hom(P, \mathbf{Z})$ be the dual. A cone $\sigma \subset N_{\mathbf{R}}$ is said to be rational if it is spanned by a finitely many elements of N. It is said to be strongly convex if it does not contain a line. A sub cone τ of a rational cone σ is called a face if there exists $a \in P$ such that $f(a) \geqq 0$ for every $f \in \sigma$ and $\tau = \{f \in \sigma \mid f(a) = 0\}$. For a rational cone $\sigma \subset N_{\mathbf{R}}$, let $\sigma^\vee \subset P_{\mathbf{R}}$ denote the dual $\{a \in P_{\mathbf{R}} \mid f(a) \geqq 0 \text{ for } f \in \sigma\}$ and set $P_\sigma = P \cap \sigma^\vee$.

A fan Σ is a finite set of strongly convex rational cones of $N_{\mathbf{R}}$ such that a face τ of an element σ of Σ is an element of Σ and the intersection $\sigma \cap \tau$ of elements σ, τ of Σ is a face of σ and of τ. We say a fan Σ is proper if the union $\bigcup_{\sigma \in \Sigma} \sigma$ is equal to $N_{\mathbf{R}}$. We say a fan Σ is smooth, if the monoid $N \cap \sigma$ is isomorphic to the product of copies of $\mathbf{N}$ for every $\sigma \in \Sigma$.

For a fan Σ and a field k, we define the toric variety $X_{\Sigma,k}$ by patching $X_{\sigma,k} = \mathrm{Spec}\ k[P_\sigma]$ along $X_{\sigma\cap\tau,k}$. The toric variety $X_{\Sigma,k}$ has a natural action of the torus $T = \mathrm{Spec}\ k[P] \subset X_{\Sigma,k}$ defined by $k[P_\sigma] \to k[P_\sigma] \otimes k[P] = k[P_\sigma \times P]$ induced by $P_\sigma \to P_\sigma \times P \colon a \mapsto (a, a)$. If Σ is a proper fan, the toric variety $X_{\Sigma,k}$ is proper. If Σ is a smooth fan, the toric variety $X_{\Sigma,k}$ is smooth. If τ is a one dimensional face of a smooth fan Σ, the ideals of $k[P_\sigma]$ generated by $\{a \in P_\sigma \mid f(a) > 0 \text{ for } f \in \tau\}$ define a smooth irreducible divisor D_τ of $X_{\Sigma,k}$.

Example 6.1 Set $P = \mathrm{Ker}(\mathrm{sum} \colon \mathbf{Z}^{n+1} \to \mathbf{Z})$ and $N = \mathbf{Z}^{n+1}/\Delta\mathbf{Z}$. For a subset $\sigma \subsetneqq \{0, \ldots, n\}$, let σ also denote abusively the cone of $N_{\mathbf{R}}$ spanned by the images of the standard basis $e_i \in \mathbf{Z}^{n+1}$ for $i \in \sigma$. Let Σ be the set of cones σ associated to subsets $\sigma \subsetneqq \{0, \ldots, n\}$. For $i = 0, \ldots, n$, set $\sigma_i = \{j \mid j \neq i\}$. Then, P_{σ_i} is generated by $x_j - x_i$ where x_j denote the standard basis of $\mathbf{Z}^{n+1} \supset P$. Thus the toric variety $X_{\Sigma,k}$ is defined by patching $\mathrm{Spec}\ k\Big[\dfrac{X_j}{X_i}\Big]$ and is nothing but the projective space $\mathbf{P}^n_k$.

Let K be a perfectoid field and Σ be a proper smooth fan. For a face $\sigma \in \Sigma$, set $P_\sigma^{1/p^\infty} = P_{\mathbf{Z}[\frac{1}{p}]} \cap \sigma^\vee$ and $K\langle P_\sigma\rangle, \mathcal{O}_K\langle P_\sigma\rangle, K\langle P_\sigma^{1/p^\infty}\rangle$ and $\mathcal{O}_K\langle P_\sigma^{1/p^\infty}\rangle$ be the ϖ-adic completions. Then, we define an adic space $X^{\mathrm{ad}}_{\Sigma,K}$ and a perfectoid space $X^{\mathrm{perf}}_{\Sigma,K}$ by patching $\mathrm{Spa}(K\langle P_\sigma\rangle, \mathcal{O}_K\langle P_\sigma\rangle)$ and $\mathrm{Spa}(K\langle P_\sigma^{1/p^\infty}\rangle, \mathcal{O}_K\langle P_\sigma^{1/p^\infty}\rangle)$ respectively. The construction of $X^{\mathrm{perf}}_{\Sigma,K}$ corresponds to the vertical arrow with ? in the diagram (1.1).

Assume that K is of mixed characteristic $(0,p)$. Then the tilt of $X^{\rm perf}_{\Sigma,K}$ is $X^{\rm perf}_{\Sigma,K^\flat}$ and we obtain morphisms of étale sites;

$$X^{\rm ad}_{\Sigma,K,{\rm \acute{e}t}} \longleftarrow X^{\rm perf}_{\Sigma,K,{\rm \acute{e}t}} \xrightarrow{\ \flat\ } X^{\rm perf}_{\Sigma,K^\flat,{\rm \acute{e}t}} \longrightarrow X^{\rm ad}_{\Sigma,K^\flat,{\rm \acute{e}t}} \tag{6.1}$$

By Corollary 5.3, the middle arrow is an isomorphism. Since a surjective radicial morphism induces an isomorphism on the étale site, the right arrow is also an isomorphism. For a prime number $\ell \neq p$, the right arrow induces an isomorphism $H^q(X^{\rm ad}_{\Sigma,\bar K,{\rm \acute{e}t}}, \mathbf{Q}_\ell) \to H^q(X^{\rm perf}_{\Sigma,\bar K,{\rm \acute{e}t}}, \mathbf{Q}_\ell)$ by the proper base change theorem. This means that we do not lose too much information by going down in the diagram (1.1).

Let k be a field and Σ be a smooth proper fan. A closed subscheme Y of $X_{\Sigma,k}$ defined by a non-zero section of an invertible sheaf defined by a linear combination of D_τ for one-dimensional faces is called a hypersurface of $X_{\Sigma,k}$.

Theorem 6.2 [7, Theorem 9.6] *Let Y be a smooth closed subscheme of codimension c of a smooth projective toric variety $X_{\Sigma,k}$ over a finite extension k of $\mathbf{Q}_p$. If there exist hypersurfaces $H_1,\ldots,H_c$ of $X_{\Sigma,k}$ such that the underlying set of Y is equal to that of the intersection $H_1\cap\ldots\cap H_c$, then the weight-monodromy conjecture is true for Y.*

We take a perfectoid extension K of k, for example the completion of $k(\zeta_{p^\infty})$. We prove Theorem 6.2 by constructing a proper smooth variety Z of dimension $\dim Y$ over $K^\flat$ defined over a dense subfield $k_0 \subset K^\flat$ that is a function field of one variable over $\mathbf{F}_p$ and a generically finite morphism $Z \to X_{\Sigma,K^\flat}$ and an injection $H^q(Y_{\bar K}, \mathbf{Q}_\ell) \to H^q(Z_{\bar K^\flat}, \mathbf{Q}_\ell)$ compatible with the canonical isomorphism ${\rm Gal}(\bar K/K) \leftarrow {\rm Gal}(\bar K^\flat/K^\flat)$ such that the image is a direct summand.

By the theory of étale cohomology of adic space, there exists an open neighborhood $\widetilde Y$ of $Y^{\rm ad}_K$ such that the pull-back $H^q(\widetilde Y_{\bar K,{\rm \acute{e}t}}, \mathbf{Q}_\ell) \to H^q(Y_{\bar K,{\rm \acute{e}t}}, \mathbf{Q}_\ell)$ is an isomorphism for every $q \geqq 0$. Let $\pi\colon X^{\rm ad}_{\Sigma,K^\flat} \to X^{\rm ad}_{\Sigma,K}$ denote the composition of continuous mappings similarly defined as (6.1). Then, by approximation, we find a closed subscheme $Z_0 \subset X_{\Sigma,K^\flat}$ of codimension c defined over a subfield $k_0 \subset K^\flat$ that is a function field of one variable over $\mathbf{F}_p$ such that $Z^{\rm ad}_{0,K^\flat}$ is contained in $\flat(\widetilde Y)$. By a theorem of de Jong, we find an alteration $Z \to Z_0$ proper smooth over k_0. The weight-monodromy conjecture for Z is known by Deligne in [1].

The diagram

$$\begin{array}{ccc}
H^q(X_{\Sigma,\bar K,{\rm \acute{e}t}}, \mathbf{Q}_\ell) & \xrightarrow{\ \simeq\ } & H^q(X_{\Sigma,\bar K^\flat,{\rm \acute{e}t}}, \mathbf{Q}_\ell) \\
\big\downarrow & & \big\downarrow \\
H^q(\widetilde Y_{\bar K,{\rm \acute{e}t}}, \mathbf{Q}_\ell) & \longrightarrow & H^q(\flat(\widetilde Y_{\bar K})_{\rm \acute{e}t}, \mathbf{Q}_\ell) \\
\simeq\big\downarrow & & \big\downarrow \\
H^q(Y_{\bar K,{\rm \acute{e}t}}, \mathbf{Q}_\ell) & & H^q(Z_{\bar K^\flat,{\rm \acute{e}t}}, \mathbf{Q}_\ell)
\end{array} \tag{6.2}$$

defines a morphism $H^q(Y_{\bar K}, \mathbf{Q}_\ell) \to H^q(Z_{\bar K^\flat}, \mathbf{Q}_\ell)$ compatible with the isomorphism (1.2). The Poincaré duality implies that, in order to show that it satisfies the required property, it suffices to show that it is non-zero for $q = 2\dim Y$. If it was zero, the composition of the right vertical arrows would be zero. This contradicts the fact that the $\dim Y$-th power of the class of an ample divisor of $X_{\Sigma,\bar K^\flat}$ is non-zero in $H^{2\dim Y}(Z_{\bar K^\flat,{\rm \acute{e}t}}, \mathbf{Q}_\ell)$. Since the weight-monodromy conjecture is known for $H^q(Z_{\bar K^\flat,{\rm \acute{e}t}}, \mathbf{Q}_\ell)$ by Deligne [1], it also holds for a direct summand $H^q(Y_{\bar K,{\rm \acute{e}t}}, \mathbf{Q}_\ell)$.

The author thanks A. Abbes, T. Mihara, K. Miyatani, K. Tokimoto and N. Umezaki for useful comments on a preliminary version.

References

1. P. Deligne, *La conjecture de Weil II,* Publ. Math. IHES, **52** (1980), 137-252.
2. G. Faltings, *Almost étale extensions,* in Cohomologies p-adiques et applications arithmétiques, II, Astérisque **279** (2002), 185-270.
3. J.-M. Fontaine, *Perfectoïdes, presque pureté et monodromie-poids, d'après Peter Scholze*, Séminaire Bourbaki, 2011-2012, n° 1057.
4. J.-M. Fontaine, J.-P. Wintenberger, *Extensions algébrique et corps des normes des extensions APF des corps locaux,* C. R. Acad. Sci. Paris Ser. A-B, 288 (1979), 441-444.
5. O. Gabber, L. Ramero, *Almost ring theory*, Lecture notes in Math. 1800, Springer, Berlin, 2003.
6. R. Huber, *Étale cohomology of rigid analytic varieties and adic spaces*, Aspects of Math. vol. E30, Vieweg, Braunschweig, 1996.
7. P. Scholze, *Perfectoid spaces*, Publ. Math. IHES, **116** (2012), 245-313
8. T. Tsuji, *Notes on almost étale extensions of Faltings,* (preprint).

THE TATE CONJECTURE FOR K3 SURFACES A SURVEY OF SOME RECENT PROGRESS

Vasudevan Srinivas

Abstract The most important open questions in the theory of algebraic cycles are the Hodge Conjecture, and its companion problem, the Tate Conjecture. Both these questions attempt to give a description of those cohomology classes on a nonsingular proper variety which are represented by algebraic cycles, in terms of intrinsic structure which is present on the cohomology of such a variety (namely, a Hodge decomposition, or a Galois representation). For the Hodge conjecture, the case of divisors (algebraic cycles of codimension 1) was settled long ago by Lefschetz and Hodge, and is popularly known as the Lefschetz $(1, 1)$ theorem, though there is little general progress beyond that case. However, even this case of divisors is an open question for the Tate Conjecture, in general, even for divisors on algebraic surfaces. After giving an introduction to these problems, I will discuss the recent progress on the Tate Conjecture for K3 surfaces, around works of M. Lieblich, D. Maulik, F. Charles and K. Pera.

1 What is the Tate Conjecture?

The goal of this article is to discuss some recent progress on the *Tate Conjecture for K3 surfaces*, as contained in the recent papers [14,5,19] (see also the related papers [12, 13]).

We begin by recalling what the Tate Conjecture is.

Let K be a field, which is given to be finitely generated over its prime subfield. Thus, K is a finite algebraic extension of either $\mathbb{Q}(x_1, \ldots, x_n)$ or $\mathbb{F}_p(x_1, \ldots, x_n)$, where $x_1, \ldots .x_n$ are algebraically independent over the prime field – we allow also the case when $n = 0$, i.e., K is either an algebraic number field or a finite field, and these are perhaps the most fundamental cases. Let $\overline{K}$ be an algebraic closure of K.

Let X be a smooth projective (or proper) variety over K. Let $\overline{X} = X \times_K \overline{K}$ denote the corresponding variety over the algebraically closed field $\overline{K}$. Note that the Galois group $G = \mathrm{Gal}(\overline{K}/K)$ acts on $\overline{X}$ compatibly with its action on $\overline{K}$.

Lecture at VIASM, Hanoi, July 20-21, 2013

V. Srinivas
Tata Institute, Mumbai, India
E-mail: srinivas@math.tifr.res.in

The "Tate Conjectures" for X refer to a certain list of statements which conjecturally describe relations between algebraic cycles on X, with rational coefficients, and the étale cohomology of $\overline{X}$, with its natural G-action.

More specifically, we fix a prime ℓ which is invertible in K, and consider on the one hand, the étale cohomology group

$$H^{2i}_{et}(\overline{X}, \mathbb{Q}_\ell(i)) H^{2i}_{et}(\overline{X}, \mathbb{Q}_\ell) \otimes_{\mathbb{Q}_\ell} \mathbb{Q}_\ell(1)^{\otimes i}$$

which is a finite dimensional $\mathbb{Q}_\ell$-vector space with a continuous action of the Galois group $G = \mathrm{Gal}(\overline{K}/K)$. The Galois action is obtained as an i-fold "Tate twist" of $H^{2i}_{et}(\overline{X}, \mathbb{Q}_\ell)$ with its natural Galois action, induced from the action on the scheme $\overline{X}$, and functoriality of étale cohomology. Here $\mathbb{Q}_\ell(1)$ denotes the 1-dimensional $\mathbb{Q}_\ell$-Galois module defined by

$$\mathbb{Q}_\ell(1) = \left(\lim_{\substack{\longleftarrow \\ n}} \mu_{\ell^n}\right) \otimes_{\mathbb{Z}} \mathbb{Q},$$

where μ_m denotes the m-th roots of unity in $\overline{K}$ with its natural G-action. (One basic reference for this theory is Milne's book [15]; see also the useful lecture notes [16], available at his website, along with other interesting material related to the Tate conjecture).

On the other hand, one has the Chow group $CH^i(X)_{\mathbb{Q}}$ of algebraic cycles of codimension i on X, with rational coefficients[1], modulo rational equivalence. If X has dimension d, one has an *intersection product*

$$CH^i(X)_{\mathbb{Q}} \otimes CH^{d-i}(X)_{\mathbb{Q}} \to \mathbb{Q}$$

which is obtained using intersections of subvarieties of X of complementary dimensions; to show that this is well-defined, and has reasonable properties, is a prominent goal of *intersection theory* in algebraic geometry (for example, this is exposed carefully in Fulton's book [8]).

This defines a natural equivalence relation on classes of algebraic cycles: two cycle classes in $CH^i(X)_{\mathbb{Q}}$ are called *numerically equivalent* if they have the same intersection number with any class in $CH^{d-i}(X)_{\mathbb{Q}}$. It is known that the quotient

$$CH^i_{num}(X)_{\mathbb{Q}} = \frac{CH^i(X)}{\text{numerical equivalence}}$$

is a finite dimensional $\mathbb{Q}$-vector space.

We recall also that a cycle class α on $\overline{X}$ is said to be *algebraically equivalent* to 0 if there is a cycle ξ on $\overline{X} \times Z$ for some irreducible projective smooth curve Z over $\overline{K}$, and points $a, b \in Z$, so that if $i_z : \overline{X} \cong \overline{X} \times \{z\} \hookrightarrow \overline{X} \times Z$, then $\alpha = i_a^*(\xi) - i_b^*(\xi)$. If we can choose $Z \cong \mathbb{P}^1$, this defines rational equivalence. It follows from standard properties of intersection numbers that if a cycle is algebrically equvalent to 0, it is also numerically equivalent to 0.

The theory of the *cycle map in étale cohomology* produces a natural homomorphism for any $0 \leq i \leq \dim X$

$$CH^i(\overline{X})_{\mathbb{Q}} \to H^{2i}_{et}(\overline{X}, \mathbb{Q}_\ell(i)) \tag{1}$$

which yields therefore a map

$$CH^i(X)_{\mathbb{Q}} \to H^{2i}_{et}(\overline{X}, \mathbb{Q}_\ell(i)), \tag{2}$$

[1] Algebraic cycles of codimension i are elements of the free abelian group on integral subschemes of codimension i; cycles with rational coefficients are obtained by tensoring the group of cycles with $\mathbb{Q}$.

for each $0 \leq i \leq \dim X$, obtained as a composition

$$CH^i(X)_{\mathbb{Q}} \hookrightarrow CH^i(\overline{X})_{\mathbb{Q}} \to H^{2i}_{et}(\overline{X}, \mathbb{Q}_\ell(i)).$$

The map (2) has the following properties:

(i) its *image* is contained in the subspace

$$H^{2i}_{et}(\overline{X}, \mathbb{Q}_\ell(i))^G$$

of *invariants* for the G-action (where $G = \mathrm{Gal}(\overline{K}/K)$)

(ii) its *kernel* consists of classes which are numerically equivalent to 0

(iii) the cycles algebraically equivalent to 0 all lie in the kernel.

The first property is because the cycle map (1) is in fact equivariant for the natural action of G on both sides, from the functorial properties of the cycle class map, while $CH^i(X)_{\mathbb{Q}}$ is clearly contained in the G-invariants in $CH^i(\overline{X})_{\mathbb{Q}}$ (in fact one can show easily that $CH^i(X)_{\mathbb{Q}}$ consists precisely of the G-invariants in $CH^i(\overline{X})_{\mathbb{Q}}$).

The second property is because of a compatibility between intersection numbers of classes of algebraic cycles, and cup products of the corresponding étale cohomology classes, i.e., there is an equality between algebraic intersection numbers and ℓ-adic intersection numbers.

The third property can be deduced from the defintion of algebraic equivalence, and the Kunneth formula for the cohomology of $\overline{X} \times Z$ for any smooth projective curve Z over $\overline{K}$.

If the ground field K has characteristic 0, and we embed K into the complex number field $\mathbb{C}$, we may associate to X a complex algebraic variety $X_{\mathbb{C}}$. Then the above statements about cycle classes have analogues, where instead of étale cohomology, we use the singular cohomology groups $H^*_B(X_{\mathbb{C}}, \mathbb{Z})$ (also referred to in the subject as the "Betti cohomology") of the corresponding complex manifold $X(\mathbb{C})$. This complex manifold is thus an even dimensional, oriented compact topological manifold, and we have a known compatibility between algebraic and topological intersection numbers for classes of algebraic cycles; the property (ii) is motivated by this "more classical" statement. Similarly (iii) follows from the Kunneth formula in topology.

The analogue of (i) for Betti cohomology is somewhat subtler. First, instead of a Galois action, the "extra structure" for Betti cohomology is given by its *Hodge decomposition* (for each $0 \leq n \leq 2\dim X$)

$$H^n_B(X_{\mathbb{C}}, \mathbb{Z}) \otimes \mathbb{C} = \oplus_{p+q=n} H^{p,q}(X_{\mathbb{C}}),$$

where we have

$$\overline{H^{p,q}(X_{\mathbb{C}})} = H^{q,p}(X_{\mathbb{C}}),$$

and natural isomorphisms

$$H^{p,q}(X_{\mathbb{C}}) \cong H^q(X_{\mathbb{C}}, \Omega^p_{X_{\mathbb{C}}/\mathbb{C}})$$

identifying the Hodge "pieces" with certain cohomology groups of (coherent) sheaves of (algebraic Kähler) differentials. A detailed account of this theory, along with interesting applications, may be found in the book [22] (in 2 parts) of C. Voisin.

The properties of the Hodge decomposition are abstracted into the notion of a *pure Hodge structure* (see Voisin's book for explanations). Using this language, we may "Tate twist" the Betti cohomology as well, defining

$$H^{2i}_B(X_{\mathbb{C}}, \mathbb{Z}(i)) = H^{2i}_B(X_{\mathbb{C}}, \mathbb{Z}) \otimes \mathbb{Z}(1)^{\otimes i},$$

where $\mathbb{Z}(1) = 2\pi\iota\mathbb{Z} \subset \mathbb{C}$ (and $\iota = \sqrt{-1}$), with a trivial Hodge decomposition, given by

$$\mathbb{Z}(1) \otimes \mathbb{C} = (\mathbb{Z}(1) \otimes \mathbb{C})^{-1,-1}$$

Then one may reinterpret the classical cycle map with values in singular cohomology as yielding a map

$$CH^i(X_{\mathbb{C}}) \to H_B^{2i}(X_{\mathbb{C}}, \mathbb{Z}(i))$$

which lands in the subgroup of *Hodge cycles*

$$Hg^i(X_{\mathbb{C}}) = H_B^{2i}(X_{\mathbb{C}}, \mathbb{Z}(i))^{0,0}$$

consisting of elements of $H_B^{2i}(X_{\mathbb{C}}, \mathbb{Z}(i))$ which map to the (0,0) Hodge subspace of $H_B^{2i}(X_{\mathbb{C}}, \mathbb{Z}(i)) \otimes \mathbb{C}$.

Viewing (as in Deligne's article [7]) a pure Hodge structure as corresponding to an $\mathbb{R}$-representation of a certain (non-split) real torus on the real vector space $V \otimes \mathbb{R}$, where V is a fintely generated abelian group, we note that the (0,0) part is precisely the space of invariants for the group action. Thus, the image of the cycle map to Betti cohomology is contained, after $\otimes\mathbb{R}$, in a certain space of group invariants. We may view this as analogous to the Galois invariance of the image of cycle classes in étale cohomology.

The famous *Hodge conjecture* is the assertion that the corresponding map with rational coefficients

$$CH^i(X_{\mathbb{C}})_{\mathbb{Q}} \to Hg^i(X_{\mathbb{C}})_{\mathbb{Q}} \tag{3}$$

is surjective. That its kernel consists exactly of cycles numerically equivalent to 0 is also a well-known conjecture, which is a consequence of Grothendieck's *Standard Conjectures* (see [9], and also [11] for an exposition, explaining also the background). This description of the kernel has been proved in case $i = 1$, i.e., for the case of divisors, both for the Betti cohomology and the étale cohomology; further, algebraic and numerical equivalence are known to coincide for divisor classes with rational coefficients (from well-known examples of Griffiths, this does not hold in general for cycle classes in codimensions i with $1 < i < \dim X$ — see [22] for more details).

The "Tate Conjecture" is the following list of statements:

(T1) The cycle map (2) induces a surjection

$$CH^i(X) \otimes \mathbb{Q}_\ell \to H_{et}^{2i}(\overline{X}, \mathbb{Q}_\ell(i))^G$$

for each $0 \leq i \leq \dim X$.

(T2) The image of the map in (T1) is identified with $CH^i_{num}(X) \otimes \mathbb{Q}_\ell$.

A related conjecture, attributed to Grothendieck and Serre, is:

(T3) The action of G on $H_{et}^{2i}(\overline{X}, \mathbb{Q}_\ell)$ is semisimple.

Tate originally made these conjectures in his 1964 lecture notes [20], and explored their consequences, e.g., for poles of zeta functions. These notes were perhaps not widely accessible earlier, but are now available online (see Milne's website, for example). The survey article [21] contains more information, including motivations and important applications to the theory of motives, as well as an overview of known results, etc. Strictly speaking, there is one set of "Tate Conjectures" for each prime ℓ as above; we will not dwell on this point explicitly[2].

[2] Tate's lecture notes do make this point as well, and make the additional conjecture that (T1)-(T3) are "independent of ℓ" in a suitable sense: that the conjectures for one ℓ are equivalent to the conjectures for any ℓ.

Tate also points out in [21] that, for example, if X is a variety over K for which (T1) holds (for a certain codimension i) for the base-change X_L to a finite algebraic extension L of K, then (T1) holds for X (in that same codimension i).

This remark also implies that if we make a base change instead to a finitely generated extension field L, then the Tate conjecture for X_L implies that for X over K. One way to see this is to note, on the one hand, that there are canonical isomorphisms

$$CH^i(\overline{X})_{\mathbb{Q}}/(\text{algebraic equivalence}) \cong CH^i(X_{\overline{L}})_{\mathbb{Q}}/(\text{algebraic equivalence}),$$

$$H^{2i}_{et}(\overline{X}, \mathbb{Q}_\ell(i)) \cong H^{2i}_{et}(X_{\overline{L}}, \mathbb{Q}_\ell(i))$$

determined by an inclusion $\overline{K} \to \overline{L}$(which determines a homomorphism $\mathrm{Gal}(\overline{L}/L) \to \mathrm{Gal}(\overline{K}/K)$, well-defined upto conjugation), and both the isomorphisms are compatible with the action of $\mathrm{Gal}(\overline{L}/L)$. Now it suffices to remark that the image of homomorphism $\mathrm{Gal}(\overline{L}/L) \to \mathrm{Gal}(\overline{K}/K)$ has finite index.

We recall here that the Hodge conjecture for divisors (i.e., the surjectivity of (3) for $i = 1$) is true, and even holds with $\mathbb{Z}$-coefficients[3]. This is the assertion of the *Lefschetz (1,1) theorem*:

Theorem *Let Y be a smooth proper variety over $\mathbb{C}$. The image of the cycle map*

$$CH^1(Y) \to H^2_B(Y, \mathbb{Z}(1))$$

coincides with $Hg^1(Y)$.

We note that since we have also put in a Tate twist in our notion of Hodge cycle, we should really call this the "Lefschetz (0,0) theorem"!

The standard proof of this theorem (see [22], for example) uses the exponential sheaf sequence, and the assertion that the isomorphism classes of holomorphic and algebraic line bundles are naturally identified (this is a particular instance of Serre's GAGA principle).

However, in contrast, the assertion (T1) is unknown in general, even for divisors on a smooth projective surface. As a result, it is sometimes the practice to use the term "Tate Conjecture" to refer to the statement (T1) alone, and that too, to the special case $i = 1$. This is the sense in which "Tate Conjecture" appears in the title of this article, and usually also in the recent papers that we intend to give an exposition of below.

2 Some earlier results: more background

One of the most important cases when (T1) is known for divisors is the following (we remind the reader that K is finitely generated over its prime subfield).

Theorem

(i) *Let X be an abelian variety over K. Then the Tate conjecture for divisors* (T1) *holds for X.*

[3] It is known, however, that the "integral Hodge conjecture" is in general false in any codimension i with $2 \le i < \dim X$, from examples of Atiyah-Hirzebruch and Kollár.

(ii) *If X, Y are two abelian varieties over K, and* $\mathrm{Hom}(X,Y)$ *denotes the group of homomorphisms of abelian varieties, then the natural map*

$$\mathrm{Hom}(X,Y)\otimes\mathbb{Q}_\ell \to \mathrm{Hom}_G(H^1_{et}(\overline{Y},\mathbb{Q}_\ell(1)), H^1_{et}(\overline{X},\mathbb{Q}_\ell(1))$$

(*with values in the Galois equivariant homomorphisms between étale cohomologies*) *is an isomorphism of* $\mathbb{Q}_\ell$ *vector spaces.*

For the case when the ground field K is a finite field, this result is due to Tate himself. Tate also showed that the statement (i) for all abelian varieties X over K is equivalent to the stement (ii) for all pairs X, Y over K, and further to the special case of (ii) for $X = Y$, for all X over K. The case when K is an algebraic number field is a famous theorem of G. Faltings, which was also a key step in his first proof of the Mordell Conjecture, and other important results. Y. Zarhin has studied the case when K is finitely generated over its prime subfield. A recent accessible survey of these results, and other interesting related work in the case of abelian varieties, particularly in characteristic p, may be found in [4]. This paper also provides useful background to understand some aspects of the papers under discussion later in this article.

We now start the discussion of the subject proper: the case of K3 surfaces. In fact, the new results are basically for the important case of *K3 surfaces over finite fields.* So we will concentrate primarily on what is needed for this case, though other cases will be considered if they are relevant to the proofs for the finite field case. In particular, the proofs involve the study of deformations, and liftings to characteristic 0, so it is important for us to study properties of K3 surfaces in characteristic 0.

Recall that a *K3 surface* over K is a smooth projective surface X over K such that
(i) X is geometrically integral, and has trivial canonical sheaf, $\Omega^2_{X/K}\cong\mathcal{O}_X$
(ii) X is algebraically simply connected, and satisfies $H^1(X,\mathcal{O}_X)=0$.
These conditions imply that

$$H^0(X,\mathcal{O}_X)\cong H^2(X,\mathcal{O}_X)\cong K$$

and that $H^2_{et}(\overline{X},\mathbb{Q}_\ell)$ has dimension 22 (from Noether's formula, relating $\chi(\mathcal{O}_X)$ to Chern numbers).

It is also useful to recall the algebraic Hodge index theorem for a K3 surface X, valid for arbitrary K: the intersection pairing on $CH^1(\overline{X})\cong\mathrm{Pic}(\overline{X})$ has signature $(1,-(r-1))$ where $r=\mathrm{rank}\,\mathrm{Pic}(\overline{X})$. (A similar assertion holds for the pairing on $CH^1_{num}(Y)$ for an arbitrary smooth projective surface Y over an algebraically closed field: see for example [10], V, Theorem 1.9.) We note that $r\le 22$, and $r\le 20$ in characteristic 0 (see below).

Standard examples of K3 surfaces are
(i) smooth quartic surfaces, i.e., smooth hypersurfaces $X\subset\mathbb{P}^3$ of degree 4
(ii) smooth sextic double planes, i.e., smooth surfaces X with an involution, whose quotient is $\mathbb{P}^2$, such that the fixed points map isomorphically onto a smooth curve $C\subset\mathbb{P}^2$ of degree 6
(iii) the surface X obtained by resolving the (mild) singularities of a quotient of an abelian variety S of dimension 2 modulo the involution given by multiplication by -1 for the group structure on S; such an X is called a *Kummer surface* (if a similar construction is made with elliptic curves instead, the quotient is $\mathbb{P}^1$, and we obtain the usual representation of the elliptic curve in Weierstrass form, and the quotient rational map is essentially the Weierstrass $\wp$ function).

If $K\subset\mathbb{C}$, the corresponding complex K3 surface has Hodge groups

$$H^{0,0}(X_{\mathbb{C}})\cong H^{2,2}(X_{\mathbb{C}})\cong H^{2,0}(X_{\mathbb{C}})\cong H^{0,2}(X_{\mathbb{C}})\cong\mathbb{C},\ \ \dim_{\mathbb{C}} H^{1,1}(X_{\mathbb{C}})=20,$$

and all other $H^{p,q}(X_{\mathbb{C}})$ vanish. Further, the cup product

$$H^2_B(X_{\mathbb{C}},\mathbb{R})\otimes H^2_B(X_{\mathbb{C}},\mathbb{R})\to\mathbb{R}$$

is a non-degenerate symmetric bilinear form with signature $(3,-19)$ — this follows from the (cohomological) *Hodge index theorem* (see [22]). The topological cycle map

$$\mathrm{Pic}(\overline{X})=\mathrm{Pic}(X_{\mathbb{C}})\to H^2_B(X_{\mathbb{C}},\mathbb{Z}(1))$$

is injective with torison-free cokernel, where $H^2_B(X_{\mathbb{C}},\mathbb{Z})\cong\mathbb{Z}^{22}$. In particular one has that $\mathrm{rank}\,\mathrm{Pic}(X_{\mathbb{C}})\leq 20$.

The Tate conjecture (T1) for divisors is known for K3 surfaces over finitely generated fields K of characteristic 0; see [1], where, using techniques from Deligne's proof of Weil's Riemann Hypothesis for K3 surfaces (prior to his proof of this in general), André is able to reduce it to the case of abelian varieties (where it is known after Faltings and Zarhin). We discuss this argument in more detail below, since these techniques also motivate proofs of the recent new results.

There is one older (1973) result on the Tate conjecture for K3 surfaces over finite fields, using methods which are apparently not directly related to the later developments (see [3]).

Theorem (Artin, Swinnerton-Dyer) *Let X be a K3 surface over a finite field K such that $\overline{X}$ supports a pencil of elliptic curves. Then the Tate conjecture for divisors (T1) holds for X.*

Corollary *Let X be a K3 surface over a finite field such that* $\mathrm{rank}\,\mathrm{Pic}(\overline{X})\geq 5$*. Then (T1) for divisors holds for X.*

In fact, by the Hasse-Minkowski theorem[4] on quadratic forms over $\mathbb{Q}$, we see that $CH^1(\overline{X})\cong\mathrm{Pic}(\overline{X})$ must contain a non-zero element with self-intersection 0. After making a finite extension of the ground field, we may assume this is represented by an element of $\mathrm{Pic}(X)$ which is not a nontrivial multiple of another class in $\mathrm{Pic}(\overline{X})$. If $\mathcal{L}$ is the corresponding invertible sheaf, then from the Riemann-Roch theorem for X, and Serre duality, we see that either $\mathcal{L}$ or its inverse has a non-zero global section, whose zero scheme is an effective Cartier divisor $D\subset X$ with $D^2=0$, and such that $|D|$ has dimension 1. This gives a pencil of curves with arithemtic genus 1 (by the adjunction formula), which is an elliptic pencil on X, except possibly in characteristics 2 or 3, when it could be a "quasi-elliptic pencil", i.e., a pencil of irreducible rational curves with cuspidal singularities (in this case, in fact $CH^1(\overline{X})\otimes\mathbb{Q}_\ell\cong H^2_{et}(\overline{X},\mathbb{Q}_\ell(1))$, so (T1) is easily seen to hold for X).

Another older pair of results are due to N. Nygaard [17], and to Nygaard and A. Ogus [18]. We may associate to a K3 surface over a finite field K its *formal Brauer group* $\hat{\mathrm{Br}}(X)$, as follows: it is a formal group which pro-represents the functor on Artinian local K-algebras

$$A\mapsto\ker(\mathrm{Br}\,(X\times_K A)\to\mathrm{Br}\,(X)),$$

where $\mathrm{Br}\,(S)=H^2_{et}(S,\mathbb{G}_m)_{tors}$ is Grothendieck's cohomological Brauer group. Artin[2] showed that for a K3 surface, this procedure does indeed yield a formal group, which is commutative and 1-dimensional, and either equals $\hat{\mathbb{G}}_a$, the additive formal group (in which case we say X has *infinite height*), or has finite height (which is in fact ≤ 20:

[4] That any indefinite quadratic form over $\mathbb{Q}$ of dimension ≥ 5 has a nontrivial zero.

however, the precise definition of height will not be relevant for our discussions below). In the first case, we call X *supersingular*; in this case, in fact the G action on $H^2_{et}(\overline{X}, \mathbb{Q}_\ell(1))$ factors through a finite quotient, and so (T1) amounts to the assertion that the cycle map

$$\mathrm{Pic}(\overline{X}) \otimes \mathbb{Q}_\ell \to H^2_{et}(\overline{X}, \mathbb{Q}_\ell(1))$$

is an *isomorphism*, or equivalently, that $\mathrm{rankPic}(\overline{X}) = 22$. It is this case which had not been fully resolved, and remained an open question for some 30 years.

In the contrary case when the formal Brauer group $\hat{\mathrm{Br}}(X)$ has finite height, the results of Nygaard and Ogus imply that (T1) holds, provided K has characteristic ≥ 5. In arbitrary characteristic, Nygaard showed earlier that if we assume the height is 0, which more concretely means that the Frobenius action on the 1-dimensional vector space $H^2(X, \mathcal{O}_X)$ is non-zero, then (T1) holds for divisors on X (we refer to such K3 surfaces as *ordinary*, by analogy with ordinary abelian varieties).

3 The Kuga-Satake construction

The basis of almost all the discussion of the Tate conjecture for divisors on K3 surfaces (starting with the works cited of Nygaard and Nygaard-Ogus) rests on properties of the *Kuga-Satake construction*, as interpreted by Deligne [6]. We now recall this construction in a convenient form, and list some of its key properties.

Let V be a free abelian group of finite rank, which underlies a Hodge structure of weight 0, such that $V_{\mathbb{C}} = V \otimes \mathbb{C}$ has

$$\dim V^{-1,1} = \dim V^{1,-1} = 1, \;\; V^{p,-p} = 0 \;\text{ if } p \notin \{0, 1, -1\}.$$

Assume further that V admits a symmetric bilinear form ψ with values in $\mathbb{Z}$, such that
(i) $\psi : V \otimes V \to \mathbb{Z}$ is a morphism of Hodge structures of weight 0, where $\mathbb{Z}$ is the trivial Hodge structure (of pure type $(0, 0)$
(ii) $\psi \otimes \mathbb{Q}$ is a non-degenerate $\mathbb{Q}$-bilinear form
(iii) $\psi \otimes \mathbb{R}$ is positive definite on

$$(V \otimes \mathbb{R}) \cap V^{0,0},$$

and negative definite on

$$(V \otimes \mathbb{R}) \cap (V^{-1,1} + V^{1,-1}) \cong \mathbb{R}^2.$$

In particular $\psi \otimes \mathbb{R}$ has signature $(n - 2, -2)$, where $\mathrm{rank}\, V = n$.

Let $C^+ = C^+(V, \psi) \subset C = C(V, \psi)$ be the inclusion of the even subalgebra into the Clifford algebra of V with respect to the quadratic form ψ. Then $C_{\mathbb{R}} = C \otimes \mathbb{R}$ contains the Clifford algebra of the 2-dimensional negative definite subspace $V_{\mathbb{R}} \cap (V^{-1,1} + V^{1,-1})$, whose even part is isomorphic to $\mathbb{C}$ as an $\mathbb{R}$-algebra. This makes $C^+_{\mathbb{R}}$ into a complex vector space (using left multiplication), and hence $B(V) := C^+_{\mathbb{R}}/C^+$ into a compact complex torus, and thereby endows C^+ with a Hodge structure of weight -1, as the homology of that torus. If we choose an element $a \in C$ with $a^* = -a$ (where $*$ is the natural (anti)-involution on the Clifford algebra), then one can show that $< x, y >= \mathrm{Trace}(ax^*y)$ defines a Riemann form on C^+, and so this complex torus $B(V)$ is in fact a complex abelian variety. Its dual abelian variety $A(V)$ is the *Kuga-Satake abelian variety* associated to $V = (V, \psi)$, and is characterized by the identification $H^1_B(A(V), \mathbb{Z}(1)) = C^+$ as Hodge structures of weight -1. (We note that C itself also supports a Hodge structure of weight 0, induced from the original one on V, and this Hodge structure is isomorphic to $\wedge^\bullet V$).

We note that this is a transcendental construction, so it may be surprising that it has relevance for questions related to arithmetic and characteristic p. However, Deligne showed in [6] that it may be used to deduce Weil's Riemann Hypothesis for K3 surfaces, reduced mod p, from a similar assertion for abelian varieties, where it was proved earlier by Weil.

We note first that $(C^+)^{op}$ acts as a ring of endomorphisms for the complex abelian variety $A(V)$, by right multiplication on $C^+(\mathbb{R})$ (which is an action through complex linear transformations), while C^+ acts through left multiplication as a ring of endomorphisms of the underlying real vector space, preserving the integral structure, and commuting with this right action. Hence we obtain a ring homomorphism

$$C^+ = C^+(V,\psi) \to \mathrm{End}_{(C^+)^{op}}(H^1_B(A(V), \mathbb{Z}(1)))$$

which is an isomorphism of finitely generated $\mathbb{Z}$-algebras. Deligne shows that this is in fact an *isomorphism of Hodge structures of weight 0*.

To apply this to a K3 surface $X_{\mathbb{C}}$ over the complex numbers, we also choose a primitive polarization η on $X_{\mathbb{C}}$, that is, an ample class in $\mathrm{Pic}(X_{\mathbb{C}})$ (which is free abelian of finite rank) that is not divisible by any integer > 1. Now let $\eta \in H^2_B(X_{\mathbb{C}}, \mathbb{Z}(1))$ also denote the Chern class of this polarization (harmless abuse of notation, since $\mathrm{Pic}(X_{\mathbb{C}}) \subset H^2_B(X_{\mathbb{C}}, \mathbb{Z}(1))$); the corresponding primitive cohomology sublattice $P = <\eta>^\perp \subset H^2_B(X_{\mathbb{C}}, \mathbb{Z}(1)) \cong \mathbb{Z}^{21}$ supports a Hodge stucture of weight 0, with the appropriate conditions on its Hodge numbers, and taking ψ to be the negative of the cup product, we also have the non-degenerate symmetric bilinear form of signature $(19, -2)$, by the (cohomological) Hodge index theorem for K3 surfaces. Hence the above construction applies to (P, ψ); the resulting abelian variety $A(X, \eta)$ is the *Kuga-Satake variety of the polarized K3 surface* $(X_{\mathbb{C}}, \eta)$. The notation does not reflect that it depends (at least as far as we have indicated so far) on the given embedding of K into $\mathbb{C}$, though one may hope that it is in fact intrinsic to (X, η) (or perhaps at least is so after we make a finite algebraic extension of K).

André points out in [1] that a similar technique also yields a Kuga-Satake variety associated to an irreducible hyperkähler variety over $\mathbb{C}$. An irreducible hyperkähler variety X over K is defined to be a smooth, projective geometrically integral variety X such that
(i) $\dim X = 2n$ is even, and X is algebraically simply connected
(ii) $H^0(X, \Omega^2_{X/K})$ is 1-dimensional, and a generator ω yields a non-degenerate symplectic form on the cotangent sheaf $\Omega^1_{X/K}$
In particular the canonical sheaf $\Omega^{2n}_{X/K}$ is isomorphic to $\mathcal{O}_X$, generated by ω^n.

The term "irreducible holomorphic symplectic projective manifold" is also used to describe the same objects.

Here again we have that if we embed K into $\mathbb{C}$, the Hodge decomposition of $H^2_B(X_{\mathbb{C}}, \mathbb{Z}(1)$ has non-zero Hodge numbers only in degree $(1, -1)$, $(0, 0)$ and $(-1, 1)$ and the Hodge groups $H^{1,-1}$, $H^{-1,1}$ are 1-dimensional. Further, if $\eta \in \mathrm{Pic}(\overline{X}) \subset H^2_B(X_{\mathbb{C}}, \mathbb{Z}(1))$ is the class of a polarization, then

$$\psi(x, y) = -x \cup y \cup \eta^{n-2}$$

is a $\mathbb{Z}$-valued bilinear form which is non-degenerate $\otimes\mathbb{Q}$ (where we identify $H^{2n}_B(X_{\mathbb{C}}, \mathbb{Z}(n))$ with $\mathbb{Z}$); if P denotes the orthogonal complement of η, for this pairing, then by the Hodge index theorem, one knows its signature is $(m, -2)$, where the second Betti number equals $m+3$ (in the K3 case, $m = 19$). The case when m is odd works exactly like the K3 case, while there are small changes in the case when m is even, due to the slight differences in the structure of Clifford algebras of quadratic forms in odd and even dimensions (over an algebraically closed field).

A standard example of a hyperkähler manifold is $X^{[n]}$, the Hilbert scheme of 0-dimensional subschemes of length n of a K3 surface X. A deformation of a hyperkähler manifold is also one, and if $n > 1$, there are algebraic deformations of $X^{[n]}$ which are not Hilbert schemes of K3 surfaces (the local moduli space may have larger dimension).

Though our main focus is K3 surfaces, ultimately over finite fields, one of the recent developments (the work of Charles) uses an extension to the hyperkähler case to get some results for K3 surfaces X; in some situation, he finds it convenient to prove (T1) for divisors for $X^{[2]}$ instead of directly for X itself, and avoid certain technical problems that way!

4 The Tate conjecture for K3 surfaces in characteristic 0

We now sketch an argument for the Tate conjecture (T1) for divisors for K3 surfaces X over a finitely generated field K of characteristic 0. As pointed out in [1], the same argument also yields
(i) the Tate conjecture for divisors, for irreducible hyperkähler manifolds
(ii) the Tate conjecture for primitive cohomology in some even degree cohomology of a hyperkähler manifold, provided we already know the corresponding Hodge conjecture for that same primitive cohomology.

The argument depends on the following key result, proved by Deligne [6]:

Theorem *Let (X, η) be a polarized K3 surface over K. Then there is a finitely generated extension L of K and an abelian variety A over L, so that for any embedding $L \hookrightarrow \mathbb{C}$, we have the following.*

Let

$$P_B = <\eta>^{\perp} \subset H^2_B(X_{\mathbb{C}}, \mathbb{Z}(1))$$

be the Hodge structure which determines the (complex) Kuga-Satake variety $A(X_{\mathbb{C}}, \eta)$. Let

$$P_\ell \subset H^2_{et}(\overline{X}, \mathbb{Z}_\ell(1))$$

be the analogous construction of $<\eta>^{\perp}$ in étale cohomology, which is canonically identified with the analogous subspace $P_B \otimes \mathbb{Z}_\ell$ of $H^2_B(X_{\mathbb{C}}, \mathbb{Z}(1)) \otimes \mathbb{Z}_\ell$.

Then we may identify the Kuga-Satake variety $A(X_{\mathbb{C}}, \eta)$ with $A_{\mathbb{C}} = A_L \times_L \mathbb{C}$ in such a way that we have a commutative diagram

$$\begin{array}{ccc} P_B & \hookrightarrow & \mathrm{End}(H^1_B(A_{\mathbb{C}}, \mathbb{Z}(1)) \\ \downarrow & & \downarrow \\ P_B \otimes \mathbb{Z}_\ell & & \mathrm{End}(H^1_B(A_{\mathbb{C}}, \mathbb{Z}(1)) \otimes \mathbb{Z}_\ell \\ \cong\downarrow & & \downarrow\cong \\ P_\ell & \hookrightarrow & \mathrm{End}(H^1_{et}(A_{\overline{L}}, \mathbb{Z}_\ell(1)) \end{array}$$

where the vertical isomorphisms are the comparison isomorphisms between Betti and étale cohomology, AND such that
(i) *the top horizonal map is an inclusion of Hodge structres of weight 0, while*
(ii) *the bottom horizontal map is an inclusion of* $\mathrm{Gal}(\overline{L}/L)$*-modules.*

This very remarkable statement easily imples the Tate conjecture (T1) for divisors, as a consequence of the Tate conjecture for divisors on abelian vareties, as follows.

We may assume after passing to a finite extension that $\mathrm{Gal}(\overline{K}/K)$ acts trivially on $\mathrm{Pic}(\overline{X})$, and similarly that $\mathrm{Gal}(\overline{L}/L)$ acts trvially on $\mathrm{End}(\overline{A})$ (where we write $\overline{A}$ for $A_{\overline{L}}$). If P_{NS} is the orthogonal for η in $\mathrm{Pic}(\overline{X})$ for the algebraic intersection pairing, then we want to identify the $\mathrm{Gal}(\overline{L}/L)$ invariants in P_ℓ with $P_{NS} \otimes \mathbb{Z}_\ell$ (this is the content of the assertion of (T1) for divisors in this case). But the Galois invariants in $\mathrm{End}(H^1_{et}(\overline{A}, \mathbb{Z}_\ell(1)))$ are known to equal $\mathrm{End}(A_L) \otimes \mathbb{Z}_\ell = \mathrm{End}(\overline{A}) \otimes \mathbb{Z}_\ell$, by the Tate conjecture for divisors on abelian varieties (as recalled above in Section §2).

Thus, we find that that the Galois invariants in P_B must be contained in $P_B(Hodge) \otimes \mathbb{Z}_\ell$, where $P_B(Hodge)$ is the group of Hodge cycles in P_B (since $P_B \cap \mathrm{End}(A_\mathbb{C}) \subset \mathrm{End}(H^1_B(A_\mathbb{C}, \mathbb{Z}(1)))$ consists of Hodge cycles). By the Lefschetz (1,1) theorem cited earlier, $P_B(Hodge) = PB_{NS}$ is exactly the orthogonal of η in $\mathrm{Pic}(\overline{X})$ for the algebraic intersection pairing, and so we have the desired description of the Galois invariants.

If we are in charateristic p, say over a finite field K, we might first try to show that the K3 surface X over K lifts to characteristic 0, along with a polarization, over a mixed characteristic complete DVR Λ (the ring of integers in a finite extension of $\mathbb{Q}_p$), yielding a polarized K3 scheme X_Λ over Λ. If $K(\Lambda)$ is the quotient field of Λ, and we embed $K(\Lambda)$ into $\mathbb{C}$, then we may associate to the complex polarized K3 surface $X_\Lambda \times_{\mathrm{Spec}\Lambda} \mathrm{Spec}\mathbb{C}$ its Kuga-Satake abelian variety, as above.

Ideally one would like to say that the resulting complex abelian variety is "naturally" defined over $K(\Lambda)$, and has good reduction over Λ, yielding an abelian scheme over $\mathrm{Spec}\,\Lambda$, whose special fiber would be an abelian variety over the original field of characteristic p. Further, one might hope that the resulting abelian variety in characteristic p is independent of the choices, and that its etale cohomology is suitably related to the primitive cohomology of X.

From the transcendental construction of the Kuga-Satake variety given above, all this might seem rather far fetched, but maybe not completely implausible, given Deligne's remarkable result above. In some fashion, the new progress on the Tate conjecture results from trying to find a way to implement something like this.

5 The Tate conjecture for K3 surfaces over finite fields

The strategy outlined at the end of the last section is carried out by Nygaard, almost along the suggested lines, in his proof of the Tate conjecture for ordinary K3 surfaces over finite fields. He uses the notion of a *canonical lifting* to the Witt ring (i.e., the integers of the corresponding unramified extension of $\mathbb{Q}_p$ with residue field K). This lifting was constructed by Deligne and Illusie, as an analogue of the Serre-Tate canonical lifting for ordinary abelian varieties, with the formal Brauer group playing a role analogous to that of the p-divisible group.

In fact, a key step is for Nygaard to show that the Kuga-Satake variety of the canonical lift of an ordinary K3 surface over a finite field again has good, ordinary reduction, and is isogenous to the canonical lifting of its reduction. This proof uses a first form of p-adic Hodge theory, as developed in more detail later by Bloch and Kato, and a criterion for an abelian variety over a p-adic field to be isogenous to a Serre-Tate lift. In particular, Nygaard is able to show that there is a divisor class on the canonical lift of the K3 which, under the map to the endomorphisms of the Kuga-Satake abelian variety, corresponds to a lifting of geometric Frobenius. Using this, he can show that the space of (arithmetic) Frobenius invariant classes are spanned by Hodge classes, and concludes using the Lefschetz (1,1) theorem.

The proof of Nygaard and Ogus [18] uses a similar strategy, but with the construction of a certain "quasi-canonical lifting", which is characterized by the property that the action of some power of arithmetic Frobenius is similarly related to the class of an absolute Hodge cycle, in a suitable sense, which again reduces one to the Lefschetz (1,1) theorem. The construction of this quasi-canonical lift is of course involved, particularly in the case of smaller characterisics. Earlier work of Ogus on crystalline analogues of Deligne's result on the Kuga-Satake construction, and results of Berthelot, Fontaine and Messing, all play a role in the discussion.

We now turn to the new results[5].

An aspect of the arguments is the study of the moduli of polarized K3 surfaces (X, η), particularly in positive and mixed characteristics. Here the degree of the polarization of (X, η), that is, the self-intersection $\deg(\eta \cup \eta)$ (which is always an even positive integer $2d$, from Riemann-Roch) plays a role: clearly the case when $p \mid 2d$ might be expected to pose special difficulties. All the proofs use developments from p-adic Hodge theory, and make use also of the Lefschetz (1,1) theorem.

In a sense, the difference between two approaches, due to Maulik and Charles on the one hand, and due to Pera on the other, seem to consist in different levels of understanding of how to implement the Kuga-Satake construction (or at least its cohomological conseqences) for a polarized family of K3 surfaces over a given arithmetic base. Pera's approach seems to rely on a more precise version of this mixed characteristic Kuga-Satake construction, obtained by him from the persective of integral models for Shimura varieties, inspired by work of Kisin and others. The other approach, pursured by Maulik, and later Charles, makes do with less refined information on this aspect, and combines this with various other arguments from algebraic geometry, related to cycles, the Hodge conjecture, etc.

Maulik's main result is the following.

Theorem *Let K be a finite field of characteristic $p \geq 5$.*
1) *Let X be a supersingular K3 surface over K, admitting a polarization of degree $2d$ with* $\operatorname{char} K = p > 2d + 4$. *Then the Picard rank of $\overline{X}$ is 22.*
2) *Assume that semistable reduction holds for smooth projective surfaces over discrete valuation rings with residue field K. Let X be a polarized supersingular K3 surface with a polarization of degree $2d$ with p relatively prime to d. Then $\overline{X}$ has Picard rank 22.*

Maulik's argument is based on a result of Artin [2], that if we have a connected family of polarized K3 surfaces, which are all supersingular, then the Picard rank of the geometric fibers in the family does not change (the specialization maps may not be isomorphisms, but are injective with p-primary torsion cokernels). Another idea is that we would expect a supersingular K3 surface, defined over the quotient field of an equicharacteristic p discrete valuation ring, to have "potentially good reduction", since supersingularity implies a strong condition on the Galois action on ℓ-adic cohomology. So we might expect to find supersingular deformations of a supersingular K3 surface over a base consisting of a positive dimensional complete variety. Finally, it suffices to find ample divisors on the moduli space of polarized K3 surfaces in characteristic p which are supported on the locus of elliptic K3 surfaces. In characterstic 0, Maulik constructs these divisors using work of Borcherds, who had constructed certain modular

[5] The discussion here will be very sketchy, but we will attempt to give more details during the lecture, and in the later, revised version of these notes.

forms on the the period domain in question. There is then the issue of reducing these divisors mod p, and trying to show that ampleness still holds; this involves the technical hypotheses on p, as well as an understanding of comparing reductions of K3 surfaces and of their associated Kuga-Satake varieties.

Assuming all this can be carried out (which Maulik manages to do, under his hypotheses) we would find that there is a positive dimensional connected family of polarized, supersingular K3 surfaces, one of which is our original one, and another is elliptic, and so has Picard rank 22 by the old result of Artin and Swinnerton-Dyer. Hence by Artin's theorem quoted at the beginning of the previous paragraph, our given K3 also has Picard rank 22.

Charles proves the following result.

Theorem *Any supersingular K3 surface over an algebraically closed field of characteristic* $p \geq 5$ *has Picard rank 22.*

He deduces this from the following result, which is for hyperkähler varieties (also called "irreducible holomorphic symplectic manifolds" as we remarked earlier).

Theorem *Let* Y *be a smooth projective variety over a finite unramified extension* L *of* $\mathbb{Q}_p$*, with good reducion* X*, such that if we embed* L *into* $\mathbb{C}$*, then* $Y_{\mathbb{C}}$ *is a complex projective irreducible holomorphic symplectic manifold of dimension* $2n$ *with second Betti number* $b_2 \geq 5$*. Let* η *be the class of a polarization,* $d = \eta^{2n}$ *its degree,* $P \subset H^2_B(Y_{\mathbb{C}}, \mathbb{Z}(1))$ *the orthogonal for* ξ *with respect to the form defined earlier,* $< x, y >= \deg x \cup y \cup \eta^{2n-2}$*. Assume further that the discriminant of the Beauville-Bogomolov form restricted to* P *is not divisible by* p*.*

Then X *satisfies Tate's conjecture (T1) for divisors.*

The Beauville-Bogomolov form q is a certain integral quadratic form which can be defined intrisically on H^2 such that the ("usual") form defined on P using the polarization is a positive integer multiple of the restriction of q to P. For K3 surfaces q is given by the cup product.

Charles' argument seems to sidestep some of Maulik's problems about making sense of Kuga-Satake for a mixed characteristic family, by just extending a certain map using Zariski's Main Theorem, instead of trying to "construct" it via some modular information about mod p reductions of K3 surfaces, in the style of the Neron-Ogg-Shafarevich criterion for good reduction of abelian varieties. This sort of issue arises in both proofs, in order to check ampleness statements for Borcherd's style divisors on the K3 moduli space in characteristic p.

Charles' second result above does not directly give the stated result for K3 surfaces, in all cases (due to the discriminant condition). But Charles uses a trick: he applies his criterion to $X^{[2]}$ for a K3 surface X whose polarization degree is divisible by p; from a explicit description of $X^{[2]}$, and its lift $Y^{[2]}$, it is possible to then find a polarization on $Y^{[2]}$ so that it satisfies the hypothesis of the above second result, and hence the Tate conjecture for divisors follows for $X^{[2]}$ — but this easily implies it for the original X!

We now discuss the second approach, due to Pera, which yields the strongest result on the Tate conjecture for K3 surfaces over finite fields:

Theorem *Let* X *be a K3 surface over a finite field of odd characteristic. Then the Tate conjecture holds for* X*.*

Pera's method, which does not depend on the works of Nygaard or Nygaard-Ogus, is to prove that a suitable version of the Kuga-Satake construction can be made to work in positive charcteristics, to yield the desired conclusions. He shows that, given a field

K of positive odd characteristic p and a polarized K3 surface (X, η) over K, there is a finite separable extension K' of K and an abelian variety A over K' which satisfies the conclusions of the "Kuga-Satake package" (analogues of Deligne's theorem) in a strong enough sense, to yield the Tate conjecture by an argument somehwat like the one over fields of characteristic 0. What seems to make the method work is that he can show that the Kuga-Satake construction yields an étale map from (a p-integral version of) a moduli stack of quasi-polarized K3 surfaces with polarization degree not divisible by p^2 (but which may be divisible by p), to a certain integral model of the Shimura variety associated to the family of Kuga-Satake abelian varieties; this latter intergal model is seen to possess good mixed characteristic properties (details to be found in Pera's other works, and based on Kisin's work – Pera's paper gives additional references), and by étaleness, induce similar good properties of the original K3 family, in a fashion compatible with the Kuga-Satake construction over $\mathbb{Q}$.

We end our account with a result due to Lieblich, Maulik and Snowden:

Theorem *Let K be a finite field of characteristic $p \geq 5$.*
(1) *There are only finitely many isomorphism classes of K3 surfaces over K which satisfy the Tate conjecture over $\overline{K}$.*
(2) *If there are only finitely many isomorphism classes of K3 surfaces over the quadratic extension K' of K, then every K3 surface over K satisfies the Tate conjecture over K'.*

As a consequence, we see that there are only finitely many isomorphism classes of K3 surfaces over any given finite field K of characteristic $p \geq 5$. The proof of this result, in both directions, is quite interesting as well.

References

1. André, Y.: On the Shafarevich and Tate conjectres for hyperkähler varieties. Math. Ann. **305**, 205-248 (1996)
2. Artin, M.: Supersingular K3 surfaces. Ann. Sci. École Norm. Sup. **7**(4), 543-567 (1974)
3. Artin, M., Swinnerton-Dyer, H.P.F.: The Tate-Safarevic conjecture for pencils of elliptic curves on K3 surfaces. Invent. Math. **20**, 279-296 (1973)
4. Chai, C.-L., Oort, F.: Moduli of abelian varieties and p-divisible groups. In Arithmetic geometry, H. Darmon, D. A. Ellwood, B. Hassett and Y. Tschinkel, eds., Clay Math. Proceedings **8**, 439-534 (2009)
5. Charles, F.: The Tate conjecture for K3 surfaces over finite fields. To appear in Invent. Math. (DOI 10.1007/s00222-012-0443-y)
6. Deligne, P.: La conjecture de Weil pour les surfaces K3. Invent. Math. **15**, 206-226 (1972)
7. Deligne, P.: Théoriè de Hodge II. Publ. Math. IHES **40**, 5-57 (1971)
8. Fulton, W.: Intersection theory. Second ed., Ergeb. Math. (3) 2, Springer-Verlag (1998)
9. Grothendieck, A.: Standard conjectures on algebraic cycles. In Algebraic Geometry, (Internat. Colloq., Tata Inst. Fund. Res., Bombay, 1968), Oxford Univ. Press, London (1969) 193-199
10. Hartshorne, R.: Algebraic Geometry. Grad. Texts in Math. No.52, Springer (1977)
11. Kleiman, S.: The standard conjectures. In Motives, U. Jannsen, S. Kleiman, J.-P. Serre (eds.), Proc. Sym. Pure Math. Vol. 55, Part 1, A.M.S. (1994) 3-20.
12. Lieblich, M., Maulik, D.: A note on the cone conjecture for K3 surfaces in positive characteristics. Preprint at arXiv:1102.3377
13. Lieblich, M., Maulik, D., Snowden, A.: Finiteness of K3 surfaces and the Tate conjecture. Preprint at arXiv:1107.1221
14. Maulik, D.: Supersingular K3 surfaces for large primes. preprint at arXiv:1203.2889
15. J. S. Milne, J.S.: Étale Cohomology. Princeton Math. Ser. 33, Princeton Univ. Press(1980)
16. Milne, J.S.: Lecture Notes on Étale Cohomology. available for download at http://www.jmilne.org/math/CourseNotes/lec.html
17. Nygaard, N.: The Tate conjecture for ordinary K3 surfaces over finite fields. Invent. Math. **74**, 213-237 (1983)

18. Nygaard, N., Ogus, A.: Tate's conjecture for K3 surfaces of finite height. Ann. Math. **122**(3), 461-507 1985
19. Pera, K.M.: The Tate conjecture for 3 surfaces in odd characteristics. Preprint arXiv:1301.6326
20. Tate, J.: Algebraic cohomology classes. Lecture Notes, A.M.S. Summer Inst. in Algebraic Goemetry (Woods Hole, 1964) 166-190
21. Tate, J.: Conjectures on Algebraic Cycles in ℓ-adic cohomology. In Motives, U. Jannsen, S. Kleiman, J.-P. Serre (eds.), Proc. Sym. Pure Math. Vol. 55, Part 1, A.M.S. (1994) 71-84
22. Voisin, C.: Hodge Theory and Complex Algeraic Geometry, Vols I, II. Cambridge Univ. Studies in advanced math. Vols. 76, 77. Cambridge univ. Press (2003)