UNIVERSITÀ DEGLI STUDI DI ROMA "LA SAPIENZA"
Dipartimento di Matematica, Dipartimento di Metodi e Modelli Matematici per le Scienze Applicate
ISTITUTO NAZIONALE DI ALTA MATEMATICA "FRANCESCO SEVERI"

RENDICONTI DI MATEMATICA

E DELLE SUE APPLICAZIONI

Serie VII - Volume 28 - Fascicolo II - 2008

Le modalità per la presentazione dei lavori e le condizioni di vendita si trovano alla web page.

Submission of papers and subscription rate information could be found on the web page.

ADDRESS:
RENDICONTI DI MATEMATICA
DIPARTIMENTO DI MATEMATICA - PIAZZALE A. MORO, 2 - 00185 ROMA, ITALIA

Fax.: +39 06 49913052
e-mail: rendmat@mat.uniroma1.it
web page: www.dmmm.uniroma1.it/~rendiconti

Direttore Responsabile: Prof. ALESSANDRO SILVA

Autorizzazione del Tribunale di Roma del 22-11-2005 (n. 456/05)

Composizione tipografica effettuata con il programma TEX © dalla
CompoMat – Loc. Braccone – 02040 Configni (RI) – Telefax 39 0746 672240

Stampato a cura del *Centro Stampa d'Ateneo* dell'Università di Roma *"La Sapienza"*

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 117–121

Integration on fuzzy subsets of the unit circle

VITTORIO CAFAGNA – GIANLUCA CATERINA

Vittorio Cafagna suddendly died in Paris on January 5, 2007.
He is deeply regretted by those who admired his mathematical curiosity and intuition

ABSTRACT: *Let* $\mathbb{T}$ *be the unit circle and* $(\mathbb{T}, \mathscr{B}, \mu)$ *be the probability space defined by the Borel* σ*-ring* $\mathscr{B}$ *and the normalized Lebesgue measure* μ*. Consider the collection* $L^1(\mathbb{T}, \mathbb{R}) \supset \mathscr{F} = \{\mathscr{E} : 0 \leq \mathscr{E}(t) \leq 1\}$*. According to Zadeh's philosophy, members of* $\mathscr{F}$ *will be called measurable fuzzy subsets of* $\mathbb{T}$ *and thought of as generalized subsets. Let* $f \in L^1(\mathbb{T}, \mathbb{R})$ *be a summable function. Define the integral of* f *on a measurable fuzzy subset* $\mathscr{E}$ *as* $\int_{\mathscr{E}} f d\mu = \int_{\mathbb{T}} \mathscr{E} f d\mu$*. In this note we prove that there exist sequences* R_n *of collections of* n *arcs, such that* $\int_{\mathscr{E}} f d\mu = \lim_{n\to\infty} \int_{R_n} f d\mu$*. The main ingredient of the convergence result is an old and remarkable theorem due, independently, to Friedman and Ghizzetti.*

1 – Introduction

Let us denote by $\mathbb{I}$ the unit interval $[0, 1]$ and by $\mathbb{Z}_2 = \{0, 1\}$ its boundary. Let X be a set and $2^X = \mathbb{Z}_2^X$ its power set. Consider now $\mathbb{I}^X$, obviously a superset of $\mathbb{Z}_2^X$. According to Zadeh [3], we call members of $\mathbb{I}^X$ *fuzzy* subsets of X. The philosophy of the definition being that, while a function $E : X \to \mathbb{Z}_2$ tells you if $x \in X$ belongs or not to the subset $E \subset X$, a function $\mathscr{E} : X \to \mathbb{I}$ tells you that $x \in X$ *belongs to* $\mathscr{E}$ *up to a certain degree* given by $0 \leq \mathscr{E}(x) \leq 1$. The collection $\mathbb{I}^X$ of fuzzy subsets of a set X will be denoted by $\mathscr{F}_X$, or simply $\mathscr{F}$

KEY WORDS AND PHRASES: *Fuzzy measures – Fourier expansions.*
A.M.S. CLASSIFICATION: 42A16, 42A99, 26E50, 28E10

when X is understood, and members of the subcollection $\mathbb{Z}_2^X$ will be called *crisp* subsets. Accordingly, we will denote by $\mathscr{C}_X$, or simply $\mathscr{C}$, the collection of crisp subsets, when needed.

Let now (X, S, μ) be a measure space, μ a positive measure and S the σ-ring of measurable subsets. Denote by $\mathscr{S}$ the collection of summable fuzzy subsets of X (summable meaning of course summable as functions from X to $\mathbb{I}$). $\mathscr{S}$ is a ring with respect to the operations $\bigvee, \bigwedge$ defined by $\mathscr{E} \bigvee \mathscr{G} = \sup_{x\in X}\{\mathscr{E}(x),\, \mathscr{G}(x)\}$ and $\mathscr{E} \bigwedge \mathscr{G} = \inf_{x\in X}\{\mathscr{E}(x),\, \mathscr{G}(x)\}$. Remark that $\bigvee, \bigwedge$ reduce, respectively, to $\bigcup, \bigcap$ in the case of crisp subsets, so that $\mathscr{S}$ is a superring of S. Therefore we feel free to re-christen the operations $\bigvee, \bigwedge$ as $\bigcup, \bigcap$ on the whole ring $\mathscr{S}$. It is easy to prove that $(\mathscr{S}, \bigcup, \bigcap)$ is also *σ-complete* in the sense that it is closed with respect to countable intersection and union of a sequence $\mathscr{E}_n$, defined, respectively, as $\cup_n \mathscr{E}_n = \limsup_{n\to\infty} \mathscr{E}_n$ and $\cap_n \mathscr{E}_n = \liminf_{n\to\infty} \mathscr{E}_n$. Moreover, one can define on $\mathscr{S}$ an involution $\mathscr{E} \to \bar{\mathscr{E}}$ by $\bar{\mathscr{E}} = 1 - \mathscr{E}$, which reduces to taking the complement on the subring of crisp subsets. It is worth remarking that the ring $\mathscr{S}$ is not Boolean, due to the failure of the excluded middle axiom. In fact, $\mathscr{E} \cap \bar{\mathscr{E}} = \emptyset \Leftrightarrow \mathscr{E}$ is a crisp subset.

One could wonder how much of classical measure and integration theory can be carried over to such a generalized setting. It might come as a surprise that, in the very special case of fuzzy subsets of the unit circle $\mathbb{T}$, one can prove that the integral of a summable function on a fuzzy subset can be defined as the limit of integrals of the summable function on a sequence of crisp subsets, actually families of arcs. The convergence result rests on a remarkable theorem due, independently, to Friedman [1] and Ghizzetti [2].

In this note we describe this theorem and show how to use it to derive the convergence theorem.

2 – Fuzzy measure spaces

Let (X, S, μ) be a measure space, μ a positive measure, and $L^1(X, \mathbb{R})$ the space of summable functions on X. Let $\mathscr{F} = \mathbb{I}^X$ the collection of the fuzzy subsets of X, as defined above, and $\mathscr{S} = \mathscr{F} \cap L^1(X, \mathbb{R})$ the σ-ring of summable fuzzy subsets of X. One can extend in a natural way the notion of measure of a set to summable fuzzy sets: define, for $\mathscr{E} \in \mathscr{S}$,

$$\mu_{\mathscr{F}}(\mathscr{E}) = \int_X \mathscr{E} d\mu.$$

Trivially $\mu_{\mathscr{F}} \mathscr{E} = \mu\mathscr{E}$ if $\mathscr{E}$ is crisp and it is also easy to verify that the defining properties of a measure are still valid on $\mathscr{S}$. In fact, if $\{\mathscr{E}_n\}_{n\in\mathbb{N}} \subset \mathscr{S}$ with $\mathscr{E}_n \cap \mathscr{E}_m = \emptyset$ for $n \neq m$, one has $\mu(\cup_n \mathscr{E}_n) = \sum_n \mu(\mathscr{E}_n)$.

One is therefore naturally tempted to state the following definitions:

Definition 1 The triple $(X, \mathscr{S}, \mu_{\mathscr{F}})$ is called the fuzzy measure space associated to the measure space (X, S, μ).

Definition 2 Let (X, S, μ) be a measure space, $(X, \mathscr{S}, \mu_{\mathscr{F}})$be the associated fuzzy measure space, $f \in L^1(X, \mathbb{R})$ a summable function (with respect to μ) and $\mathscr{E} \in \mathscr{S}$ a fuzzy subset of X. The quantity

$$\int_{\mathscr{E}} f d\mu_{\mathscr{F}} = \int_X \mathscr{E} f d\mu$$

is the integral of f over $\mathscr{E}$ with respect to $\mu_{\mathscr{F}}$.

Let us stress that a fuzzy measure space is **not**, in accordance with the red herring principle, a measure space: the formal σ-ring $\mathscr{S}$ contains not only sets but also functions (*a fortiori* a fuzzy measure is **not** a measure.) It is just a formal mathematical object which behaves somewhat like a measure space. To what extent, can be considered as a wide-open question. Even more, the integral of a function over a fuzzy set is **not even formally** the Lebesgue integral with respect to the fuzzy measure. It is not even clear how to give a sound definition of a characteristic function of a fuzzy subset, if any. It is just a procedure to associate a positive real number to the couple $(\mathscr{E}, f)$ formed by a fuzzy subset and a summable function (with respect to the *bona fide* measure μ.) That it behaves somewhat like a Lebesgue integral is due to the fact that is actually a Lebesgue integral (with respect to the *bona fide* measure μ) of a product of functions. It might be safe to bet that every proposition of standard measure and integration theory which does not use in a crucial manner the excluded middle axiom has a painless extension to the fuzzy setting. Still, the theoretical framework is rather vague and unclear (not to say *fuzzy*.) Therefore it seems a good (and also a rather intriguing) surprise that, in the special case of the unit circle, one has that the integral of a summable function over a fuzzy subset is the limit of the integrals of the function over a sequence of arcs.

This result is a byproduct of the results of Friedman and Ghizzetti on the Fourier coefficients of a bounded function which we shall describe in the next section.

3 – Fourier coefficients of a bounded function: a theorem by Friedman and Ghizzetti

Let $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ be the unit circle, $\mathscr{B}$ be the Borel σ-ringand μ the Lebesgue measure, normalized so that $\mu(\mathbb{T}) = 1$. Let $L^1(\mathbb{T})$ be the space of summable complex-valued functions on $\mathbb{T}$ and let

$$c_k(f) = \int_{\mathbb{T}} f e^{ikt} d\mu$$

denote the k-th Fourier coefficient $(k \in \mathbb{Z})$ of a $f \in L^1(\mathbb{T})$.

We are now in a position to state a theorem due to Friedman [1] and Ghizzetti [2])

Theorem 3. *Let $f \in L^1(\mathbb{T})$ be real-valued and such that $0 \leq f(t) \leq 1$. Then there exist a sequence $\{R_n\}_{n\in\mathbb{N}}$, each R_n being the characteristic function of a set of n arcs, such that*

$$c_k(f) = c_k(R_n), \ \ |k| < n.$$

For the proof, we refer to the original quoted papers. A few words on the different attitudes: while the statements in both the papers of Friedman and Ghizzetti are essentially very similar, the proofs are very different. The proof by Friedman is substantially simpler, relying only on differential calculus. The proof by Ghizzetti is a much more sophisticated one, using some deep intuitions on rather intriguing arguments of complex analysis. Moreover, the proof by Ghizzetti is a constructive one, describing a procedure which, given a bounded function, allows to define step by step sequences (actually uncountably many) of sets of arcs with the desired property. On the other side, it seems possible, at least in principle and with a lot of work, to foresee a generalization of Friedman's proof to compact Riemannian manifolds, while the most that one can expect from Ghizzetti's is maybe a generalization, with even harder work, to some well-behaved compact abelian group. The authors of this note are actually trying to work in both these perspectives.

4 – Integration over fuzzy subsets of the unit circle

Let $(\mathbb{T}, \mathscr{B}, \mu)$ be the probability space defined by the normalized Lebesgue measure on the Borel σ-ring on the unit circle and $(\mathbb{T}, \mathscr{S}, \mu_{\mathscr{F}})$ the associated fuzzy measure space. The immediate consequence of the Friedman-Ghizzetti theorem for the argument of this note, restated in fuzzy parlance, is

Theorem 4. *For every fuzzy subset $\mathscr{E} \in \mathscr{S}$ of the unit circle, there exists a sequence of crisp subsets R_n such that $c_k\mathscr{E} = c_k(R_n)$, $|k| < n$.*

This result makes possible a proof of the following convergence theorem announced in the Introduction:

Theorem 5. *Let $\mathscr{E} \in \mathscr{S}$ be a fuzzy subset of the unit circle and $f \in L^1(\mathbb{T})$ a real-valued function. Then there exists a sequence R_n of crisp subsets such that*

$$\int_{\mathscr{E}} f d\mu = \lim_{n\to\infty} \int_{R_n} f d\mu.$$

PROOF. By definition of the integral of a function f over a fuzzy subset $\mathscr{E}$

$$\int_{\mathscr{E}} f d\mu = \int_{\mathbb{T}} \mathscr{E} f d\mu.$$

By the density of trigonometric polynomials in $L^1(\mathbb{T})$ one can write $f(t) = \lim_{n\to\infty} \sum_{k=0}^{n-1} \gamma_k e^{ikt}$ and the last expression becomes

$$\int_{\mathbb{T}} \mathscr{E}(t)\big(\lim_{n\to\infty} \sum_{k=0}^{n-1} \gamma_k e^{ikt}\big) d\mu = \lim_{n\to\infty} \sum_{k=0}^{n-1} \gamma_k \int_{\mathbb{T}} \mathscr{E}(t) e^{ikt} d\mu = \lim_{n\to\infty} \sum_{k=0}^{n-1} \gamma_k c_k \mathscr{E}$$

By the Friedman-Ghizzetti theorem, this is equal to

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} \gamma_k c_k(R_n) = \lim_{n\to\infty} \sum_{k=0}^{n-1} \gamma_k \int_{\mathbb{T}} R_n(t) e^{ikt} d\mu = \lim_{n\to\infty} \int_{\mathbb{T}} R_n(t) \sum_{k=0}^{n-1} \gamma_k e^{ikt} d\mu =$$

$$= \int_{\mathbb{T}} \lim_{n\to\infty} \big(R_n(t) \sum_{k=0}^{n-1} \gamma_k e^{ikt}\big) d\mu = \int_{\mathbb{T}} \lim_{n\to\infty} R_n(t) \lim_{n\to\infty} \sum_{k=0}^{n-1} \gamma_k e^{ikt} d\mu =$$

$$= \int_{\mathbb{T}} \lim_{n\to\infty} R_n(t) f(t) d\mu = \lim_{n\to\infty} \int_{\mathbb{T}} R_n(t) f(t) d\mu = \lim_{n\to\infty} \int_{R_n} f d\mu.$$

REMARK 1 The theorem above can be also rephrased as: *the probability measure on $\mathbb{T}$ defined by $\mathscr{E} d\mu$ is the is the weak*-limit of the sequence of probability measures defined by $R_n d\mu$.*

REFERENCES

[1] B. FRIEDMAN: *Fourier coefficients of bounded functions, Bull. Amer. Math. Soc.*, **47** (1941) 84-92.

[2] A. GHIZZETTI: *Sui coefficienti di Fourier di una funzione limitata, compresa fra limiti assegnati*, Ann. Scuola Norm. Sup. Pisa, **4** (1950) 131-156.

[3] L. A. ZADEH: *Fuzzy sets*, Information and Control, **8** (1965) 338-353.

Lavoro pervenuto alla redazione il 26 ottobre 2007
ed accettato per la pubblicazione il 21 novembre 2007.
Bozze licenziate il 30 settembre 2008

INDIRIZZO DEGLI AUTORI:

Vittorio Cafagna – DMI, Department of Mathematics and Informatics – University of Salerno – via Ponte Don Melillo, 84084 Fisciano (SA) – Italy

Gianluca Caterina – Department of Mathematics – Tufts University – Cambridge, MA 0213, USA
E-mail: gianluca.caterina@tufts.edu

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 123–131

Surfaces with a family of nongeodesic biharmonic curves

J. MONTERDE

ABSTRACT: *The only surface whose level curves of the Gauss curvature are nongeodesic biharmonic curves and such that the gradient lines are geodesics is, up to local isometries, the revolution surface defined by Caddeo-Montaldo-Piu.*

1 – Introduction

In a recent paper ([2]) the authors study the notion of biharmonic curves on surfaces. If we consider isometric immersions $\gamma : I \to S$ from an interval I to a surface S, then the bienergy functional is defined by

$$E_2(\gamma) = \frac{1}{2}\int_S |\tau_\gamma|^2 \, dv,$$

where $\tau_\gamma = \nabla_{\dot\gamma}\dot\gamma$ is the tension field associated to the curve γ. A curve is called biharmonic if it is a critical point of the bienergy functional.

In the cited paper it is proved that along a nongeodesic biharmonic curve the Gauss curvature is constant and equal to the square of the geodesic curvature. Therefore, nongeodesic biharmonic curves are level curves of the Gauss curvature.

Moreover, biharmonic curves on revolution surfaces also are therein studied. In particular the unique revolution surfaces with all parallels nongeodesic biharmonic curves are determined.

KEY WORDS AND PHRASES: *Biharmonic curves – Surfaces of revolution.*
A.M.S. CLASSIFICATION: 58E20

The two conditions: nonvanishing constant geodesic curvature and Gauss curvature equal to the square of the geodesic curvature along level curves seem to be hard conditions. Apart from the previously cited revolution surface, the authors in [2] are able to find just some few such curves in, for instance, revolution surfaces with constant Gauss curvature.

In this note, we first determine the local expression of the metric tensor of a two-dimensional Riemannian manifold whose level curves of the Gauss curvature are nongeodesic biharmonic curves.

The coefficients of the metric only depend on the function which assigns to each level curve its constant geodesic curvature and on another function on the same parameter transversal to the level curves.

Since the Gauss curvature is positive, then the two-dimensional Riemannian manifolds can be locally realized as regular surfaces (see [4]). If in addition we ask for the gradient lines of the Gauss curvature to be geodesics, then the only surface, up to local isometries, is the revolution surface defined in [2]. A final example shows that this last condition, orthogonal lines are geodesics, is necessary.

2 – Surfaces of revolution for which all parallels are biharmonic curves

PROPOSITION 1. (*See* [2]) *Let* $\gamma : I \to (M^2, g)$ *be a differentiable curve in a surface* M^2. *Then, if* γ *is a nongeodesic biharmonic curve, along* γ *the Gauss curvature is constant, positive and equal to the square of the geodesic curvature of* γ.

So, a nongeodesic biharmonic curve, γ, is characterized by

$$\begin{cases} k_g(t) & = \text{constant} \neq 0, \\ k_g^2(t) & = K(\gamma(t)), \end{cases}$$

for all $t \in I$ and where k_g denotes the geodesic curvature of γ and K denotes the Gauss curvature.

THEOREM 1. (*See* [2]) *Let* $M^2 \subset \mathbb{R}^3$ *be a surface of revolution obtained by rotating the arc length parametrized curve* $\alpha(v) = (f(v), 0, g(v))$ *in the* xz*-plane around the* z*-axis. Then all parallels of* M *are biharmonic curves if and only if either*

1. f *is constant and* M *is a right circular cylinder or*
2. $f(v) = \pm c\sqrt{v}$ *and*

$$g(v) = v\sqrt{\frac{4v - c^2}{4v}} - \frac{c^2}{8} \ln\left(8v + 8v\sqrt{\frac{4v - c^2}{4v}} - c^2\right) + c_1,$$

where c *and* c_1 *are positive constants.*

Remark 1. The surfaces introduced in Theorem 1, (2), will be called CMP-revolution surfaces. If we consider the parametrization

$$\overrightarrow{\mathbf{x}}(u,v)=\Big(f(v)\cos\frac{u}{c},f(v)\sin\frac{u}{c},g(v)\Big),$$

then, a simple computation shows that the coefficients of the metric are independent of the values of the two constants c and c_1:

$$g_{11}(u,v)=v,\qquad g_{12}(u,v)=0,\qquad g_{22}(u,v)=1.$$

Therefore any pair of CMP-revolution surfaces are isometrics.

Let us consider the parallel generated by $(f(v),0,g(v))$. The geodesic curvature of the parallel is $-\frac{1}{2v}$, the sign depends on the orientation, and the Gauss curvature is $\frac{1}{4v^2}$.

Another parametrization of the revolution surface can be obtained by changing $v=\frac{c^2}{4}\cosh^2(t)$ and modifying the constants $c=\sqrt{2}a, c_1=c^2(\frac{c_2}{8}+\frac{1}{4}\ln c)$. The new parametrization of the generating curve is

$$f(t)=\pm a^2\cosh t,\quad g(t)=-\frac{a^2}{2}(t-2\sinh(2t))+c_2.$$

3 – Two-dimensional Riemannian manifolds with a family of nongeodesic biharmonic curves

The surface in fig. 1 is the only surface of revolution with nongeodesic biharmonic parallel lines. The natural question is to ask if there are more surfaces, obviously not of revolution, with a family of coordinate lines which are nongeodesic biharmonic curves.

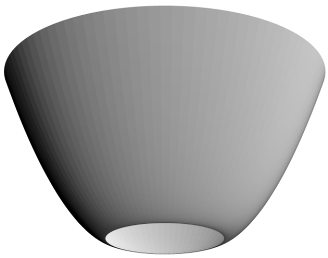

Fig. 1: Plot of a piece of the unique revolution surface with nongeodesic biharmonic parallel lines, for $c=1$ and $c_1=0$.

In any surface the level curves of the Gauss curvature define a foliation, maybe degenerated, on it. At same time, the integral curves of the gradient vector field are orthogonal to the level curves. We are interested in studying the case when the level curves are curves with non zero constant geodesic curvature whose square is the value of the Gauss curvature.

We will use the notation i to denote partial derivatives with respect to the variable u_i. Thus, $g_{12,1}$ denotes $\frac{\partial g_{12}}{\partial u_1}$.

PROPOSITION 2. *Let (M^2, g) be a two-dimensional Riemannian manifold such that the level curves of the Gauss curvature are nongeodesic biharmonic curves. Then, for any $p \in M$, regular point of the Gauss curvature, there exists a parametrization of a neighborhood of p, $V \subset M$, $\overrightarrow{\mathbf{x}} : U \to V \subset M$, such that all the coordinate lines $v = v_0$, v_0 constant, are nongeodesic biharmonic curves, and the coefficients of the metric are*

$$
\begin{aligned}
g_{11} &\equiv 1, \\
g_{12}(u,v) &= \frac{\sqrt{2}m(v)\left(\sin(\sqrt{2}k(v)(u-n(v))) + \sin(\sqrt{2}\,n(v)k(v))\right)}{2} + \\
&\quad + \frac{uk'(v)}{2k(v)}, \\
g_{22}(u,v) &= g_{12}^2(u,v) + \left(\frac{g_{12,1}(u,v)}{k(v)}\right)^2,
\end{aligned}
\tag{3.1}
$$

where $m(v) = \sec(\sqrt{2}k(v)n(v))\left(1 - \frac{k'(v)}{2k^2(v)}\right)$ and where $k(v_0)$ is the geodesic curvature of the coordinate line $v = v_0$.

Reciprocally, if a metric is of the kind 3.1*, then the level curves of the Gauss curvature are nongeodesic biharmonic curves.*

PROOF. Let $\alpha : I \to S$ be the gradient line of the Gauss curvature passing trough the point p, and let us suppose that it is parametrized by arc-length. Since p is a regular point for the Gauss curvature, there is a neighborhood of p, V, such that all points $q \in V$ are also regular. For each point $\alpha(v) \in V$, let σ^v be the level curve passing trough $\alpha(v)$ and parametrized by arc-length.

Finally, let us consider $\overrightarrow{\mathbf{x}} : U \to M$ defined by $\overrightarrow{\mathbf{x}}(u,v) = \sigma^v(u)$.

Since all the coordinate lines $v = v_0$ are parametrized by arc-length, then the coefficient g_{11} of the metric is equal to 1.

The geodesic curvature of a curve $\overrightarrow{\mathbf{x}}(u_1(t), u_2(t))$, not necessarily parametrized by the arc-length, can be computed from the formula (see [3], formula (49.7))

$$
\begin{aligned}
k_g &= \frac{1}{||\alpha'||^3} < \frac{D\alpha'}{dt}, \alpha' \wedge (N \circ \alpha) > \\
&= \frac{\sqrt{g_{11}g_{22} - g_{12}^2}}{||\alpha'||^3}\left((u_1'' + \sum_{j,k=1}^{2}\Gamma_{jk}^1 u_j' u_k')u_2' - (u_2'' + \sum_{j,k=1}^{2}\Gamma_{jk}^2 u_j' u_k')u_1'\right).
\end{aligned}
$$

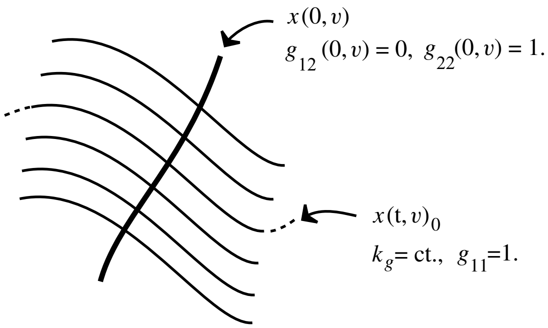

Fig. 2: Schematic description of the definition of the parametrization.

Therefore, the geodesic curvature of a coordinate line $t \to \overrightarrow{\mathbf{x}}(t, v_0)$ reduces to

$$k_g(t) = -\Gamma^2_{11}(t, v_0)\sqrt{g_{22}(t, v_0) - g^2_{12}(t, v_0)} = -\frac{g_{12,1}(t, v_0)}{\sqrt{g_{22}(t, v_0) - g^2_{12}(t, v_0)}}. \tag{3.2}$$

Since we are supposing that the geodesic curvature of a coordinate line $v = v_0$ is constant, then $k_g(t) = k(v_0)$, where k is the function assigning to each coordinate line $v = v_0$ its geodesic curvature.

From eq. 3.2 we get

$$g_{22}(t, v) = g^2_{12}(t, v) + \left(\frac{g_{12,1}(t, v)}{k(v)}\right)^2. \tag{3.3}$$

Note that the area element reduces to

$$\sigma := \sqrt{g_{11}g_{22} - g^2_{12}} = \frac{g_{12,1}(t, v)}{k(v)}. \tag{3.4}$$

The computation of the Gauss curvature by the Gauss formula

$$K = -\frac{1}{g_{11}}((\Gamma^2_{12})_1 - (\Gamma^2_{11})_2 + \Gamma^1_{12}\Gamma^2_{11} + \Gamma^2_{12}\Gamma^2_{12} - \Gamma^1_{11}\Gamma^2_{12} - \Gamma^2_{11}\Gamma^2_{22}), \tag{3.5}$$

gives a simple expression:

$$K = -\frac{g_{12,111}(t, v) + k^2(v)g_{12,1}(t, v) - k(v)k'(v)}{g_{12,1}(t, v)}.$$

Condition $K(t, v) = k^2(v)$ implies

$$g_{12,111}(t, v) + k(v)\left(2k(v)g_{12,1}(t, v) - k'(v)\right) = 0. \tag{3.6}$$

From eq. 3.4, eq. 3.6 can be rewritten in terms of the area element σ as

$$\frac{\partial^2\sigma}{\partial t^2}(t,v) + 2k^2(v)\sigma(t,v) - k'(v) = 0. \tag{3.7}$$

Note that the differential equation is of the kind $f'' + a^2 f = b$ whose general solution can be written as $f(t) = C_1 \ \cos(at - C_2) + \frac{b}{a^2}$. Therefore, the general solution of eq. 3.7 is

$$\sigma(t,v) = m(v)\cos\left(\sqrt{2}\ k(v)\ (t - n(v))\right) + \frac{k'(v)}{2k^2(v)},$$

or some functions $m(v)$ and $n(v)$.

Along the curve $v \to \overrightarrow{\mathbf{x}}(0,v)$ the area element, $\sigma(0,v)$, is equal to 1, therefore,

$$m(v) = \sec\left(\sqrt{2}k(v)n(v)\right)\left(1 - \frac{k'(v)}{2k^2(v)}\right).$$

Therefore, from eq. 3.4,

$$g_{12}(t,v) = \frac{\sqrt{2}k(v)m(v)\sin(\sqrt{2}k(v)(t-n(v))) + tk'(v)}{2k(v)} + c(v),$$

for some functions $n(v)$ and $c(v)$.

Since the curve $v \to \overrightarrow{\mathbf{x}}(0,v)$ is orthogonal to all coordinate lines $v = v_0$, then $g_{12}(0,v) = 0$. This implies that

$$c(v) = \frac{m(v)\sin(\sqrt{2}\ n(v)k(v))}{\sqrt{2}}.$$

Reciprocally, note that if the coefficients of a metric are of the kind 3.1, then the Gauss curvature is $K(t,v) = k^2(v)$. Therefore, the coordinate curves $t \to \overrightarrow{\mathbf{x}}(t,v_0)$ are level curves of the Gauss curvature. Moreover, since the geodesic curvature of the curves $t \to \overrightarrow{\mathbf{x}}(t,v_0)$ is $k(v_0)$, then they are nongeodesic biharmonic curves.

REMARK 2. In the case $k' \equiv 0$, then the Gauss curvature is constant. Minding's theorem states that, up to local isometries, the models for surfaces with constant Gauss curvature are the revolution surfaces with constant Gauss curvature. It is possible to obtain parametrizations with coefficients of the metric like in the statement of Proposition 2. See the final Example 3.1.

REMARK 3. Note that in the CMP-revolution surfaces, the gradient lines of the Gauss curvature, ie., the meridian curves, are geodesic curves. So, we shall ask for all gradient lines being geodesic curves, i.e., $\frac{\mathrm{grad}K}{|\mathrm{grad}K|}$ is a geodesic vector field. As it is pointed out in [1], Section 3, this condition is equivalent to the assertion that the regular levels of K are parallel, or to the eiconal equation for K: grad(|grad$K$|) is a multiple of grad$K$.

THEOREM 2. *Let (M^2, g) be a two-dimensional manifold with $|\mathrm{grad}K|(p) \neq 0$ for all $p \in M$ and such that the level curves of the Gauss curvature are nongeodesic biharmonic curves, then (M^2, g) is locally isometric to the* CMP-*revolution surface if and only if $\frac{\mathrm{grad}K}{|\mathrm{grad}K|}$ is a geodesic vector field.*

PROOF. In the CMP-revolution surface the gradient lines of the Gaussian curvature are the meridian lines and they are geodesics, so, grad(|grad$K$|) is a geodesic vector field.

Reciprocally, let us consider one of the parametrizations, $\overrightarrow{\mathbf{x}}$, given by Proposition 2. Gradient lines are orthogonal to level curves, i.e., to the coordinate lines with $\overrightarrow{\mathbf{x}}_1$ as tangent vector. Therefore, any gradient line, $\beta(t) = \overrightarrow{\mathbf{x}}(u(t), v(t))$, parametrized by arc-length, has as tangent vector

$$\frac{k}{g_{12,1}}\left(-g_{12}\overrightarrow{\mathbf{x}}_1 + \overrightarrow{\mathbf{x}}_2\right).$$

An straightforward computation of its geodesic curvature using eq. 3.2 with

$$u'(t) = -\frac{k(v(t))g_{12}(u(t), v(t))}{g_{12,1}(u(t), v(t))}, \qquad v'(t) = \frac{k(v(t))}{g_{12,1}(u(t), v(t))},$$

gives us

$$k_g^\beta(t) = \frac{g_{12,11}}{g_{12,1}}(u(t), v(t)).$$

Now, by eq. 3.1, $k_g^\beta(t) \equiv 0$ if and only if

$$-\frac{1}{\sqrt{2}}\frac{\cos(\sqrt{2}\ k(v(t))\ (u(t) - n(v(t))))}{\sin(\sqrt{2}\ k(v(t))\ n(v(t)))}\left(2k^2(v(t)) - k'(v(t))\right) = 0.$$

If all the gradient lines are geodesics, then $2k^2(v) - k'(v) = 0$ for all v. Therefore,

$$k(v) = -\frac{1}{2v + a}.$$

A simple change of parameter v allows to put

$$k(v) = \frac{1}{2v}.$$

Now, the coefficients of the metric are

$$g_{11} = 1, \qquad g_{12}(u, v) = -\frac{u}{2v}, \qquad g_{22}(u, v) = 1 + \frac{u^2}{4v^2}.$$

A change of parameter $u \to u\sqrt{v}$ transform them into

$$g_{11} = v, \qquad g_{12} = 0, \qquad g_{22} = 1,$$

the same coefficients than the ones of the CMP-revolution surface. Therefore both surfaces are locally isometric.

3.1 – Necessary condition

The condition: "orthogonal lines to the nongeodesic biharmonic curves are geodesics" is necessary. Let us show an example where the orthogonal lines are not geodesic curves.

The example can be built using the sphere and the parallel of latitude $\frac{\pi}{4}$. It is already known that it is a nongeodesic biharmonic curve on the sphere. The image of this parallel under a rotation around the y-axis, an isometry, gives another nongeodesic biharmonic curve. So, we can construct a uniparametric family of nongeodesic biharmonic curves on the sphere.

Let us denote by $R^y_\theta : \mathbb{R}^3 \to \mathbb{R}^3$ the rotation with y-axis and angle θ. The parametrization

$$\begin{aligned}\overrightarrow{\mathbf{x}}(u,v) &= R^y_v\left(\frac{\sqrt{2}}{2}(\cos(u),\sin(u),1)\right)\\ &= \frac{\sqrt{2}}{2}(\cos(u)\cos(v)+\sin(v),\sin(u),\cos(v)-\cos(u)\sin(v)),\end{aligned}$$

for $u \in \,]\frac{\pi}{2}, -\frac{\pi}{2}[$, and $v \in \mathbb{R}$, verifies that the coordinate lines $t \to \overrightarrow{\mathbf{x}}(t,v)$ are nongeodesic biharmonic curves. (See fig. 3)

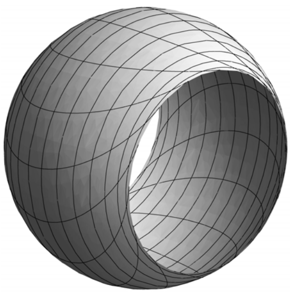

Fig. 3: The parametrization of the central section of the sphere with nongeodesic biharmonic curves as coordinate lines.

Of course, in this example we can not talk about gradient lines of the Gauss curvature because it is a constant function. Instead, we can study orthogonal curves to the coordinate lines. Since the geodesics in the sphere are great circles and they are not orthogonal to the coordinate lines of the parametrization $\overrightarrow{\mathbf{x}}$, then the family of orthogonal lines to the nongeodesic biharmonic curves is not made of geodesic curves.

Acknowledgements

The author wish to thank the referee for comments that have improved the paper.

REFERENCES

[1] P. Baird – D. Kamissoko: *On constructing biharmonic maps and metrics*, Annals of global Analysis and Geometry, **23** (2003), 65-75.

[2] R. Caddeo – S. Montaldo – P. Piu: *Biharmonic curves on a surface*, Rendiconti di Matematica, Serie VII, **21** (2001), 143-157.

[3] E. Kreyszig: *Differential geometry*, Dover Publications Inc., New York, 1991.

[4] C. S. Lin: *The local isometric embeddings in* $\mathbb{R}^3$ *of two-dimensional Riemannian manifolds with nonnegative curvature*, J. of Differential Geometry, **21** (1985), 213-230.

Lavoro pervenuto alla redazione il 6 novembre 2007
ed accettato per la pubblicazione il 10 aprile 2008.
Bozze licenziate il 30 settembre 2008

Indirizzo dell'Autore:

J. Monterde – Dep. de Geometria i Topologia – Universitat de València – Avd. Vicent Andrés Estellés 1 – E-46100-Burjassot (València) Spain
E-mail: monterde@uv.es

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 133–138

A sufficient condition for the Dunford-Pettis property in Banach spaces

N. L. BRAHA

ABSTRACT: *In this paper we will give a sufficient condition for Dunford-Pettis property in Banach spaces. More precisely, if Banach space X has a basic, normalized system of vectors (x_n), which is $f(n)$-approximate l_1, then X has the Dunford-Pettis property.*

1 – Introduction

A Banach space X is said to have the Dunford-Pettis property (DP) if for any Banach space Y, every weakly compact operator $T : X \to Y$ is completely continuous, i.e., T maps weakly compact subsets of X onto norm compact subsets of Y. Equivalently, X has Dunford-Pettis property iff for any weakly null sequences $(x_n) \in X$ and $(x_n^*) \in X^*$, one has $x_n^*(x_n) \to 0$.

In [12], was proved that if A is a disk algebra, compact Hausdorff space and μ a Borel measure on Ω, then the dual of $C(\Omega, A)$ has the Dunford-Pettis property. The DP property was studied for so called the polynomial DP property (see [8]). DP property also was studied in the tensor product of Banach spaces (see [10], [3]) etc. The reader will find further details on DPP and related properties in the survey [5]. In this paper we will give a sufficient condition under which Banach space X has the DP property. More precisely, if a Banach space X has a basic, normalized system of vectors (x_n), which is $f(n)$-approximate l_1, then X has the Dunford-Pettis property.

KEY WORDS AND PHRASES: *Dunford-Pettis property – $f(n)$-approximate l_1 sequence.*
A.M.S. CLASSIFICATION: 46B22 – 46A32

2 – Notation

Throughout the paper we will denote by B the closed unit ball, and with S the closed unit sphere. Recall by [4] that if $\{X_\alpha : \alpha \in \Lambda\}$ is a family of Banach spaces, and $1 \leq p < \infty$, one defines the Banach spaces:

$$\left(\bigoplus_{\alpha\in\Lambda} X_\alpha\right)_p = \left\{x = (x_\alpha) \in \prod_{\alpha\in\Lambda} X_\alpha : \sum_{\alpha\in\Lambda} ||x_\alpha||^p < \infty\right\},$$

with norm $||x|| = (\sum_\alpha ||x_\alpha||^p)^{\frac{1}{p}}$.

Definition 2.1 ([7]). Suppose that $(f(n))_{n=1}^\infty$ is a strictly positive nondecreasing sequence satisfying $\lim_{n\to\infty} f(n) = \infty$. Let $(x_i)_{i\in\mathbb{N}}$ be a sequence of unit vectors in a Banach space X. We say that (x_i) is a $f(n)$-approximate l_1 system if

$$\left\|\sum_{i\in A} \pm x_i\right\| \geq |A| - f(|A|) \tag{1}$$

for all finite sets $A \subset I$ and for all choices of signs.

Definition 2.2 ([2]). Let X be a Banach space. A sequence $(x_n)_{n\in\mathbb{N}}$ in X is p-colacunary if there is a $\delta > 0$ such that

$$\left\|\sum_{i\leq n} a_i x_i\right\| \geq \delta \left(\sum_{i\leq n} |a_i|^p\right)^{\frac{1}{p}}, \tag{2}$$

for any sequence of scalars $a_0, a_1, \cdots, a_n$ and $1 \leq p < \infty$.

Theorem 2.3 ([6], Dunford). *If X has a bounded complete basis (x_n), then X has the Radon-Nikodym property. All other notations are like in* [11].

3 – Results

Lemma 3.1. *Let $(x_n)_{n\in\mathbb{N}}$ be a normalized, basic sequence of vectors in the Banach space X, which satisfies the condition* (1). *Let us denote by f a positive, nondecreasing function, such that for every $0 < \delta \leq 1$, it satisfies condition $0 < f(\delta) < 1$. Then the following relation*

$$\left\|\sum_{i\leq n} \pm x_i\right\| \geq f(\delta) \left(\sum_{i\leq n} ||x_i||\right), \tag{3}$$

holds.

PROOF. Let us denote by $f(n) = (1 - f(\delta)) \cdot \left(\sum_{i\leq n} ||x_i||\right)$. It follows that $f(n)$, is a positive, nondecreasing function, such that $\lim_{n\to\infty} f(n) = \infty$. Function $f(n)$ defined as above, satisfies conditions used into the relation (1), respectively it satisfies the following relation:

$$\left\| \sum_{i\leq n} \pm x_i \right\| \geq n - f(n) \Rightarrow$$

$$\Rightarrow \left\| \sum_{i\leq n} \pm x_i \right\| \geq n - (1 - f(\delta)) \cdot \left(\sum_{i\leq n} ||x_i|| \right) = f(\delta) \sum_{i\leq n} ||x_i||.$$

LEMMA 3.2. *Let $(y_n)_{n\in\mathbb{N}}$ be a basic sequence of vectors in Banach space X. Let us denote by f a positive, nondecreasing function, such that for every $0 < \delta \leq 1$, it satisfies condition $0 < f(\delta) < 1$. If the normalized sequence of vectors $({y_i}^0) = \left(\frac{y_i}{||y_i||}\right)$, obtained from (y_n) satisfies relation (1), then the following relation*

$$\left\| \sum_{i\leq n} \pm y_i \right\| \geq \inf_{i\leq n} ||y_i|| \cdot f(\delta) \left(\sum_{i\leq n} ||{y_i}^0|| \right) \tag{4}$$

holds.

PROOF. Since ${y_i}^0 = \frac{y_i}{||y_i||}$ is a normalized sequence of vectors which satisfies (1), we use Lemma 3.1 to get the inequality

$$\left\| \sum_{i\leq n} \pm \frac{y_i}{||y_i||} \right\| \geq \sum_{i\leq n} \left\| \frac{y_i}{||y_i||} \right\| - f(n) \geq f(\delta) \left(\sum_{i\leq n} ||{y_i}^0|| \right). \tag{5}$$

From Hahn-Banach theorem, there exists a functional $x^* \in X^*$, $||x^*|| = 1$, such that the following estimation is valid:

$$\begin{aligned} \left\| \sum_{i\leq n} \pm \frac{y_i}{||y_i||} \right\| &= x^* \left(\sum_{i\leq n} \pm \frac{y_i}{||y_i||} \right) = \\ &= \sum_{i\leq n} \frac{1}{||y_i||} |x^*(y_i)| \leq \sup_{i\leq n} \frac{1}{||y_i||} \cdot \sum_{i\leq n} |x^*(y_i)|. \end{aligned} \tag{6}$$

Using relations (5) and (6), we have:

$$f(\delta)\left(\sum_{i\le n}||{y_i}^0||\right)\le \sup_{i\le n}\frac{1}{||y_i||}\cdot\sum_{i\le n}|x^*(y_i)| = \sup_{i\le n}\frac{1}{||y_i||}\cdot x^*\left(\sum_{i\le n}\pm y_i\right)\le$$

$$\le \sup_{i\le n}\frac{1}{||y_i||}\cdot\left|x^*\left(\sum_{i\le n}\pm y_i\right)\right|\le \sup_{i\le n}\frac{1}{||y_i||}\cdot\left\|\sum_{i\le n}\pm y_i\right\|.$$

Finally we get the following estimation:

$$\left\|\sum_{i\le n}\pm y_i\right\|\ge \inf_{i\le n}||y_i||\cdot f(\delta)\left(\sum_{i\le n}||{y_i}^0||\right).$$

Theorem 3.3. *Let $(x_n)_{n\in\mathbb{N}}$ be a normalized, basic sequence of vectors in Banach space X, which satisfies the condition* (1)*. Then the following relation*

$$K_1\cdot\sum_{i=1}^n|a_i|\le\left\|\sum_{i=1}^n a_i\cdot x_i\right\|\le\sum_{i=1}^n|a_i| \tag{7}$$

holds, for any sequence of scalars (a_i) and some constant K_1.

Proof. The right hand side of (7) is obvious, in what follows we will prove the left hand side of relation (7). Let (a_n) be any sequence of scalars. From Lemma 3.2 we obtain the following estimation:

$$\left\|\sum_{i=1}^n a_i\cdot x_i\right\| = \left\|\sum_{i=1}^n \pm\cdot|a_i|\cdot x_i\right\|\ge\inf_{i\le n}\||a_i|\cdot x_i\|\cdot f(\delta)\cdot\sum_{i\le n}\left\|\frac{|a_i|\cdot x_i}{|||a_i|\cdot x_i||}\right\|\ge$$

$$\ge f(\delta)\cdot\inf_{i\le n}|a_i|\cdot\inf_{i\le n}\frac{1}{|a_i|}\cdot\sum_{i\le n}|a_i|\Rightarrow$$

thus

$$\left\|\sum_{i=1}^n a_i\cdot x_i\right\|\ge f(\delta)\cdot\inf_{i\le n}|a_i|\cdot\inf_{i\le n}\frac{1}{|a_i|}\cdot\sum_{i\le n}|a_i|. \tag{8}$$

From relation (8) it follows that the inequality

$$\left\|\sum_{i=1}^n a_i\cdot x_i\right\|\ge K_1\cdot\sum_{i\le n}|a_i|,$$

where $K_1 = f(\delta)\cdot\frac{\inf_{i\le n}|a_i|}{\sup_{i\le n}|a_i|}$, holds.

Corollary 3.4. *Let* $(x_n)_{n\in\mathbb{N}}$ *be a normalized, basic sequence of vectors in Banach space X, which satisfies the condition* (1). *Then X admits the Radon-Nikodym property.*

Theorem 3.5. *Let* $(x_n)_{n\in\mathbb{N}}$ *be a sequence of normalized vectors in Banach space X, which satisfies condition* (1). *Then X admits a Dunford-Pettis property.*

Proof. From Theorem 3.3 it is known that the basic sequence of vectors (x_n) in X, is equivalent to the standard unit vector basis (e_i) of l_1 (relation (7)). It means that Banach spaces X and l_1 are isomorphic. Now proof of the Theorem follows from the fact that Dunford-Pettis property is preserved by isomorphism and from fact that l_1 admits the Dunford-Pettis property (see [1]).

In [9], was proved the following: Let $\{X_\alpha : \alpha \in \Lambda\}$, ($\Lambda$-is a family of indexes), be a family of Banach spaces which admits the DP1 property. Then $X = \left(\bigoplus_\alpha X_\alpha\right)_p$, $1 \le p < \infty$ has the DP1, too (see [9], for more details).

Corollary 3.6. *Let* $(x_n^\alpha)_{\alpha\in\Lambda}$ *be a basic, normalized sequence of vectors in Banach space* X_α. *If* (x_n^α) *satisfies condition* (1), *for every* $\alpha \in \Lambda$, *then there* $X = \left(\bigoplus_\alpha X_\alpha\right)_p$, $1 \le p < \infty$, *admits a* DP1 *property.*

REFERENCES

[1] F. Albaic – N. J. Kalton: *Topics in Banach space theory*, Springer-Verlag, 2006.

[2] D. J. Aldous – D. H. Fremlin: *Colacunary sequences in L-spaces*, Studia Math. T., **LXXI** (1982), 297-304.

[3] F. Bombal – I. Villanueva: *On the Dunford-Pettis property of the tensor product of $C(K)$ spaces*, arXiv:math.FA/0004087v1 2000.

[4] J. B. Conway: *A course in functional analysis*, Springer, New York, 1985.

[5] J. Diestel: *A survey of results related to the Dunford-Pettis property*, in Contemp. Math., Amer. Math. Soc., **2** (1980), 15-60.

[6] J. Diestel – J. J. Uhl: *Vector measure*, Providence, Rhode Island, 1977.

[7] S. J. Dilworth – D. Kutzarova – P. Wojtaszczyk: *On Approximate l_1 systems in Banach Spaces*, J. Approx. Theory, **114** (2002), 214-241.

[8] J. Farmer – W. B. Johnson: *Polynomial Schur and polynomial Dunford-Pettis properties*, arXiv:math/9211210v1 1992.

[9] W. Freedman:*An alternative Dunford-Pettis Property*, Studia Math.,**125** (1997), 143-159.

[10] M. Gonzalez – J. Gutierrez: *The Dunford-Pettis property on tensor products*, arXiv:math/0004101v1 2000.

[11] J. Lindenstrauss – L. Tzafriri: *Classical Banach spaces*, Part I, Springer-Verlag, Berlin, Heidelberg, New York, 1977.

[12] N. Randrianantoanina: *Some remarks on the Dunford-Pettis property*, arXiv:math.FA/9501214 v1 1995.

Lavoro pervenuto alla redazione il 25 gennaio 2008
ed accettato per la pubblicazione il 26 febbraio 2008.
Bozze licenziate il 30 settembre 2008

Indirizzo dell'Autore:

N. L. Braha – Department of Mathematics and Computer Sciences – Avenue "Mother Theresa" 5 – Prishtinë – 10000 Kosova
E-mail: nbraha@yahoo.com

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 139–162

Closed differential forms on moduli spaces of sheaves

FRANCESCO BOTTACIN

Abstract: *Let X be a smooth projective variety, and let $\mathcal{M}$ be a moduli space of stable sheaves on X. For any flat family E of coherent sheaves on X parametrized by a smooth scheme Y, and for any integer m, with $1 \leq m \leq \dim X$, we construct a closed differential form $\Omega = \Omega_E$ on Y with values in $H^m(X, \mathcal{O}_X)$. By using the vector-valued differential form Ω we then prove that the choice of a (non-zero) differential m-form σ on X, $\sigma \in H^0(X, \Omega^m_X)$, determines, in a natural way, a closed differential m-form Ω_σ on $\mathcal{M}$.*

– Introduction

Let X be a smooth projective variety, defined over an algebraically closed field k of characteristic 0, and let $\mathcal{M}$ be a moduli space of stable sheaves on X. It is by now well known that many geometric properties of the moduli space $\mathcal{M}$ are determined by similar geometric properties of the base variety X.

A beautiful example of this general fact was discovered by S. Mukai in [Mu1], where he shows that if X is an abelian or K3 surface (hence it is a symplectic algebraic surface), then the choice of a symplectic structure (i.e., a non-degenerate 2-form) on X determines a symplectic structure on the moduli space $\mathcal{M}$. This result was later generalized by S. Kobayashi [K] to the case of a compact Kähler manifold X with a holomorphic symplectic structure: in this case too, the choice of a symplectic structure on X determines a symplectic structure on the nonsingular part of the moduli space of stable vector bundles on X.

Key Words and Phrases: *Closed forms – Differential forms – Moduli spaces of sheaves – Vector bundles.*
A.M.S. Classification: 14J60 – 14D20 – 58A10

In [B1] we generalized Mukai's result to the case of Poisson surfaces, by proving that, if the algebraic surface X admits a Poisson structure, then the choice of such a structure on X determines, in a natural way, a Poisson structure on the moduli space $\mathcal{M}$. Similar results also hold for other types of moduli spaces, such as moduli spaces of framed vector bundles [B3], moduli spaces of parabolic bundles [B4] and Hilbert schemes of points of a surface [B2].

In this paper we provide another example showing how geometric structures on X often determine similar geometric structures on moduli spaces of sheaves on X. We prove that, if X admits non-zero differential forms of degree m, then the choice of any such m-form σ determines a closed differential m-form Ω_σ on the moduli space $\mathcal{M}$ of stable sheaves on X.

A particularly interesting special case of this result is obtained by taking X to be a Calabi-Yau n-fold. In this case there is a canonical choice (up to scalar multiples) of a n-form σ on X. This implies that there is also a closed n-form Ω_σ on the moduli space $\mathcal{M}$. This result may be considered as a higher dimensional generalization of Mukai's result in [Mu1].

This paper is organized as follows: in Section 1 we recall some useful results about cup-products and trace maps. Then we introduce the symmetrized trace map and study its graded commutativity properties. This is the main technical tool needed to construct closed differential forms on moduli spaces of stable sheaves on X.

In Section 2 we construct, for any flat family E of coherent sheaves on X parametrized by a smooth scheme Y, and for any integer m, with $1 \leq m \leq \dim X$, a differential form $\Omega = \Omega_E$ on Y with values in $H^m(X, \mathcal{O}_X)$. The main part of this section is then devoted to prove that $d\Omega = 0$ (in order to symplify the exposition this is proven under the additional assumptions that the base field k is the complex field and that E is a family of locally free sheaves).

Finally, in Section 3, we use the vector-valued differential form Ω to construct ordinary (i.e., scalar-valued) differential forms on the moduli space $\mathcal{M}$ of stable sheaves on X. More precisely, we prove that the choice of a (non-zero) differential m-form σ on X, $\sigma \in H^0(X, \Omega_X^m)$, determines, in a natural way, a differential m-form Ω_σ on $\mathcal{M}$, defined by using the vector-valued m-form Ω. Then the closure of Ω, proved in Section 2, immediately implies the closure of the m-form Ω_σ.

The case of a smooth Calabi-Yau n-fold X is especially interesting. In fact, in this case there is a canonical choice (up to scalar multiples) of the n-form σ on X, hence there is also a closed differential n-form Ω_σ on the moduli space of stable sheaves on X. The natural question of the non-degeneracy of Ω_σ is then discussed.

In the last part of this section we provide an example in order to explain how the construction of the differential forms Ω_σ can be generalized to other types of moduli spaces. We analyze in detail the case of moduli spaces of stable framed vector bundles.

1 – Preliminaries

1.1 – Cup-product and trace maps

In this section we shall recall, without proofs, some standard facts about trace maps and cup-products. For more details (and proofs) we refer the reader to [A] or [HL2, p. 217].

Let X be a smooth n-dimensional projective variety over an algebraically closed field k of characteristic 0. In the sequel, whenever we speak of a sheaf on a scheme S we shall always mean a sheaf of $\mathcal{O}_S$-modules.

For any coherent sheaf E on X, the usual trace map

$$\mathrm{tr} : \mathcal{E}nd(E) \to \mathcal{O}_X$$

induces natural maps (also called trace maps and denoted by the same symbol)

$$\mathrm{tr} : \mathrm{Ext}^i(E, E) \to H^i(X, \mathcal{O}_X).$$

For any i and j there is a natural cup-product (or Yoneda composition) map

$$\mathrm{Ext}^i(E, E) \times \mathrm{Ext}^j(E, E) \xrightarrow{\circ} \mathrm{Ext}^{i+j}(E, E)$$

(which can be easily defined in terms of Čech cocycles by replacing E with a finite locally free resolution $E^{\cdot}$), and the composition of cup-product and trace is graded commutative in the following sense: if $\alpha \in \mathrm{Ext}^i(E, E)$ and $\beta \in \mathrm{Ext}^j(E, E)$, then

$$\mathrm{tr}(\alpha \circ \beta) = (-1)^{ij}\, \mathrm{tr}(\beta \circ \alpha) \tag{1.1}$$

as an element of $H^{i+j}(X, \mathcal{O}_X)$ (cf. [HL2, pp. 216-217]).

Analogous maps can also be defined in a relative situation. Let us consider a scheme Y of finite type over k and denote by $p : X \times Y \to X$ and $q : X \times Y \to Y$ the canonical projections.

If E is aY-flat family of coherent sheaves on X, we shall denote by $\mathcal{E}xt^i_q(E,E)$ the i-th relative Ext-sheaf, i.e., the sheaf on Y associated to the presheaf

$$U \mapsto \mathrm{Ext}^i(E|_{X\times U}, E|_{X\times U}),$$

for any open subset $U \subset Y$.

Since any such E admits a finite locally free resolution, the preceding constructions of the cup-product and the trace maps carry over to this relative situation. More precisely, for any i and j we have a cup-product map

$$\mathcal{E}xt^i_q(E, E) \times \mathcal{E}xt^j_q(E, E) \xrightarrow{\circ} \mathcal{E}xt^{i+j}_q(E, E)$$

and a trace map

$$\mathrm{tr} : \mathcal{E}xt^i_q(E, E) \to R^i q_* \mathcal{O}_{X\times Y} \cong H^i(X, \mathcal{O}_X) \otimes_k \mathcal{O}_Y$$

satisfying (1.1) for any sections α and β of $\mathcal{E}xt^i_q(E, E)$ and $\mathcal{E}xt^j_q(E, E)$, respectively.

1.2 – The symmetrized trace map

Let E be a coherent sheaf on X. For any integer $m \geq 1$ let us consider the "symmetrized composition map"

$$\underbrace{\mathcal{E}nd(E) \times \cdots \times \mathcal{E}nd(E)}_{m} \xrightarrow{S} \mathcal{E}nd(E) \tag{1.2}$$

defined by setting

$$S(\phi_1, \phi_2, \ldots, \phi_m) = \frac{1}{m!} \sum_{\sigma \in \mathfrak{S}_m} \phi_{\sigma(1)} \circ \phi_{\sigma(2)} \circ \cdots \circ \phi_{\sigma(m)},$$

where the sum runs over the group $\mathfrak{S}_m$ of permutations of m elements.

We define the "symmetrized trace", denoted by Str, to be the composition of S with the usual trace map:

$$\operatorname{Str}(\phi_1, \phi_2, \ldots, \phi_m) = \frac{1}{m!} \sum_{\sigma \in \mathfrak{S}_m} \operatorname{tr}(\phi_{\sigma(1)} \circ \phi_{\sigma(2)} \circ \cdots \circ \phi_{\sigma(m)}). \tag{1.3}$$

The map

$$\operatorname{Str} : \mathcal{E}nd(E) \times \cdots \times \mathcal{E}nd(E) \to \mathcal{O}_X \tag{1.4}$$

is totally symmetric and multilinear.

It is easy to prove that, for a general $m \geq 2$,

$$\operatorname{Str}(\phi_1, \phi_2, \ldots, \phi_m) = \frac{1}{(m-1)!} \sum_{\sigma} \operatorname{tr}(\phi_1 \circ \phi_{\sigma(2)} \circ \cdots \circ \phi_{\sigma(m)}), \tag{1.5}$$

where the sum runs over all permutations σ of the set $\{2, 3, \ldots, m\}$.

The symmetrized trace map (1.4) induces a map, also denoted by Str,

$$\operatorname{Str} : \operatorname{Ext}^{i_1}(E, E) \times \cdots \times \operatorname{Ext}^{i_m}(E, E) \to H^{i_1 + \cdots + i_m}(X, \mathcal{O}_X). \tag{1.6}$$

This map satisfies a kind of graded commutativity property similar to the one stated in (1.1).

Proposition 1.1. *Let $\phi_h \in \operatorname{Ext}^{i_h}(E, E)$, for $h = 1, \ldots, m$. For any integer p, with $1 \leq p \leq m-1$, we have:*

$$\operatorname{Str}(\phi_1, \ldots, \phi_p, \phi_{p+1}, \ldots, \phi_m) = (-1)^{i_p i_{p+1}} \operatorname{Str}(\phi_1, \ldots, \phi_{p+1}, \phi_p, \ldots, \phi_m),$$

i.e., whenever we mutually exchange two adjacent elements ϕ_p and ϕ_{p+1}, the value of Str *acquires the factor $(-1)^{\deg(\phi_p)\deg(\phi_{p+1})} = (-1)^{i_p i_{p+1}}$.*

Proof. Let $E^\cdot$ be a finite locally free resolution of E and set $A^\cdot = \mathrm{Hom}^\cdot(E^\cdot, E^\cdot)$. Then we have $\mathrm{Ext}^i(E,E) = H^i(A^\cdot)$, the i-th hypercohomology group of the complex $A^\cdot$.

The symmetrized trace map

$$\mathrm{Str} : H^{i_1}(A^\cdot) \otimes \cdots \otimes H^{i_m}(A^\cdot) \to H^{i_1+\cdots+i_m}(X, \mathcal{O}_X)$$

is the composition of the following three maps: first, the multiplication

$$\mu : H^{i_1}(A^\cdot) \otimes \cdots \otimes H^{i_m}(A^\cdot) \to H^{i_1+\cdots+i_m}(A^\cdot \otimes \cdots \otimes A^\cdot), \tag{1.7}$$

then the symmetrized composition map

$$S : H^{i_1+\cdots+i_m}(A^\cdot \otimes \cdots \otimes A^\cdot) \to H^{i_1+\cdots+i_m}(A^\cdot) \tag{1.8}$$

and finally the usual trace map

$$\mathrm{tr} : H^{i_1+\cdots+i_m}(A^\cdot) \to H^{i_1+\cdots+i_m}(X, \mathcal{O}_X). \tag{1.9}$$

Let us denote by $\pi = \pi_{p,p+1}$ the twist operator that exchanges the factors at places p and $p+1$:

$$\begin{aligned}\pi : H^{i_1}(A^\cdot) \otimes \cdots \otimes H^{i_p}(A^\cdot) \otimes H^{i_{p+1}}(A^\cdot) \otimes \cdots \otimes H^{i_m}(A^\cdot) \to \\ \to H^{i_1}(A^\cdot) \otimes \cdots \otimes H^{i_{p+1}}(A^\cdot) \otimes H^{i_p}(A^\cdot) \otimes \cdots \otimes H^{i_m}(A^\cdot).\end{aligned}$$

We have

$$\pi(\alpha_1 \otimes \cdots \otimes \alpha_p \otimes \alpha_{p+1} \otimes \cdots \otimes \alpha_m) = (-1)^{i_p i_{p+1}} \alpha_1 \otimes \cdots \otimes \alpha_{p+1} \otimes \alpha_p \otimes \cdots \otimes \alpha_m,$$

where $i_p = \deg(\alpha_p)$ and $i_{p+1} = \deg(\alpha_{p+1})$. Then the following diagram is commutative:

$$\begin{array}{ccc} H^{i_1}(A^\cdot) \otimes \cdots \otimes H^{i_m}(A^\cdot) & \xrightarrow{\ \mu\ } & H^{i_1+\cdots+i_m}(A^\cdot \otimes \cdots \otimes A^\cdot) \\ \downarrow{\scriptstyle \pi_{p,p+1}} & & \downarrow{\scriptstyle H(\pi_{p,p+1})} \\ H^{i_1}(A^\cdot) \otimes \cdots \otimes H^{i_m}(A^\cdot) & \xrightarrow{\ \mu\ } & H^{i_1+\cdots+i_m}(A^\cdot \otimes \cdots \otimes A^\cdot). \end{array}$$

By composing with the maps S and tr of (1.8) and (1.9), it follows that $\mathrm{Str} = \mathrm{Str} \circ \pi_{p,p+1}$, which is precisely what we had to prove.

In the sequel we shall be interested in a special case of (1.6). By taking all i_h equal to 1, we get the map

$$\operatorname{Str} : \underbrace{\operatorname{Ext}^1(E,E) \times \cdots \times \operatorname{Ext}^1(E,E)}_{m} \to H^m(X, \mathcal{O}_X),$$

satisfying

$$\operatorname{Str}(\phi_1, \ldots, \phi_p, \phi_{p+1}, \ldots, \phi_m) = -\operatorname{Str}(\phi_1, \ldots, \phi_{p+1}, \phi_p, \ldots, \phi_m),$$

for every $p \in [1, m-1]$.

Corollary 1.2. *For any $m \geq 1$, the map*

$$\operatorname{Str} : \underbrace{\operatorname{Ext}^1(E,E) \times \cdots \times \operatorname{Ext}^1(E,E)}_{m} \to H^m(X, \mathcal{O}_X)$$

is alternating, i.e., for any permutation σ of $\{1, 2, \ldots, m\}$, we have:

$$\operatorname{Str}(\alpha_{\sigma(1)}, \alpha_{\sigma(2)}, \ldots, \alpha_{\sigma(m)}) = \operatorname{sgn}(\sigma) \operatorname{Str}(\alpha_1, \alpha_2, \ldots, \alpha_m).$$

Let now Y be a scheme of finite type over k and let E be a Y-flat family of coherent sheaves on X. Since any such E admits a finite resolution by locally free sheaves, the preceding constructions can be generalized to this relative situation (exactly as in the case of the usual trace map, described in Section 1.1). We leave the details to the reader and just state the relative version of Corollary 1.2:

Corollary 1.3. *Let E be a Y-flat family of coherent sheaves on X. For any $m \geq 1$, the map*

$$(1.10) \qquad \begin{aligned} \operatorname{Str} : \underbrace{\mathcal{E}xt^1_q(E,E) \times \cdots \times \mathcal{E}xt^1_q(E,E)}_{m} &\to R^m q_*(\mathcal{O}_{X\times Y}) \cong \\ &\cong H^m(X, \mathcal{O}_X) \otimes_k \mathcal{O}_Y \end{aligned}$$

is alternating, i.e., for any permutation σ of $\{1, 2, \ldots, m\}$, we have:

$$\operatorname{Str}(\alpha_{\sigma(1)}, \alpha_{\sigma(2)}, \ldots, \alpha_{\sigma(m)}) = \operatorname{sgn}(\sigma) \operatorname{Str}(\alpha_1, \alpha_2, \ldots, \alpha_m).$$

2 – Vector-valued differential forms

Let Y be a *smooth* scheme of finite type over k and E a Y-flat family of coherent sheaves on X. In this section we shall define, for any such E and any integer m, with $1 \leq m \leq n = \dim X$, a vector-valued m-form on Y (more precisely, a differential form of degree m on Y with values in $H^m(X, \mathcal{O}_X)$). Then we shall prove that these differential forms are closed.

Let us begin by recalling that, for any sheaf F on X, the set of isomorphism classes of infinitesimal deformations of F is canonically identified with $\mathrm{Ext}^1(F, F)$. It follows that, for any family E of coherent sheaves on X parametrized by Y, there is a map, known as the Kodaira-Spencer map,

$$\rho : TY \to \mathcal{E}xt^1_q(E, E), \tag{2.1}$$

that sends a tangent vector $v \in T_y Y$ to the class $\rho(v) \in \mathrm{Ext}^1(E_y, E_y)$ corresponding to the infinitesimal deformation of the sheaf E_y along the direction of v.

Now, for any m as above, we define an $H^m(X, \mathcal{O}_X)$-valued differential m-form $\Omega = \Omega_E$ on Y by setting

$$\Omega : \underbrace{TY \times \cdots \times TY}_{m} \to \mathcal{E}xt^1_q(E, E) \times \cdots \times \mathcal{E}xt^1_q(E, E) \to H^m(X, \mathcal{O}_X) \otimes_k \mathcal{O}_Y,$$

where the first map is induced by the Kodaira-Spencer map (2.1), and the second one is the symmetrized trace map (1.10). In other words, we set

$$\Omega(v_1, \ldots, v_m) = \mathrm{Str}(\rho(v_1), \ldots, \rho(v_m)), \tag{2.2}$$

for any sections $v_1, \ldots, v_m$ of the tangent bundle TY. It follows from Corollary 1.3 that Ω is a vector-valued differential form of degree m.

Remark 2.1. Let E be a Y-flat family of sheaves on X and L be a line bundle on Y. We can define another Y-flat family of sheaves E' on X by setting $E' = E \otimes q^*(L)$. These two families of sheaves may be considered as equivalent because, for every closed point $y \in Y$, the sheaves E_y and E'_y on X are isomorphic. Under these hypotheses, the differential m-forms Ω_E and $\Omega_{E'}$ are equal.

Remark 2.2. Let us observe that in the definition of Ω_E we do not use directly the sheaf E, but rather the sheaf $\mathcal{E}xt^1_q(E, E)$. This is very important because in most interesting applications, when we take as Y a suitable moduli space of stable sheaves on X, a global universal family E does not exist (at least as an ordinary sheaf), but the sheaf $\mathcal{E}xt^1_q(E, E)$ on Y is, nevertheless, well defined (cf. Remark 3.1). It follows that our definition of the differential form Ω_E remains valid also in these more general situations.

The rest of this section will be devoted to the proof that, for any E and any integer m, $d\Omega_E = 0$.

In order to try to simplify the exposition as far as possible, we shall assume, from now on, that k is the complex field $\mathbb{C}$, even if our proof works, with only minor modifications, for any algebraically closed field k of characteristic 0.

So let X be a smooth n-dimensional complex projective variety and let Y be a smooth complex variety of dimension N. We shall use the complex analytic topology on the complex manifolds associated to the algebraic varieties X and Y.

Let E be a Y-flat family of sheaves on X. We shall assume that E is locally free of rank r. Let $\Omega = \Omega_E$ be the m-form on Y defined above.

First of all, we remark that, since the closure of Ω is a local property on Y, we may freely replace Y be an open neighborhood of any one of its points, hence we may assume that a local system of holomorphic coordinates $y = (y_i)_{i=1,\dots,N}$ is given on Y.

Now, by eventually replacing Y with a smaller open subset, if necessary, we may find an open covering $(U_i)_{i\in I}$ of X such that the restriction of the vector bundle E to $U_i \times Y$ is trivial. Let us denote by

$$f_i : E|_{U_i\times Y} \xrightarrow{\sim} (U_i \times Y) \times \mathbb{C}^r$$

the trivialization isomorphisms. On the intersection $U_{ij} = U_i \cap U_j$ of two open subsets, we have the isomorphism

$$g_{ij} : (U_{ij} \times Y) \times \mathbb{C}^r \xrightarrow{\sim} (U_{ij} \times Y) \times \mathbb{C}^r$$

defined by setting $g_{ij} = f_i|_{U_{ij}\times Y} \circ f_j^{-1}|_{U_{ij}\times Y}$. For any point $(x, y) \in U_{ij} \times Y$ and any $v \in \mathbb{C}^r$, we have

$$g_{ij} : ((x, y), v) \mapsto ((x, y), \tilde{g}_{ij}(x, y)\, v),$$

where

$$\tilde{g}_{ij} : U_{ij} \times Y \to \mathrm{GL}(r, \mathbb{C})$$

is a holomorphic function. In the sequel, by abuse of notation, we shall identify g_{ij} with $\tilde{g}_{ij}$. The functions g_{ij} are called the transition functions of the vector bundle E (with respect to the given open covering).

The transition functions $(g_{ij})_{i,j\in I}$ satisfy the usual cocycle identities

$$g_{ii} = 1, \qquad g_{ji} = g_{ij}^{-1}, \qquad g_{ij} \circ g_{jk} = g_{ik}, \tag{2.3}$$

on the intersection of three open subsets $U_{ijk} = U_i \cap U_j \cap U_k$.

Now we can give an explicit description of the Kodaira-Spencer map

$$\rho : TY \to \mathcal{E}xt^1_q(E, E) = R^1 q_* \mathcal{E}nd(E).$$

At any point $y \in Y$ the derivations $\partial_\alpha = \frac{\partial}{\partial y_\alpha}$, for $\alpha = 1, \dots, N$, form a basis of the tangent space $T_y Y$. If we consider the vector bundle E on $X \times Y$ as a family of vector bundles $\{E_y\}_{y \in Y}$ on X varying holomorphically with $y \in Y$, the infinitesimal deformation of the vector bundle E_y corresponding to the tangent vector ∂_α is given by $\frac{\partial E_y}{\partial y_\alpha}$. To give a meaning to this symbol, let us consider the transition functions $\{g_{ij}(\cdot, y)\}_{i,j \in I}$ of the vector bundle E_y on X. Then the infinitesimal deformation $\frac{\partial E_y}{\partial y_\alpha}$ is given (by definition) by the collection of functions $\{\partial_\alpha g_{ij}(\cdot, y)\}_{i,j \in I}$.

By deriving the (multiplicative) cocycle identities (2.3), we obtain the following (additive) cocycle identities for the functions $\partial_\alpha g_{ij}(\cdot, y)$:

$$\partial_\alpha g_{ji} = -g_{ij}^{-1} \circ (\partial_\alpha g_{ij}) \circ g_{ij}^{-1}, \quad (\partial_\alpha g_{ij}) \circ g_{jk} + g_{ij} \circ (\partial_\alpha g_{jk}) = \partial_\alpha g_{ik}. \tag{2.4}$$

By recalling that $g_{ij} = f_i \circ f_j^{-1}$, these identities can be rewritten as

$$f_j^{-1}(\partial_\alpha g_{ji}) f_i = -f_i^{-1}(\partial_\alpha g_{ij}) f_j, \quad f_i^{-1}(\partial_\alpha g_{ij}) f_j + f_j^{-1}(\partial_\alpha g_{jk}) f_k = f_i^{-1}(\partial_\alpha g_{ik}) f_k.$$

Finally, if we set

$$\eta_{ij}^\alpha = f_i^{-1}(\partial_\alpha g_{ij}) f_j,$$

we obtain a collection of sections $\eta_{ij}^\alpha(\cdot, y) \in \Gamma(U_{ij}, \mathcal{E}nd(E_y))$ satisfying the usual (additive) cocycle relations

$$\eta_{ji}^\alpha = -\eta_{ij}^\alpha, \quad \eta_{ij}^\alpha + \eta_{jk}^\alpha = \eta_{ik}^\alpha. \tag{2.5}$$

The cohomology class $\bar{\eta}^\alpha \in H^1(X, \mathcal{E}nd(E_y))$ represented by the Čech 1-cocycle $\{\eta_{ij}^\alpha\}_{i,j \in I}$ is the image of the tangent vector $\partial_\alpha \in T_y Y$ by the Kodaira-Spencer map

$$\rho : T_y Y \to (R^1 q_* \mathcal{E}nd(E))_y = H^1(X, \mathcal{E}nd(E_y)).$$

Remark 2.3. Other equivalent representations of the infinitesimal deformation $\frac{\partial E_y}{\partial y_\alpha}$ are possible. For instance, in [Mu2], Mukai represents the cohomology class $\rho(\partial_\alpha) \in H^1(X, \mathcal{E}nd(E_y))$ by using the 1-cocycle $\{a_{ij}\}_{i,j \in I}$ defined by setting

$$a_{ij} = g_{ij}^{-1} \partial_\alpha g_{ij}.$$

This is equivalent to our representation, except for the fact that the a_{ij}'s satisfy a kind of twisted cocycle identities, given by

$$a_{ji} = -g_{ij} a_{ij} g_{ij}^{-1}, \qquad g_{jk}^{-1} a_{ij} g_{jk} + a_{jk} = a_{ik}.$$

(See [Mu2, Section 3, p. 151]).

We are now able to give an explicit description of the m-form $\Omega = \Omega_E$ on Y in terms of Čech cocycles.

For any closed point $y \in Y$ and any $\partial_\alpha \in T_y Y$ we have set $\eta_{ij}^\alpha = f_i^{-1}\, \partial_\alpha g_{ij}\, f_j$, and we have seen that the 1-cocycle $\{\eta_{ij}^\alpha\}_{i,j\in I}$ represents the cohomology class $\bar\eta^\alpha = \rho(\partial_\alpha) \in H^1(X, \mathcal{E}nd(E_y))$. It follows from the definition (2.2) of Ω that, for any $\alpha_1, \dots, \alpha_m \in [1, N]$, we have:

$$\Omega(\partial_{\alpha_1}, \dots, \partial_{\alpha_m}) = \mathrm{Str}(\bar\eta^{\alpha_1}, \dots, \bar\eta^{\alpha_m}).$$

We shall now recall the standard expression of the cup-product in terms of Čech cocycles:

Lemma 2.4. *Let $\{\psi_{ij}^h\}_{i,j\in I}$, for $h = 1, \dots, m$, be Čech 1-cocycles representing the cohomology classes $\bar\psi^h \in H^1(X, \mathcal{E}nd(F))$, for some locally free sheaf F on X. Then the cup-product (or Yoneda composition)*

$$\bar\psi^1 \circ \bar\psi^2 \circ \cdots \circ \bar\psi^m \in H^m(X, \mathcal{E}nd(F))$$

is the cohomology class represented by the Čech m-cocycle

$$\{\psi^1_{i_1 i_2} \circ \psi^2_{i_2 i_3} \circ \cdots \circ \psi^m_{i_m i_{m+1}}\}_{i_1,\dots,i_{m+1}\in I}.$$

From this result, and the definition (1.3) of the map Str, we obtain the following explicit expression of Ω:

$$\begin{aligned}\Omega(\partial_{\alpha_1}, \dots, \partial_{\alpha_m}) &= \mathrm{Str}(\bar\eta^{\alpha_1}, \dots, \bar\eta^{\alpha_m}) = \\ &= \frac{1}{m!} \sum_{\sigma\in\mathfrak{S}_m} \Big\{ \mathrm{tr}(\eta^{\alpha_{\sigma(1)}}_{i_1 i_2} \circ \eta^{\alpha_{\sigma(2)}}_{i_2 i_3} \circ \cdots \circ \eta^{\alpha_{\sigma(m)}}_{i_m i_{m+1}}) \Big\}_{i_1, i_2, \dots, i_{m+1}}.\end{aligned}$$

To simplify the notation, we introduce a multiindex $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_m)$. Then, for every permutation $\sigma \in \mathfrak{S}_m$, we set

$$\sigma(\boldsymbol{\alpha}) = (\alpha_{\sigma(1)}, \alpha_{\sigma(2)}, \dots, \alpha_{\sigma(m)})$$

and

$$\boldsymbol{\eta}^{\sigma(\boldsymbol{\alpha})}_{i_1, i_2, \dots, i_{m+1}} = \eta^{\alpha_{\sigma(1)}}_{i_1 i_2} \circ \eta^{\alpha_{\sigma(2)}}_{i_2 i_3} \circ \cdots \circ \eta^{\alpha_{\sigma(m)}}_{i_m i_{m+1}}.$$

We shall also write $\boldsymbol{\partial}_{\boldsymbol{\alpha}} = (\partial_{\alpha_1}, \partial_{\alpha_2}, \dots, \partial_{\alpha_m})$ and $\mathbf{dy}_{\boldsymbol{\alpha}} = dy_{\alpha_1} \wedge dy_{\alpha_2} \wedge \cdots \wedge dy_{\alpha_m}$.

With these notations we may write

$$\Omega(\partial_{\alpha_1}, \dots, \partial_{\alpha_m}) = \Omega(\boldsymbol{\partial}_{\boldsymbol{\alpha}}) = \frac{1}{m!} \sum_{\sigma\in\mathfrak{S}_m} \Big\{ \mathrm{tr}(\boldsymbol{\eta}^{\sigma(\boldsymbol{\alpha})}_{i_1, i_2, \dots, i_{m+1}}) \Big\}_{i_1, i_2, \dots, i_{m+1}}.$$

Since the derivations ∂_α, for $\alpha = 1, \dots, N$, are a basis of the tangent space T_yY, we have:

$$
\begin{aligned}
\Omega &= \sum_{\alpha_1 < \cdots < \alpha_m} \Omega(\partial_{\alpha_1}, \dots, \partial_{\alpha_m})\, dy_{\alpha_1} \wedge \cdots \wedge dy_{\alpha_m} = \\
&= \frac{1}{m!} \sum_{\alpha_1, \dots, \alpha_m} \Omega(\partial_{\alpha_1}, \dots, \partial_{\alpha_m})\, dy_{\alpha_1} \wedge \cdots \wedge dy_{\alpha_m} = \\
&= \frac{1}{m!} \sum_{\boldsymbol{\alpha}} \Omega(\boldsymbol{\partial_\alpha})\, \mathbf{dy}_{\boldsymbol{\alpha}} = \\
&= \frac{1}{(m!)^2} \sum_{\boldsymbol{\alpha}} \sum_{\sigma \in \mathfrak{S}_m} \Big\{ \operatorname{tr}(\boldsymbol{\eta}^{\sigma(\boldsymbol{\alpha})}_{i_1, i_2, \dots, i_{m+1}}) \Big\}_{i_1, i_2, \dots, i_{m+1}} \mathbf{dy}_{\boldsymbol{\alpha}}.
\end{aligned} \tag{2.6}
$$

An equivalent expression of Ω in terms of the transition functions g_{ij} can be given (cf. [Mu2, p. 154]). This will be useful in order to simplify the computation of $d\Omega$.

In fact, by recalling that $\eta^\alpha_{ij} = f_i^{-1} \partial_\alpha g_{ij} f_j$, we have:

$$
\begin{aligned}
\operatorname{tr}(\boldsymbol{\eta}^{\sigma(\boldsymbol{\alpha})}_{i_1, i_2, \dots, i_{m+1}}) &= \operatorname{tr}(\eta^{\alpha_{\sigma(1)}}_{i_1 i_2} \circ \eta^{\alpha_{\sigma(2)}}_{i_2 i_3} \circ \cdots \circ \eta^{\alpha_{\sigma(m)}}_{i_m i_{m+1}}) = \\
&= \operatorname{tr}\Big(f_{i_1}^{-1} \frac{\partial g_{i_1 i_2}}{\partial y_{\alpha_{\sigma(1)}}} \frac{\partial g_{i_2 i_3}}{\partial y_{\alpha_{\sigma(2)}}} \cdots \frac{\partial g_{i_m i_{m+1}}}{\partial y_{\alpha_{\sigma(m)}}} f_{i_{m+1}} \Big),
\end{aligned}
$$

and, since $\operatorname{tr}(\phi) = \operatorname{tr}(\psi \phi \psi^{-1})$, we also have:

$$
\begin{aligned}
\operatorname{tr}\Big(f_{i_1}^{-1} \frac{\partial g_{i_1 i_2}}{\partial y_{\alpha_{\sigma(1)}}} \cdots \frac{\partial g_{i_m i_{m+1}}}{\partial y_{\alpha_{\sigma(m)}}} f_{i_{m+1}} \Big) &= \operatorname{tr}\Big(\frac{\partial g_{i_1 i_2}}{\partial y_{\alpha_{\sigma(1)}}} \cdots \frac{\partial g_{i_m i_{m+1}}}{\partial y_{\alpha_{\sigma(m)}}} f_{i_{m+1}} f_{i_1}^{-1} \Big) = \\
&= \operatorname{tr}\Big(\frac{\partial g_{i_1 i_2}}{\partial y_{\alpha_{\sigma(1)}}} \cdots \frac{\partial g_{i_m i_{m+1}}}{\partial y_{\alpha_{\sigma(m)}}} g_{i_{m+1} i_1} \Big).
\end{aligned}
$$

It follows that:

$$
\Omega = \frac{1}{(m!)^2} \sum_{\boldsymbol{\alpha}} \sum_{\sigma \in \mathfrak{S}_m} \Big\{ \operatorname{tr}\Big(\frac{\partial g_{i_1 i_2}}{\partial y_{\alpha_{\sigma(1)}}} \cdots \frac{\partial g_{i_m i_{m+1}}}{\partial y_{\alpha_{\sigma(m)}}} g_{i_{m+1} i_1} \Big) \Big\}_{i_1, i_2, \dots, i_{m+1}} \mathbf{dy}_{\boldsymbol{\alpha}}. \tag{2.7}
$$

We shall now compute the exterior differential of Ω.

Lemma 2.5. *For Ω as above, we have*

$$
d\Omega = c \sum_{\boldsymbol{\alpha}} \sum_{\sigma \in \mathfrak{S}_m} \Big\{ \operatorname{tr}\Big(\frac{\partial g_{i_1 i_2}}{\partial y_{\alpha_{\sigma(1)}}} \cdots \frac{\partial g_{i_m i_{m+1}}}{\partial y_{\alpha_{\sigma(m)}}} \frac{\partial g_{i_{m+1} i_1}}{\partial y_{\alpha_{m+1}}} \Big) \Big\}_{i_1, i_2, \dots, i_{m+1}} \mathbf{dy}_{\boldsymbol{\alpha}},
$$

where $c = \frac{(-1)^m}{(m!)^2}$ and where the first sum runs now over all multiindices $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_{m+1})$.

Proof. By taking the exterior differential of the expression (2.7), we obtain

$$d\Omega=\frac{1}{(m!)^2}\sum_{\boldsymbol{\alpha}}\sum_{\sigma\in\mathfrak{S}_m}\Big\{\sum_{\alpha_{m+1}}\frac{\partial}{\partial y_{\alpha_{m+1}}}\mathrm{tr}\Big(\frac{\partial g_{i_1 i_2}}{\partial y_{\alpha_{\sigma(1)}}}\cdots\frac{\partial g_{i_m i_{m+1}}}{\partial y_{\alpha_{\sigma(m)}}}g_{i_{m+1}i_1}\Big)\Big\}dy_{\alpha_{m+1}}\wedge\mathbf{dy}_{\boldsymbol{\alpha}},$$

where we recall that $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\dots,\alpha_m)$.

The sum over α_{m+1} followed by the sum over all multiindices $\boldsymbol{\alpha} = (\alpha_1,\alpha_2,\dots,\alpha_m)$ is equivalent to a sum over all multiindices (still denoted by the same symbol) $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\dots,\alpha_{m+1})$, and since $dy_{\alpha_{m+1}}\wedge\mathbf{dy}_{\boldsymbol{\alpha}}=dy_{\alpha_{m+1}}\wedge dy_{\alpha_1}\wedge\cdots\wedge dy_{\alpha_m}=(-1)^m dy_{\alpha_1}\wedge\cdots\wedge dy_{\alpha_m}\wedge dy_{\alpha_{m+1}}$, we can write

$$d\Omega=c\sum_{\boldsymbol{\alpha}}\sum_{\sigma\in\mathfrak{S}_m}\Big\{\frac{\partial}{\partial y_{\alpha_{m+1}}}\,\mathrm{tr}\,\Big(\frac{\partial g_{i_1 i_2}}{\partial y_{\alpha_{\sigma(1)}}}\cdots\frac{\partial g_{i_m i_{m+1}}}{\partial y_{\alpha_{\sigma(m)}}}g_{i_{m+1}i_1}\Big)\Big\}\,\mathbf{dy}_{\boldsymbol{\alpha}}.$$

Now, to complete the proof, it suffices to observe that when we compute the partial derivative

$$\frac{\partial}{\partial y_{\alpha_{m+1}}}\,\mathrm{tr}\,\Big(\frac{\partial g_{i_1 i_2}}{\partial y_{\alpha_{\sigma(1)}}}\cdots\frac{\partial g_{i_m i_{m+1}}}{\partial y_{\alpha_{\sigma(m)}}}g_{i_{m+1}i_1}\Big)$$

we obtain the term

$$\mathrm{tr}\,\Big(\frac{\partial g_{i_1 i_2}}{\partial y_{\alpha_{\sigma(1)}}}\cdots\frac{\partial g_{i_m i_{m+1}}}{\partial y_{\alpha_{\sigma(m)}}}\frac{\partial g_{i_{m+1}i_1}}{\partial y_{\alpha_{m+1}}}\Big)$$

plus other terms involving second-order partial derivatives of the transition functions g_{ij}. But when we finally sum over all multiindices $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\dots,\alpha_{m+1})$, for every term of the form

$$\mathrm{tr}\,\Big(\cdots\frac{\partial^2 g_{i_h i_{h+1}}}{\partial y_\mu\partial y_\nu}\cdots g_{i_{m+1}i_1}\Big)dy_{\alpha_1}\wedge\cdots\wedge dy_\nu\wedge\cdots\wedge dy_{\alpha_m}\wedge dy_\mu$$

there is another term equal to

$$\mathrm{tr}\,\Big(\cdots\frac{\partial^2 g_{i_h i_{h+1}}}{\partial y_\nu\partial y_\mu}\cdots g_{i_{m+1}i_1}\Big)dy_{\alpha_1}\wedge\cdots\wedge dy_\mu\wedge\cdots\wedge dy_{\alpha_m}\wedge dy_\nu,$$

and these two terms add to zero because

$$\frac{\partial^2 g_{i_h i_{h+1}}}{\partial y_\mu\partial y_\nu}=\frac{\partial^2 g_{i_h i_{h+1}}}{\partial y_\nu\partial y_\mu}$$

while

$$dy_{\alpha_1}\wedge\cdots\wedge dy_\nu\wedge\cdots\wedge dy_{\alpha_m}\wedge dy_\mu=-dy_{\alpha_1}\wedge\cdots\wedge dy_\mu\wedge\cdots\wedge dy_{\alpha_m}\wedge dy_\nu.$$

Having computed $d\Omega$ in terms of the partial derivatives of the trivialization functions g_{ij}'s, we can now switch back to the representation in terms of the Čech cocycles $\eta_{ij}^{\alpha} = f_i^{-1}(\partial_\alpha g_{ij})f_j$ (this is convenient essentially because the cocycle identities (2.5) for the η_{ij}^{α}'s are simpler then the analogous cocycle identities (2.4) for the functions $\partial_\alpha g_{ij}$'s).

Since

$$\operatorname{tr}(\eta_{i_1 i_2}^{\alpha_{\sigma(1)}} \circ \cdots \circ \eta_{i_m i_{m+1}}^{\alpha_{\sigma(m)}} \circ \eta_{i_{m+1} i_1}^{\alpha_{m+1}}) = \operatorname{tr}\Big(f_{i_1}^{-1} \frac{\partial g_{i_1 i_2}}{\partial y_{\alpha_{\sigma(1)}}} \cdots \frac{\partial g_{i_m i_{m+1}}}{\partial y_{\alpha_{\sigma(m)}}} \frac{\partial g_{i_{m+1} i_1}}{\partial y_{\alpha_{m+1}}} f_{i_1}\Big) = \\ = \operatorname{tr}\Big(\frac{\partial g_{i_1 i_2}}{\partial y_{\alpha_{\sigma(1)}}} \cdots \frac{\partial g_{i_m i_{m+1}}}{\partial y_{\alpha_{\sigma(m)}}} \frac{\partial g_{i_{m+1} i_1}}{\partial y_{\alpha_{m+1}}}\Big),$$

we have

$$d\Omega = c \sum_{\boldsymbol{\alpha}} \sum_{\sigma \in \mathfrak{S}_m} \Big\{ \operatorname{tr}(\eta_{i_1 i_2}^{\alpha_{\sigma(1)}} \circ \cdots \circ \eta_{i_m i_{m+1}}^{\alpha_{\sigma(m)}} \circ \eta_{i_{m+1} i_1}^{\alpha_{m+1}}) \Big\} \mathbf{dy}_{\boldsymbol{\alpha}}. \tag{2.8}$$

Now, it follows from the cocycle relations (2.5) that, on $U_{i_1 i_2 \cdots i_{m+1}} = U_{i_1} \cap \cdots \cap U_{i_{m+1}}$, we have

$$\eta_{i_{m+1} i_1}^{\alpha_{m+1}} = -\eta_{i_1 i_{m+1}}^{\alpha_{m+1}} = -(\eta_{i_1 i_2}^{\alpha_{m+1}} + \eta_{i_2 i_3}^{\alpha_{m+1}} + \cdots + \eta_{i_m i_{m+1}}^{\alpha_{m+1}}).$$

By inserting this expression into equation (2.8), we have

$$d\Omega = -c \sum_{\boldsymbol{\alpha}} \sum_{\sigma \in \mathfrak{S}_m} \sum_{k=1}^{m} \Big\{ \operatorname{tr}(\eta_{i_1 i_2}^{\alpha_{\sigma(1)}} \circ \cdots \circ \eta_{i_m i_{m+1}}^{\alpha_{\sigma(m)}} \circ \eta_{i_k i_{k+1}}^{\alpha_{m+1}}) \Big\} \mathbf{dy}_{\boldsymbol{\alpha}}.$$

Finally, by exchanging the order of summations, we can write

$$d\Omega = -c \sum_{k=1}^{m} A_k, \tag{2.9}$$

where

$$A_k = \sum_{\boldsymbol{\alpha}} \sum_{\sigma \in \mathfrak{S}_m} \Big\{ \operatorname{tr}(\eta_{i_1 i_2}^{\alpha_{\sigma(1)}} \circ \cdots \circ \eta_{i_m i_{m+1}}^{\alpha_{\sigma(m)}} \circ \eta_{i_k i_{k+1}}^{\alpha_{m+1}}) \Big\} \mathbf{dy}_{\boldsymbol{\alpha}}.$$

Lemma 2.6. *For $2 \le k \le m-1$, we have $A_k = 0$.*

PROOF. Let us fix $k \in [2, m-1]$. In the expansion of A_k, for every multiindex $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_{m+1})$ there is a multiindex $\boldsymbol{\beta} = \boldsymbol{\beta}_{r,s} = (\beta_1, \dots, \beta_{m+1})$ differing from $\boldsymbol{\alpha}$ only for the exchange of two elements, at places r and s with $1 \leq r < s \leq m$; $\beta_r = \alpha_s$, $\beta_s = \alpha_r$, and $\beta_j = \alpha_j$ for $j \neq r, s$. Note that we have:

$$\begin{aligned} \mathbf{dy}_{\boldsymbol{\beta}} &= dy_{\beta_1} \wedge \cdots \wedge dy_{\beta_r} \wedge \cdots \wedge dy_{\beta_s} \wedge \cdots \wedge dy_{\beta_{m+1}} = \\ &= dy_{\alpha_1} \wedge \cdots \wedge dy_{\alpha_s} \wedge \cdots \wedge dy_{\alpha_r} \wedge \cdots \wedge dy_{\alpha_{m+1}} = \\ &= -dy_{\alpha_1} \wedge \cdots \wedge dy_{\alpha_r} \wedge \cdots \wedge dy_{\alpha_s} \wedge \cdots \wedge dy_{\alpha_{m+1}} = \\ &= -\mathbf{dy}_{\boldsymbol{\alpha}}. \end{aligned}$$

Now, for any such pair of multiindices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ and for any permutation σ of $\{1, 2, \dots, m\}$, let $\tau = \tau_{r,s}$ be the permutation given by the composition $\tau = \pi_{r,s} \circ \sigma$, where $\pi_{r,s}$ is the permutation that exchanges the elements at places r and s:

$$\pi_{r,s}(1, \dots, r, \dots, s, \dots, m) = (1, \dots, s, \dots, r, \dots, m).$$

It follows that

$$\eta_{i_1 i_2}^{\beta_{\tau(1)}} \circ \cdots \circ \eta_{i_m i_{m+1}}^{\beta_{\tau(m)}} \circ \eta_{i_k i_{k+1}}^{\beta_{m+1}} = \eta_{i_1 i_2}^{\alpha_{\sigma(1)}} \circ \cdots \circ \eta_{i_m i_{m+1}}^{\alpha_{\sigma(m)}} \circ \eta_{i_k i_{k+1}}^{\alpha_{m+1}},$$

hence the two terms

$$\operatorname{tr}(\eta_{i_1 i_2}^{\alpha_{\sigma(1)}} \circ \cdots \circ \eta_{i_m i_{m+1}}^{\alpha_{\sigma(m)}} \circ \eta_{i_k i_{k+1}}^{\alpha_{m+1}})\, \mathbf{dy}_{\boldsymbol{\alpha}}$$

and

$$\operatorname{tr}(\eta_{i_1 i_2}^{\beta_{\tau(1)}} \circ \cdots \circ \eta_{i_m i_{m+1}}^{\beta_{\tau(m)}} \circ \eta_{i_k i_{k+1}}^{\beta_{m+1}})\, \mathbf{dy}_{\boldsymbol{\beta}}$$

add to zero. Since all terms in the expansion of A_k can be paired in this way, we conclude that $A_k = 0$.

Now, it remains to consider the two terms A_1 and A_m.

LEMMA 2.7. *We have $A_1 = -A_m$.*

PROOF. Let us consider

$$A_1 = \sum_{\boldsymbol{\alpha}} \sum_{\sigma \in \mathfrak{S}_m} \Big\{ \operatorname{tr}(\eta_{i_1 i_2}^{\alpha_{\sigma(1)}} \circ \cdots \circ \eta_{i_m i_{m+1}}^{\alpha_{\sigma(m)}} \circ \eta_{i_1 i_2}^{\alpha_{m+1}}) \Big\} \mathbf{dy}_{\boldsymbol{\alpha}}.$$

By recalling the usual symmetry property of the trace, we may write

$$\operatorname{tr}(\eta_{i_1 i_2}^{\alpha_{\sigma(1)}} \circ \cdots \circ \eta_{i_m i_{m+1}}^{\alpha_{\sigma(m)}} \circ \eta_{i_1 i_2}^{\alpha_{m+1}}) = \operatorname{tr}\big((\eta_{i_1 i_2}^{\alpha_{m+1}} \circ \eta_{i_1 i_2}^{\alpha_{\sigma(1)}}) \circ \cdots \circ \eta_{i_m i_{m+1}}^{\alpha_{\sigma(m)}}\big),$$

hence, if we set $\xi_{ij}^{\alpha,\sigma} = \eta_{ij}^{\alpha_{m+1}} \circ \eta_{ij}^{\alpha_{\sigma(1)}} \in \Gamma(U_{ij}, \mathcal{E}nd(E))$, we have

$$A_1 = \sum_{\boldsymbol{\alpha}} \sum_{\sigma \in \mathfrak{S}_m} \Big\{ \operatorname{tr}(\xi_{i_1 i_2}^{\alpha,\sigma} \circ \eta_{i_2 i_3}^{\alpha_{\sigma(2)}} \circ \cdots \circ \eta_{i_m i_{m+1}}^{\alpha_{\sigma(m)}}) \Big\} \mathbf{dy}_{\boldsymbol{\alpha}}.$$

Now, from the skew-symmetry of the map Str in (1.10) (see Corollary 1.3), it follows that

$$\sum_{\sigma\in\mathfrak{S}_m} \operatorname{tr}(\xi^{\alpha,\sigma}_{i_1i_2}\circ\eta^{\alpha_{\sigma(2)}}_{i_2i_3}\circ\cdots\circ\eta^{\alpha_{\sigma(m)}}_{i_mi_{m+1}}) = -\sum_{\sigma\in\mathfrak{S}_m} \operatorname{tr}(\eta^{\alpha_{\sigma(m)}}_{i_1i_2}\circ\eta^{\alpha_{\sigma(2)}}_{i_2i_3}\circ\cdots\circ\xi^{\alpha,\sigma}_{i_mi_{m+1}}),$$

hence we have

$$A_1 = -\sum_{\boldsymbol{\alpha}}\sum_{\sigma\in\mathfrak{S}_m}\left\{\operatorname{tr}\left(\eta^{\alpha_{\sigma(m)}}_{i_1i_2}\circ\eta^{\alpha_{\sigma(2)}}_{i_2i_3}\circ\cdots\circ(\eta^{\alpha_{m+1}}\circ\eta^{\alpha_{\sigma(1)}})_{i_mi_{m+1}}\right)\right\}\mathbf{dy}_{\boldsymbol{\alpha}}.$$

On the other hand, we have

$$A_m = \sum_{\boldsymbol{\beta}}\sum_{\tau\in\mathfrak{S}_m}\left\{\operatorname{tr}\left(\eta^{\beta_{\tau(1)}}_{i_1i_2}\circ\eta^{\beta_{\tau(2)}}_{i_2i_3}\circ\cdots\circ(\eta^{\beta_{\tau(m)}}\circ\eta^{\beta_{m+1}})_{i_mi_{m+1}}\right)\right\}\mathbf{dy}_{\boldsymbol{\beta}}.$$

Now, for every multiindex $\boldsymbol{\alpha} = (\alpha_1,\ldots,\alpha_{m+1})$ and every permutation $\sigma\in\mathfrak{S}_m$ there is a multiindex $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_{m+1})$ and a permutation $\tau\in\mathfrak{S}_m$ such that

$$\beta_{\tau(1)}=\alpha_{\sigma(m)}, \beta_{\tau(2)}=\alpha_{\sigma(2)},\ldots,\beta_{\tau(m-1)}=\alpha_{\sigma(m-1)}, \beta_{\tau(m)}=\alpha_{m+1}, \beta_{m+1}=\alpha_{\sigma(1)}.$$

With this choice of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, σ and τ, we have

$$\begin{aligned}\mathbf{dy}_{\boldsymbol{\beta}} &= dy_{\beta_1}\wedge dy_{\beta_2}\wedge\cdots\wedge dy_{\beta_{m-1}}\wedge dy_{\beta_m}\wedge dy_{\beta_{m+1}} = \\ &= dy_{\alpha_m}\wedge dy_{\alpha_2}\wedge\cdots\wedge dy_{\alpha_{m-1}}\wedge dy_{\alpha_{m+1}}\wedge dy_{\alpha_1} = \\ &= (-1)^{2m-2} dy_{\alpha_1}\wedge dy_{\alpha_2}\wedge\cdots\wedge dy_{\alpha_{m+1}} = \\ &= \mathbf{dy}_{\boldsymbol{\alpha}}.\end{aligned}$$

It follows that $A_1 = -A_m$.

Now, from equation (2.9) and Lemmas 2.6 and 2.7, we obtain the following result:

THEOREM 2.8. *For any Y-flat family E of locally free sheaves on X and any integer m, with $1\le m\le \dim X$, the $H^m(X,\mathcal{O}_X)$-valued m-form $\Omega=\Omega_E$ on Y is closed, i.e., $d\Omega = 0$.*

3 – Differential Forms on Moduli Spaces

In this section we shall apply the results of Section 2 to the construction of closed holomorphic differential forms on moduli spaces of sheaves on a smooth complex projective variety X (we choose to work over the complex field because Theorem 2.8 was proved under the assumption $k=\mathbb{C}$, but everything we shall say holds true for any algebraically closed field k of characteristic 0).

So, let X be a smooth n-dimensional complex projective variety, let $\mathcal{O}_X(1)$ be a very ample invertible sheaf on X and let $\mathcal{M}$ be a moduli space of stable sheaves on X (with some fixed moduli data).

REMARK 3.1. On the moduli space $\mathcal{M}$ there does not exist, in general, a universal family of sheaves $\mathcal{E}$, not even locally in the Zariski topology. In any case, a universal family $\mathcal{E}$ on $\mathcal{M}$ exists locally in the complex analytic topology (or in the étale topology, if we are working over an algebraically closed field k of characteristic zero) [S, Theorem 1.21]. As noted in Remark 2.1, these local universal families are not uniquely determined, in fact they are defined only up to tensoring with the pull-back of a line bundle on $\mathcal{M}$. In general, these ambiguities prevent the local universal families to glue together to a globally defined one (see [Ma, Theorem 6.11] or [HL2, Section 4.6] for numerical conditions ensuring the existence of a global universal family on $\mathcal{M}$). On the other hand, when we consider the relative Ext-sheaves $\mathcal{E}xt^i_q(\mathcal{E},\mathcal{E})$ (or the sheaf $\mathcal{E}nd(\mathcal{E})$), these ambiguities disappear, and these locally defined sheaves glue together to a globally defined one on $\mathcal{M}$. For this reason, we shall abuse the notation and write $\mathcal{E}xt^i_q(\mathcal{E},\mathcal{E})$ (resp. $\mathcal{E}nd(\mathcal{E})$) even if the universal family $\mathcal{E}$ does not exist on $\mathcal{M}$.

Since the moduli space $\mathcal{M}$ is, in general, not smooth, we shall denote by $\mathcal{M}^{sm}$ its smooth locus. Analogously, we shall denote by $\mathcal{M}_{lf}$ the open subscheme of $\mathcal{M}$ parametrizing isomorphism classes of locally free sheaves and by $\mathcal{M}^{sm}_{lf}$ the smooth part of it.

For any $E \in \mathcal{M}^{sm}$ the Kodaira-Spencer map (2.1) gives a natural isomorphism

$$T_E\mathcal{M}^{sm} \cong \mathrm{Ext}^1(E,E). \tag{3.1}$$

If $E \in \mathcal{M}^{sm}_{lf}$, we also have

$$T_E\mathcal{M}^{sm}_{lf} \cong H^1(X,\mathcal{E}nd(E)), \tag{3.2}$$

because, for a locally free sheaf E, there are canonical isomorphisms

$$\mathrm{Ext}^i(E,E) \cong H^i(X,\mathcal{E}nd(E)).$$

The global versions of the Kodaira-Spencer isomorphisms (3.1) and (3.2) provide natural isomorphisms

$$T\mathcal{M}^{sm} \cong \mathcal{E}xt^1_q(\mathcal{E},\mathcal{E}),$$

and

$$T\mathcal{M}^{sm}_{lf} \cong R^1q_*\mathcal{E}nd(\mathcal{E}).$$

We can now apply the results of the preceding section to construct natural holomorphic differential forms on $\mathcal{M}^{sm}$.

More precisely, by setting $Y = \mathcal{M}^{sm}$ and denoting by $\mathcal{E}$ a locally defined universal family on Y (cf. also Remark 2.2), we have, for any m with $1 \leq m \leq n = \dim X$, a vector-valued m-form

$$\Omega : T\mathcal{M}^{sm} \times \cdots \times T\mathcal{M}^{sm} \to H^m(X, \mathcal{O}_X) \otimes_k \mathcal{O}_{\mathcal{M}^{sm}}. \tag{3.3}$$

Let us now assume that there exists a holomorphic m-form σ on X, $\sigma \in H^0(X, \Omega^m_X)$. The multiplication by σ defines a map

$$H^m(X, \mathcal{O}_X) \xrightarrow{\sigma} H^m(X, \Omega^m_X). \tag{3.4}$$

Finally, if we denote by $\eta_X \in H^1(X, \Omega^1_X)$ the cohomology class of the polarization $\mathcal{O}_X(1)$ (the cohomology class of the Kähler $(1,1)$-form on X), we have a map

$$H^m(X, \Omega^m_X) \xrightarrow{\eta_X^{n-m}} H^n(X, \Omega^n_X) \cong \mathbb{C}. \tag{3.5}$$

By composing the vector-valued differential form Ω with the maps (3.4) and (3.5), we obtain an ordinary (scalar-valued) m-form, which we denote by Ω_σ:

$$\Omega_\sigma : T\mathcal{M}^{sm} \times \cdots \times T\mathcal{M}^{sm} \to \mathcal{O}_{\mathcal{M}^{sm}}.$$

Since, by Theorem 2.8, the restriction of Ω to the open subset $\mathcal{M}^{sm}_{lf}$ parametrizing locally free sheaves is closed, it follows that the restriction of Ω_σ to $\mathcal{M}^{sm}_{lf}$ is a closed homolorphic m-form.

We can summarize these results as follows:

THEOREM 3.2. *For any holomorphic m-form σ on the complex projective variety X there is a holomorphic m-form Ω_σ on the smooth locus $\mathcal{M}^{sm}$ of the moduli space of stable sheaves on X. The restriction of Ω_σ to the smooth locus $\mathcal{M}^{sm}_{lf}$ of the moduli space of stable vector bundles on X is closed.*

Actually this result also holds with no assumption on the smoothness of the moduli space $\mathcal{M}$ or on the locally freeness of the stable sheaves. The method of proof is, however, completely different, and will be described in a different paper.

Let us describe now some particularly interesting special cases of Theorem 3.2.

Let us take $m = n = \dim X$ and assume that there exists a non-zero section σ of the canonical line bundle $K_X = \Omega^n_X$. In this case the map (3.5) is the identity, hence the n-form Ω_σ is given by the composition of Ω in (3.3) with the map

$$H^n(X, \mathcal{O}_X) \xrightarrow{\sigma} H^n(X, K_X) \cong \mathbb{C}.$$

Even more interesting is the case when the canonical line bundle of X is trivial, i.e., when X is a smooth Calabi-Yau n-fold. In fact, in this case there is a canonical choice (up to scalars) of the n-form σ on X, namely $\sigma = 1 \in H^0(X, K_X) \cong \mathbb{C}$, hence there is also a n-form Ω_σ on the moduli space $\mathcal{M}^{sm}$.

The natural question that arises at this point is to know under what conditions the canonical n-form Ω_σ on the moduli space $\mathcal{M}^{sm}$ is non-degenerate. We recall that, for $n = 2$, i.e., when X is an abelian or a K3 surface, this is always the case [Mu1]. For $n \geq 3$, on the other hand, there is no hope that Ω_σ be always non-degenerate; in fact there are examples of moduli spaces of stable sheaves (even stable vector bundles) on a smooth Calabi-Yau n-fold that are isomorphic to projective spaces.

We do not know the answer to this question but, in order to investigate the non-degeneracy of the n-form Ω_σ, when X is a Calabi-Yau n-fold or in the more general case of a smooth projective variety X with an effective canonical divisor, it may be helpful to use the following algebraic result (whose proof is elementary):

Proposition 3.3. *Let V be a finite dimensional k-vector space and*

$$\omega : \underbrace{V \times \cdots \times V}_{m} \to k$$

be an alternating, or symmetric, multilinear form. Let us define

$$\tilde{\omega} : \underbrace{V \times \cdots \times V}_{m-1} \to V^*$$

by setting

$$\langle v_1, \tilde{\omega}(v_2, \ldots, v_m) \rangle = \omega(v_1, v_2, \ldots, v_m),$$

for any $v_1, \ldots, v_m \in V$. Then the transpose of $\tilde{\omega}$ is the map

$$\tilde{\omega}^t : V \to \underbrace{V^* \times \cdots \times V^*}_{m-1}$$

given by

$$\langle \tilde{\omega}^t(v_1), (v_2, \ldots, v_m) \rangle = \omega(v_1, v_2, \ldots, v_m),$$

for any $v_1, \ldots, v_m \in V$, and we have

$$\mathrm{Ker}(\omega) = \mathrm{Ker}(\tilde{\omega}^t) = (\mathrm{Im}(\tilde{\omega}))^\perp,$$

where

$$\mathrm{Ker}(\omega) = \{v \in V \mid \omega(v, v_2, \ldots, v_m) = 0, \forall v_2, \ldots, v_m \in V\}.$$

Hence ω is non-degenerate if and only if $\tilde{\omega}$ is surjective or, equivalently, if and only if $\tilde{\omega}^t$ is injective.

In order to apply this result to our situation let us prove the following lemma (inspired by a similar result in [T]):

LEMMA 3.4. *Let $m = n = \dim X$ and let $\sigma \in H^0(X, K_X)$, with $\sigma \neq 0$. Let $E \in \mathcal{M}^{sm}$ and, using the notations of the preceding proposition, let us set $V = \mathrm{Ext}^1(E,E)$ and $\omega = \Omega_\sigma(E)$. Then, by Serre duality, we have $V^* \cong \mathrm{Ext}^{n-1}(E, E \otimes K_X)$, and the map*

$$\tilde{\omega} : \underbrace{\mathrm{Ext}^1(E,E) \times \cdots \times \mathrm{Ext}^1(E,E)}_{n-1} \to \mathrm{Ext}^{n-1}(E, E \otimes K_X) \tag{3.6}$$

is the composition of the map

$$S : \mathrm{Ext}^1(E,E) \times \cdots \times \mathrm{Ext}^1(E,E) \to \mathrm{Ext}^{n-1}(E,E)$$

induced by the symmetrized composition map (1.2)*, with the map*

$$\mathrm{Ext}^{n-1}(E,E) \xrightarrow{\sigma} \mathrm{Ext}^{n-1}(E, E \otimes K_X)$$

given by the multiplication by $\sigma \in H^0(X, K_X)$.

PROOF. The duality between $V = \mathrm{Ext}^1(E,E)$ and $V^* = \mathrm{Ext}^{n-1}(E, E \otimes K_X)$ is given by

$$\langle \phi, \phi^* \rangle = \mathrm{tr}(\phi \circ \phi^*) \in H^n(X, K_X) \cong \mathbb{C},$$

for any $\phi \in V$ and $\phi^* \in V^*$.

Let us now take $\phi_1, \ldots, \phi_n \in V$. By recalling the definition of Ω_σ, we have:

$$\Omega_\sigma(\phi_1, \ldots, \phi_n) = \sigma \ \mathrm{Str}(\phi_1, \ldots, \phi_n) = \sigma \, \frac{1}{n!} \sum_{\tau \in \mathfrak{S}_n} \mathrm{tr}(\phi_{\tau(1)} \circ \phi_{\tau(2)} \circ \cdots \circ \phi_{\tau(n)}).$$

By recalling the expression for the map Str given in (1.5), we can also write

$$\Omega_\sigma(\phi_1, \ldots, \phi_n) = \sigma \, \frac{1}{(n-1)!} \sum_{\tau} \mathrm{tr}(\phi_1 \circ \phi_{\tau(2)} \circ \cdots \circ \phi_{\tau(n)}), \tag{3.7}$$

where now the sum runs over all permutations τ of the set $\{2, 3, \ldots, n\}$.

Let us now set $\phi^* = \tilde{\omega}(\phi_2, \ldots, \phi_n)$. By recalling the definition of $\tilde{\omega}$, we have

$$\langle \phi_1, \phi^* \rangle = \omega(\phi_1, \phi_2, \ldots, \phi_n) = \Omega_\sigma(\phi_1, \ldots, \phi_n).$$

On the other hand, by the expression of Serre duality given above, we have

$$\langle \phi_1, \phi^* \rangle = \mathrm{tr}(\phi_1 \circ \phi^*). \tag{3.8}$$

Now, by comparing (3.8) with the expression in (3.7), we find that

$$\phi^* = \sigma \, \frac{1}{(n-1)!} \sum_{\tau} \phi_{\tau(2)} \circ \cdots \circ \phi_{\tau(n)} = \sigma \, S(\phi_2, \ldots, \phi_n).$$

It seems difficult to investigate, in general, the surjectivity of the map $\tilde{\omega}$ in (3.6). Obviously, a necessary condition is that the map

$$\operatorname{Ext}^{n-1}(E,E) \xrightarrow{\sigma} \operatorname{Ext}^{n-1}(E, E\otimes K_X) \tag{3.9}$$

be surjective. This is equivalent to requiring that the transpose of this map, i.e., the map

$$\operatorname{Ext}^1(E,E) \xrightarrow{\sigma} \operatorname{Ext}^1(E, E\otimes K_X),$$

be injective. By applying the functor $\operatorname{Hom}(E,\cdot)$ to the standard exact sequence

$$0 \longrightarrow E \xrightarrow{\sigma} E\otimes K_X \longrightarrow E\otimes K_X|_D \longrightarrow 0,$$

where $D \in |K_X|$ is the divisor defined by σ, we see that the map above fits into the long exact sequence

$$\begin{aligned} 0 \to \operatorname{Hom}(E,E) &\xrightarrow{\sigma} \operatorname{Hom}(E, E\otimes K_X) \to \operatorname{Hom}(E, E\otimes K_X|_D) \to \\ \to \operatorname{Ext}^1(E,E) &\xrightarrow{\sigma} \operatorname{Ext}^1(E, E\otimes K_X). \end{aligned}$$

From the stability of E it follows that $\operatorname{Hom}(E,E) \cong \mathbb{C}$, but it is difficult to get informations on $\operatorname{Hom}(E, E\otimes K_X)$ and $\operatorname{Hom}(E, E\otimes K_X|_D)$, in general. Obviously, this problem simply disappears when X is Calabi-Yau, i.e., when $K_X \cong \mathcal{O}_X$ and $\sigma = 1$ (in this case the map (3.9) is the identity).

Remark 3.5. If X is a smooth Calabi-Yau n-fold, it may happen that, for a suitable choice of moduli data, the corresponding moduli space $\mathcal{M}$ of stable sheaves on X has an irreducible component Y which is smooth, projective and of dimension n. Under these hypotheses, if the restriction of the canonical n-form Ω_σ to Y is non-degenerate, then Y will be a Calabi-Yau n-fold.

An example of this situation can be found in [T, Theorem 4.23]. In this case $n = 3$ and the Calabi-Yau 3-fold X is a K3 fibration over $\mathbb{P}^1$. The moduli space $\mathcal{M}$ is a relative moduli space of stable sheaves on X supported on the fibers (with suitable moduli data). The claim is that $\mathcal{M}$ is again a Calabi-Yau 3-fold. To prove this result, Thomas explicitly constructs a holomorphic 3-form on the moduli space $\mathcal{M}$ and shows that it is non-degenerate.

A similar (and more general) problem has been investigated by T. Bridgeland and A. Maciocia, under the additional assumptions that X is a flat Calabi-Yau fibration over a base S, with fibers of dimension ≤ 2, and $\mathcal{M}$ is a relative moduli space of stable sheaves supported on the fibers of $\pi : X \to S$. We refer to [BM] for details.

Remark 3.6 K. Yoshioka has constructed in [Y] moduli spaces of stable twisted sheaves on a smooth complex projective variety X. These are quasi-projective schemes and can be compactified, in the usual way, by adding S-equivalence classes of semistable twisted sheaves. In greater generality, moduli spaces of twisted sheaves have also been constructed by M. LIEBLICH in [Li], using the language of algebraic stacks. In any case, it turns out that the tangent space to such a moduli space at a point corresponding to a twisted sheaf E is canonically identified with $\mathrm{Ext}^1(E,E)$. From this fact it should follow immediately that our construction of closed differential forms Ω_σ on moduli spaces of stable sheaves can be generalized, in a straightforward way, to moduli spaces of stable twisted sheaves. When X is a K3 surface, this is explicitly proven in [Y]; in this case, the moduli space of stable twisted sheaves on X has a canonical symplectic structure (just as in the untwisted case [Mu1]).

We end this section with an example explaining how the construction of the differential forms Ω_σ can be adapted to other moduli spaces. We describe the case of moduli spaces of framed bundles.

3.1 – Moduli spaces of framed vector bundles

Let us now see how the construction of the differential forms Ω_σ can be extended to the case of moduli spaces of framed vector bundles.

Let X be a smooth n-dimensional complex projective variety with a very ample invertible sheaf $\mathcal{O}_X(1)$, and let $D \subset X$ be a smooth hypersurface. Let us denote by F a fixed vector bundle on D.

DEFINITION 3.7. A framed vector bundle on X is a pair (E,ϕ) consisting of a locally free sheaf E on X and an isomorphism $\phi : E|_D \xrightarrow{\sim} F$.

Moduli spaces of stable framed vector bundles on a smooth projective variety X were constructed in [Lu] and, in a more general context, in [HL1], to which we refer for definitions and results.

Let us denote by $\mathcal{FB}$ the moduli space of stable framed vector bundles on X (with fixed Hilbert polynomial) and by $\mathcal{FB}^{sm}$ its smooth locus. The moduli space $\mathcal{FB}$ is a quasi-projective variety, and it is actually a fine moduli space, i.e., there exists a (global) universal family of framed vector bundles on $\mathcal{FB}$.

Standard infinitesimal deformation theory gives the following result:

PROPOSITION 3.8. *For any $(E,\phi) \in \mathcal{FB}$ there is a canonical identification*

$$T_{(E,\phi)}\mathcal{FB} \cong H^1(X, \mathcal{E}nd(E) \otimes \mathcal{O}_X(-D)),$$

and the obstruction to the smoothness of the moduli space $\mathcal{FB}$ at the point (E,ϕ) lies in $H^2(X, \mathcal{E}nd_0(E) \otimes \mathcal{O}_X(-D))$, where $\mathcal{E}nd_0(E)$ denotes the sheaf of traceless endomorphisms of E.

In the sequel we shall denote $\mathcal{E}nd(E) \otimes \mathcal{O}_X(-D)$ simply by $\mathcal{E}nd(E)(-D)$.

In this situation we can define a $H^m(X, \mathcal{O}_X(-mD))$-valued differential m-form Ω on $\mathcal{FB}^{sm}$ by setting, for any $(E, \phi) \in \mathcal{FB}^{sm}$,

$$\begin{aligned}\Omega(E,\phi) : H^1(X, \mathcal{E}nd(E)(-D)) \times \cdots \times H^1(X, \mathcal{E}nd(E)(-D)) \to \\ \to H^m(X, \mathcal{O}_X(-mD)),\end{aligned}$$

where this map is the composition of the map

$$\begin{aligned}S : H^1(X, \mathcal{E}nd(E)(-D)) \times \cdots \times H^1(X, \mathcal{E}nd(E)(-D)) \to \\ \to H^m(X, \mathcal{E}nd(E) \otimes \mathcal{O}_X(-mD))\end{aligned}$$

induced by the symmetrized composition map, and the usual trace map

$$\mathrm{tr} : H^m(X, \mathcal{E}nd(E)(-mD)) \to H^m(X, \mathcal{O}_X(-mD)).$$

The proof of the closure of the m-form Ω on $\mathcal{FB}^{sm}$ is formally the same as the proof of the closure of the m-form Ω on Y given in Section 2.

Now, to construct from Ω a scalar-valued differential form on $\mathcal{FB}^{sm}$, we just need a section $\sigma' \in H^0(X, \Omega^m_X(mD))$, i.e., a global differential form of degree m on X, with poles bounded by mD. Then, the multiplication by σ' defines a map

$$H^m(X, \mathcal{O}_X(-mD)) \xrightarrow{\sigma'} H^m(X, \Omega^m_X),$$

hence by composing Ω with this map and then with the map η_X^{n-m} of (3.5), we obtain an ordinary (scalar-valued) m-form on $\mathcal{FB}^{sm}$, which we shall denote by $\Omega_{\sigma'}$. In conclusion, we have proved the following result:

THEOREM 3.9. *Let X, D and F be as above. For any meromorphic m-form σ' on X with poles bounded by mD, $\sigma' \in H^0(X, \Omega^m_X(mD))$, there is a closed holomorphic m-form $\Omega_{\sigma'}$ on the smooth locus $\mathcal{FB}^{sm}$ of the moduli space of stable framed vector bundles on X.*

Finally, let us investigate the relations between the differential forms constructed on the moduli spaces $\mathcal{M}^{sm}$ and $\mathcal{FB}^{sm}$.

Let X be as above and let $\sigma \in H^0(X, \Omega^m_X)$. Let D be a smooth hypersurface of X defined by a section $s \in H^0(X, \mathcal{O}_X(D))$. In this case there is an obvious choice for a global section σ' of $\Omega^m_X(mD)$, namely $\sigma' = \sigma\, s^m$.

Let us set

$$\mathcal{FB}^{sm}_0 = \{(E,\phi) \in \mathcal{FB}^{sm} \mid E \text{ is a stable vector bundle}\}.$$

(In general $\mathcal{FB}^{sm}_0 \neq \mathcal{FB}^{sm}$ because there can be framed bundles (E, ϕ) that are stable as framed bundles, but such that E is not stable as a vector bundle).

Then we have a natural map

$$\pi : \mathcal{FB}_0^{sm} \to \mathcal{M}$$

that forgets the framing, i.e., that sends a framed bundle (E, ϕ) to E.

With the natural identifications explained above, the tangent map to π at a point (E, ϕ) is the map

$$H^1(X, \mathcal{E}nd(E)(-D)) \xrightarrow{s} H^1(X, \mathcal{E}nd(E))$$

induced by the multiplication by $s \in H^0(X, \mathcal{O}_X(D))$.

Then we have a commutative diagram

$$\begin{array}{ccc} H^1(X, \mathcal{E}nd(E)(-D)) \times \cdots \times H^1(X, \mathcal{E}nd(E)(-D)) & \xrightarrow{\text{Str}} & H^m(X, \mathcal{O}_X(-mD)) \\ \big\downarrow{\scriptstyle s\times\cdots\times s} & & \big\downarrow{\scriptstyle s^m} \\ H^1(X, \mathcal{E}nd(E)) \times \cdots \times H^1(X, \mathcal{E}nd(E)) & \xrightarrow{\text{Str}} & H^m(X, \mathcal{O}_X). \end{array}$$

From this diagram, and from the preceding definitions of the m-forms on the moduli spaces, it is evident that the pull-back by π of the m-form Ω_σ defined on $\mathcal{M}^{sm}$ is equal to the m-form $\Omega_{\sigma'}$ defined on $\mathcal{FB}_0^{sm}$.

REFERENCES

[A] I. V. Artamkin: *On deformation of sheaves*, Math. USSR Izvestiya, **32** (1989), 663-668.

[B1] F. Bottacin: *Poisson structures on moduli spaces of sheaves over Poisson surfaces*, Invent. Math., **121** (1995), 421-436.

[B2] F. Bottacin: *Poisson structures on Hilbert schemes of points of a surface and integrable systems*, Manuscripta Math., **97** (1998), 517-527.

[B3] F. Bottacin: *Poisson structures on moduli spaces of framed vector bundles on surfaces*, Math. Nachr., **220** (2000), 33-44.

[B4] F. Bottacin: *Poisson structures on moduli spaces of parabolic bundles on surfaces*, Manuscripta Math., **103** (2000), 31-46.

[BM] T. Bridgeland, A. Maciocia: *Fourier-Mukai transforms for K3 and elliptic fibrations*, J. Algebraic Geom., **11** (2002), 629–657.

[C] A. Căldăraru: *Derived categories of twisted sheaves on Calabi-Yau manifolds*, Ph.D. thesis, Cornell University, 2000.

[HL1] D. Huybrechts, M. Lehn: *Framed modules and their moduli*, Int. J. Math., **6** (1995), 297-324.

[HL2] D. Huybrechts, M. Lehn: *The geometry of moduli spaces of sheaves*, Aspects of Mathematics, Vol. E31, Vieweg, 1997.

[K] S. Kobayashi: *Simple vector bundles over symplectic Kähler manifolds*, Proc. Japan Acad., **62** (1986), 21-24.

[Li] M. Lieblich: *Moduli of twisted sheaves*, Duke Math. J., **138** (2007), 23-118.

[Lu] M. Lübke: *The analytic moduli space of framed vector bundles*, J. Reine Angew. Math., **441** (1993), 45-59.

[M] S. Mac Lane: *Homology*, Springer-Verlag, 1967.

[Ma] M. Maruyama: *Moduli of stable sheaves II*, J. Math. Kyoto Univ., **18** (1978), 557-614.

[Mu1] S. Mukai: *Symplectic structure of the moduli space of sheaves on an abelian or $K3$ surface*, Invent. Math., **77** (1984), 101-116.

[Mu2] S. Mukai: *Moduli of vector bundles on $K3$ surfaces and symplectic manifolds*, Sugaku Expositions 1, **2** (1988), 138-174.

[S] C. Simpson: *Moduli of representations of the fundamental group of a smooth projective variety I*, Publ. Math. I.H.E.S., **79** (1994), 47-129.

[T] R.P. Thomas: *A holomorphic Casson invariant for Calabi-Yau 3-folds, and bundles on K3 fibrations*, J. Differential Geom., **53** (1999), 367-438.

[Y] K. Yoshioka: *Moduli spaces of twisted sheaves on a projective variety*, Moduli spaces and arithmetic geometry, 1-30, Adv. Stud. Pure Math., **45** Math. Soc. Japan, Tokyo, 2006.

Lavoro pervenuto alla redazione il 14 febbraio 2008
ed accettato per la pubblicazione il 4 marzo 2008.
Bozze licenziate il 30 settembre 2008

Indirizzo dell'Autore:

Francesco Bottacin – Dip. di Matematica e Informatica – Università degli Studi di Salerno – Via Ponte don Melillo – I–84084 Fisciano (SA) – Italy
E-mail: fbottacin@unisa.it

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 163–188

On the analogy between Arithmetic Geometry and foliated spaces

ERIC LEICHTNAM

ABSTRACT: *Christopher Deninger has developed an infinite dimensional cohomological formalism which allows to prove the expected properties of the arithmetical Zeta functions (including the Riemann Zeta function). These cohomologies are (in general) not yet constructed. Deninger has argued that these cohomologies might be constructed as leafwise cohomologies of suitable foliated spaces. We shall review some recent results which support this hope.*

1 – Introduction

Christopher Deninger's approach to the study of arithmetic zeta functions proceeds in two steps.

In the **first step**, he postulates the existence of infinite dimensional cohomology groups satisfying some "natural properties". From these data, he has elaborated a formalism allowing him to prove the expected properties for the arithmetic zeta functions: functional equation, conjectures of Artin, Beilinson, Riemann ... etc. There it is crucial to interpret the so called explicit formulae for the arithmetic zeta function as a Lefschetz trace formula.

The **second step** consists in constructing these cohomologies. Deninger has given some hope that these cohomologies might be constructed as leafwise cohomologies of suitable foliated spaces. Very little is known in this direction

KEY WORDS AND PHRASES: *Foliations – Dynamical – Lefschetz – Trace.*
A.M.S. CLASSIFICATION: 57R30 – 11F72 – 14H52

at the moment, but this second step seems to be a good motivation to develop interesting mathematics even if they are far from the ultimate goal.

In Section 2 we recall Deninger's cohomological formalism in the case of the Riemann zeta function. We point out a dissymmetry in the explicit formula (1) between the coefficients of $\delta_{k \log p}$ and $\delta_{-k \log p}$, see Comment 1.

In Section 3 is devoted to the description of the Lesfchetz trace formula for a flow acting on a codimension one foliated space. In Section 3.1 we recall the Guillemin-Sternberg trace formula which is indeed an important computational tool for this goal. By comparison with (1), it suggests that there should exist a flow $(\phi^t)_{t\in\mathbb{R}}$ acting on a certain space $S_{\mathbb{Q}}$ with the following property. To each prime number p [resp. the archimedean place of $\mathbb{Q}$] there should correspond a closed orbit with length $\log p$ [resp. a stationary point] of the flow ϕ^t.

In Section 3.2 we recall the theorem of Alvarez-Lopez and Kordyukov.

They consider a flow $(\phi^t)_{t\in\mathbb{R}}$ acting on $(X, \mathcal{F})$ where the compact three dimensonal manifold X is foliated by Riemann surfaces. They assume that $(\phi^t)_{t\in\mathbb{R}}$ preserves globally the foliation and is transverse to the foliation. Then Alvarez-Lopez and Kordyukov define a suitable leafwise Hodge cohomology on which ϕ^t acts and they prove a Lefschetz trace (à la Atiyah-Bott) formula which has some similarities with (1) for t real positive. But the dissymmetry mentioned above for (1) does not hold.

In Section 3.2 we consider the case of an elliptic curve E_0 over a finite field $\mathbb{F}_q$. The explicit formula (7) for its zeta function ζ_{E_0} exhibits a dissymmetry between the coefficients of $\delta_{k \log Nw}$ and $\delta_{-k \log Nw}$, where w is a closed point of E_0. It is quite analogous to the one mentioned above for (1). We review briefly our result which, using the work of Deninger and results from Alvarez-Lopez and Kordyukov, allows to interpret (7) as an Atiyah-Bott Lefschetz trace formula and to provide a dynamical interpretation of this dissymmetry.

In Section 4, we first recall the statement of Lichtenbaum's conjecture for a number field K. Then we explain briefly how Deninger proved an analogue of this conjecture in the case of a foliation $(X, \mathcal{F}, \phi^t)$ with the following properties. X is a smooth compact 3-dimensional manifold endowed with a codimension 1 foliation $\mathcal{F}$ and the flow ϕ^t preserves globally the foliation and is transverse to it. We shall explain how the reduced leafwise cohomology enters as a crucial ingredient of the proof.

In Section 5, we make a synthesis of various results of Deninger. We state several axioms for a laminated foliated space $(S_{\mathbb{Q}}, \mathcal{F}, g, \phi^t)$ which (if satisfied!) allow to construct the required cohomology groups for the Riemann zeta function. We compare carefully the contribution of the archimedean place of $\mathbb{Q}$ in (1) with the contribution of a stationary point in the Guillemin-Sternberg formula.

2 – Deninger's Cohomological formalism in the case of the Riemann zeta function

The (completed) Riemann zeta function is given by:

$$\widehat{\zeta}(s) = 2^{-1/2}\pi^{-s/2}\Gamma(s/2)\prod_{p\in\mathcal{P}}\frac{1}{1-p^{-s}}$$

where $\mathcal{P} = \{2, 3, 5, \dots\}$ denotes the usual set of prime numbers. The following well known explicit formulas express a connection between $\mathcal{P}\cup\{\infty\}$ and the zeroes of $\widehat{\zeta}$. Let $\alpha \in C^\infty_{\text{compact}}(\mathbb{R},\mathbb{R})$ and for real s, set $\Phi(s) = \int_{\mathbb{R}} e^{st}\alpha(t)dt$; Φ belongs to the Schwartz class $\mathcal{S}(\mathbb{R})$. Then one can prove the following formula:

$$\begin{aligned}&\Phi(0) - \sum_{\rho\in\widehat{\zeta}^{-1}\{0\},\,\Re\rho\geq 0}\Phi(\rho) + \Phi(1) = \\ &= \sum_{p\in\mathcal{P}}\log p\left(\sum_{k\geq 1}\alpha(k\log p) + \sum_{k\leq -1}p^k\alpha(k\log p)\right) + W_\infty(\alpha),\end{aligned} \tag{1}$$

where

$$W_\infty(\alpha) = \alpha(0)\log\pi + \int_0^{+\infty}\left(\frac{\alpha(t)+e^{-t}\alpha(-t)}{1-e^{-2t}} - \alpha(0)\frac{e^{-2t}}{t}\right)dt.$$

Now recall the standard Lefschetz trace formula for a smooth map with non degenerate fixed points $\phi : V \to V$ where V is an oriented compact Riemann surface:

$$\sum_{j=0}^{2}(-1)^j\,\mathrm{TR}(\phi^* : H^j(V,\mathbb{R})) = \sum_{\phi(v)=v}(-1)\mathrm{sign}\det(\mathrm{Id} - D\phi)(v).$$

Deninger's philosophy is motivated by the fact that the left hand side of (1)

$$\Phi(0) - \sum_{\rho\in\widehat{\zeta}^{-1}\{0\},\,\Re\rho\geq 0}\Phi(\rho) + \Phi(1)$$

is reminiscent of a Lefschetz trace formula of the form

$$\mathrm{TR}\int_{\mathbb{R}}\alpha(t)e^{t\,\Theta_0}\,dt - \mathrm{TR}\int_{\mathbb{R}}\alpha(t)e^{t\,\Theta_1}\,dt + \mathrm{TR}\int_{\mathbb{R}}\alpha(t)e^{t\,\Theta_2}\,dt,$$

where the following two assumptions should be satisfied.

- $\Theta_0 = 0$ acts on $H^0 = \mathbb{R}$, $\Theta_2 = \mathrm{Id}$ acts on $H^2 = \mathbb{R}$.
- The unbounded operator, Θ_1 acts on an infinite dimensional real vector (pre-Hilbert) space H^1, for any $\alpha \in C^\infty_{\text{compact}}(\mathbb{R}, \mathbb{R})$ the operator $\int_{\mathbb{R}} \alpha(t) e^{t\,\Theta_1}\, dt$ is trace class. The eigenvalues of $\Theta_1 \otimes Id_{\mathbb{C}}$ acting on $H^1 \otimes_{\mathbb{R}} \mathbb{C}$ coincide with the non trivial zeroes of $\widehat{\zeta}$.

Moreover in Deninger's approach one first assumes the existence of a Poincare duality pairing:

$$H^1 \times H^1 \to H^2$$

$$(\alpha, \beta) \to \alpha \cup \beta$$

satisfying

$$\forall \alpha, \beta \in H^1,\ e^{t\,\Theta_1}\alpha \cup e^{t\,\Theta_1}\beta = e^t(\alpha \cup \beta), \tag{2}$$

where the e^t is dictated by the fact that $\Phi_2 = Id$.

Second, one assumes the existence of a Hodge star $\star$ on H^1 such that $e^{t\,\Theta_1}\star = \star e^{t\,\Theta_1}$ and $< \alpha; \beta >= \alpha \cup \star\beta$ defines a real scalar product on the real vector space H^1.

Then with these data, Deninger's formalism implies the following:

$$\forall \alpha \in H^1,\ \langle e^{t\,\Theta_1}\alpha; e^{t\,\Theta_1}\alpha \rangle = e^t \langle \alpha; \alpha \rangle. \tag{3}$$

Therefore,

$$\frac{d}{dt} \langle e^{t\,\Theta_1}\alpha; e^{t\,\Theta_1}\alpha \rangle_{t=0} = \langle \Theta_1(\alpha); \alpha \rangle + \langle \alpha; \Theta_1(\alpha) >=< \alpha; \alpha \rangle,$$

and

$$\langle (\Theta_1 - 1/2)(\alpha); \alpha \rangle + \langle \alpha; (\Theta_1 - 1/2)(\alpha) \rangle = 0.$$

Thus one gets that $\Theta_1 - \frac{1}{2}$ is antisymmetric on the real vector space H^1. Therefore, the eigenvalues s of Θ_1 (which coincide by (1) to the non trivial zeroes of $\widehat{\zeta}$) satisfy $s - 1/2 \in i\mathbb{R}$ or equivalently: $\Re s = \frac{1}{2}$. Therefore Deninger's formalism should imply the Riemann hypothesis!!! This argument comes from an idea of Serre [Se60] and has been formalized in the foliation case in [De-Si02]. Of course, we have described only a very small part of Deninger's formalism which deals also with L-functions of motives, Artin conjecture, Beilinson conjectures ... etc.

Comment 1. There is a dissymmetry in (1) between the coefficients of $\alpha(k \log p)$ and $\alpha(-k \log p)$ for $k \in \mathbb{N}^*$. In the framework of Deninger's formalism the explanation is the following. Equation (2) allows to prove (3) which in turn implies that the transpose of $e^{t\Theta_1}$ is $e^t e^{-t\Theta_1}$. Therefore, if we have a Lefschetz cohomological interpretation of (1) in Deninger's formalism for a test function α with support in $]0, +\infty[$ then we have also a cohomological proof of (1) for α

with support in $]-\infty, 0[$. In this formalism, (2) (and the above dissymmetry) is quite connected to the Riemann hypothesis.

Notice that Ralf Meyer [Meyer03] has provided a nice spectral interpretation of the zeroes of $\widehat{\zeta}$ and an original proof of (1). Unfortunately, he cannot prove the Riemann Hypothesis because he is obliged to work with Frechet spaces rather than Hilbert spaces. To our opinion, the geometry underlying his constructions is not sufficient. Recall that also Alain Connes [Co99] has reduced the validity of the Riemann hypothesis (for $L-$function of the Hecke characters) to a trace formula.

The idea of the proof of (1) is the following: apply the residue theorem to

$$\left(\int_0^{+\infty} \sqrt{t}\,\alpha(\log t) t^s \frac{d\,t}{t}\right) \frac{\widehat{\zeta}'}{\widehat{\zeta}}(s)$$

on the interior of the rectangle of $\mathbb{C}$ defined by the four points:

$$1+\epsilon+iT,\ -\epsilon+iT,\ -\epsilon-iT,\ 1+\epsilon-iT,$$

then use the functional equation $\widehat{\zeta}(s) = \widehat{\zeta}(1-s)$ and the formula:

$$\frac{\Gamma'}{\Gamma}\left(\frac{s}{2}+\frac{1}{4}\right) = \int_0^{+\infty} \left(\frac{e^{-u}}{u} - \frac{e^{-u(\frac{s}{2}+\frac{1}{4})}}{1-e^{-u}}\right) du,$$

lastly let T goes to $+\infty$.

3 – Analogy with the foliation case

3.1 – The Guillemin-Sternberg trace formula

Consider a smooth compact manifold X with a smooth action:

$$\phi : X \times \mathbb{R} \to X,\ (x,t) \to \phi^t(x),$$

so that $\phi^{t+t'} = \phi^t \circ \phi^{t'}$ for any $t, t' \in \mathbb{R}$. Let $D_y\phi^t$ denote (for fixed $t \in \mathbb{R}$) the differential of the map $y \in X \to \phi^t(y)$. One has: $D_y\phi^t(\partial_s \phi^s_{|s=0}(x)) = \partial_s \phi^s_{|s=0}(x)$. In other words, the vector field associated with the flow ϕ^t belongs to $\ker(D_y\phi^t - \mathrm{Id})$.

Consider also a smooth vector bundle $E \to X$. Assume that E is endowed with a smooth family of maps

$$\psi^t : (\phi^t)^* E \to E,\ t \in \mathbb{R},$$

satisfying the following cocycle condition:

$$\forall u \in C^\infty(X;E),\ \forall t,t' \in \mathbb{R},\ \psi^{t'}(\psi^t(u\circ\phi^t)\circ\phi^{t'}) = \psi^{t+t'}(u\circ\phi^{t+t'}).$$

So we require that the maps $K^t : u \to \psi^t(u\circ\phi^t) = K^t(u)$ define an action of the additive group $\mathbb{R}$ on $C^\infty(X;E)$. Notice that in the case of $E = \wedge^* T^* X$ and $\psi^s = {}^t D\phi^s$ (the transpose of the differential of ϕ^s) this condition is satisfied.

We shall assume that the graph of ϕ meets transversally the "diagonal" $\{(x,x,t),\ x\in X,\ t\in\mathbb{R}\setminus\{0\}\}$. Guillemin-Sternberg have checked ([G-S77]) that the trace $\mathrm{Tr}(K^t|C^\infty(X;E))$ is defined as a distribution of $t\in\mathbb{R}\setminus\{0\}$ by the formula:

$$\mathrm{Tr}(K^t|C^\infty(X;E)) = \int_X K^t(x,x)$$

where $K^t(x,y)$ denote Schwartz (density) kernel of K^t. We warn the reader that, in general, for $\alpha\in C^\infty_{\mathrm{compact}}(\mathbb{R})\setminus\{0\}$, $\int_{\mathbb{R}}\alpha(t)K^t dt$ is not trace class.

Now, we give the name $T^0_x = \partial_t\phi^t(x)_{t=0}\,\mathbb{R}$ to the real line generated by the vector field $\partial_t\phi^t(x)_{t=0}$ of ϕ^t at a point x where $\partial_t\phi^t(x)_{t=0}\neq 0$.

Proposition 1 (Guillemin-Sternberg [G-S77]). *The following formula holds in $\mathcal{D}'(\mathbb{R}\setminus\{0\})$.*

$$\mathrm{Tr}(K^t;C^\infty(X;E)) = \sum_\gamma l(\gamma)\sum_{k\in\mathbb{Z}^*}\frac{\mathrm{Tr}(\psi^{kl(\gamma)}_{x_\gamma};E_{x_\gamma})}{|\det(1-D_y\phi^{kl(\gamma)}(x_\gamma)\,;\,T_{x_\gamma}X/T^0_{x_\gamma})|}\delta_{kl(\gamma)} + \\ + \sum_x\frac{\mathrm{Tr}(\psi^t_x;E_x)}{|\det(1-D_y\phi^t(x);T_xX)|}.$$

In the first sum, γ runs over the periodic primitive orbits of ϕ^t, x_γ denotes any point of γ, $l(\gamma)$ is the length of γ, $\phi^{l(\gamma)}(x_\gamma)=x_\gamma$. In the second sum, x runs over the fixed points of the flow: $\phi^t(x)=x$ for any $t\in\mathbb{R}$.

Comment 2. Recall that $D_y\phi^t$ denotes, for fixed t, the differential of the map $y(\in X)\to\phi^t(y)$. The non vanishing of the two determinants in Proposition 1 is equivalent to the fact that the graph of ϕ meets transversally the "diagonal" $\{(x,x,t),\ x\in X,\ t\in\mathbb{R}\setminus\{0\}\}$.

Note that the following elementary observation is the main ingredient of the proof the Proposition 1. It is important with respect to Subsection 3.3. Let $A\in GL_n(\mathbb{R})$ and $\delta_0(\cdot)$ denote the Dirac mass at $0\in\mathbb{R}^n$. Then one computes the distribution $\delta_0(A\cdot)$ in the following way. For any $f\in C^\infty_{\mathrm{comp}}(\mathbb{R}^n)$, one has:

$$\langle\delta_0(A\cdot);f(\cdot)\rangle = \int_{\mathbb{R}^n}\delta_0(Ax)f(x)dx = \int_{\mathbb{R}^n}\delta_0(y)f(A^{-1}y)\frac{1}{\mathrm{Jac}(A)}dy = \frac{1}{\mathrm{Jac}(A)}f(0)$$

where dy denotes the Lebesgue measure. Therefore: $\delta_0(A\cdot) = \frac{1}{\mathrm{Jac}(A)}\delta_0(\cdot)$.

3.2 – The Lesfchetz trace formula of Alvarez-Lopez and Kordyukov

Now we shall assume that X is a (still compact) three dimensional manifold and endowed with a codimension one foliation $(X, \mathcal{F})$. We shall also assume that the flow ϕ^t preserves the foliation $(X, \mathcal{F})$, is transverse to it and thus has no fixed point. Therefore $(X, \mathcal{F})$ is a Riemannian foliation. We shall apply later Proposition 1 with $E = \wedge^* T^* \mathcal{F} \to X$.

COMMENT 3. A typical example is $X = \frac{L \times \mathbb{R}^{+*}}{\Lambda}$, where Λ is a subgroup of $(\mathbb{R}^{+*}, \times)$ and $\phi^t(l, x) = (l, xe^{-t})$. See Section 4.

Now, we get a so called bundle like metric g_X on $(X, \mathcal{F})$ in the following way. We require that $g_X(\partial_t \phi^t(z)) = 1$, $\partial_t \phi^t(z) \perp T\mathcal{F}$ for any $(t, z) \in \mathbb{R} \times X$, and that $(g_X)_{|T\mathcal{F}}$ is a given leafwise metric. By construction, with respect to g_X, the foliation $(X, \mathcal{F})$ is defined locally by riemannian submersions.

In this setting Alvarez-Lopez and Kordyukov [A-K01] have proved the following Hodge decomposition theorem $(0 \leq j \leq 2)$:

$$C^\infty(X, \wedge^j T^* \mathcal{F}) = \ker \Delta_\tau^j \oplus \overline{\operatorname{Im} \Delta_\tau^j} \tag{4}$$

where Δ_τ^j denotes the leafwise Laplacian. Since we have $\frac{\ker d_\mathcal{F}}{\overline{\operatorname{Im} d_\mathcal{F}}} = \ker \Delta_\tau^j$, we call the vector space $H_\tau^j = \ker \Delta_\tau^j$ a reduced leafwise cohomology group.

Let π_τ^j denote the projection of the vector space of leafwise differential forms $C^\infty(X, \wedge^j T^* \mathcal{F})$ onto $H_\tau^j = \ker \Delta_\tau^j$ according to (4) with $0 \leq j \leq 2$. Then Alvarez-Lopez and Kordyukov [A-K00] have proved the following Lefschetz trace formula.

THEOREM 1 ([A – K00]). *Let* $\alpha \in C^\infty_{\text{compact}}(\mathbb{R})$ *Then the operators*

$$\int_{\mathbb{R}} \alpha(s)\, \pi_\tau^j \circ (\phi^s)^* \circ \pi_\tau^j \, ds$$

are trace class for $0 \leq j \leq 2$. *Let* χ_Λ *denote the leafwise measured Connes Euler characteristic of* $(X, \mathcal{F})$ ([Co94]). *Then one has*:

$$\begin{aligned} &\sum_{j=0}^{2} (-1)^j \mathrm{TR} \int_{\mathbb{R}} \alpha(s)\, \pi_\tau^j \circ (\phi^s)^* \circ \pi_\tau^j \, ds = \\ &= \chi_\Lambda \alpha(0) + \sum_\gamma \sum_{k \geq 1} l(\gamma) \left(\epsilon_{-k\gamma} \alpha(-kl(\gamma)) + \epsilon_{k\gamma} \alpha(kl(\gamma)) \right) \end{aligned} \tag{5}$$

where γ *runs over the primitive closed orbits of* ϕ^t, $l(\gamma)$ *is the length of* γ, $x_\gamma \in \gamma$ *and* $\epsilon_{\pm k\gamma} = \operatorname{sign} \det(\mathrm{id} - D\phi^{\pm kl(\gamma)}_{|\mathrm{T}_{x_\gamma} \mathcal{F}})$.

Proof. (Sketch of the idea). The case where the support of α is included in a suitably small interval $[-\epsilon, +\epsilon]$ is treated separately. When the (compact) support of α is included in $\mathbb{R} \setminus \{0\}$, the authors show by highly non trivial arguments based on (4), that

$$\sum_{j=0}^{2}(-1)^j \mathrm{TR} \int_{\mathbb{R}} \alpha(s)\, \pi_\tau^j \circ (\phi^s)^* \circ \pi_\tau^j \, ds \; = \sum_{j=0}^{2}(-1)^j \mathrm{Tr} \int_{\mathbb{R}} \alpha(s) (\phi^s)^* \, ds.$$

But Proposition 1 (with $E = \wedge^j T^* \mathcal{F}$) shows that the right handside is equal to:

$$\sum_\gamma l(\gamma) \sum_{k \in \mathbb{Z}^*} \sum_{j=0}^{2} (-1)^j \, \frac{\mathrm{Tr}\big((D_y \phi^{\pm kl(\gamma)}(x_\gamma))^* : \wedge^j \mathrm{T}^*_{x_\gamma} \mathcal{F} \mapsto \wedge^j \mathrm{T}^*_{x_\gamma} \mathcal{F}\big)}{|\det(\mathrm{id} - D_y \phi^{\pm kl(\gamma)}_{|\mathrm{T}_{x_\gamma}\mathcal{F}})|} \alpha(kl(\gamma)).$$

One then gets immediately the result.

Comment 4. Notice that here (unlike in (1)) there is no dissymmetry for the coefficients of $\alpha(-kl(\gamma))$ and $\alpha(kl(\gamma)$. The reason for this absence of dissymmetry is explained by the Guillemin-Sternberg formula as we have seen in the proof. Here is another way to rephrase this explanation when X is orientable. If we had a leafwise metric g satisfying $(\phi^t)^* g = e^t g$ then we should get in (5) the same dissymmetry as the one already mentioned in (1). Assume that for a fix real β one has $\forall t \in \mathbb{R}$, $(\phi^t)^* g = e^{\beta t} g$. Consider the bundlelike metric g_X on X as above. Its volume form ω_X is such that $(\phi^t)^* \omega_X = e^{\beta t} \omega_X$, $\forall t \in \mathbb{R}$. But we know that the degree $\frac{\int_X (\phi^t)^* \omega_X}{\int_X \omega_X}$ has to be an integer for any $t \in \mathbb{R}$. Therefore $\beta = 0$.

The Ruelle zeta function is defined by

$$\zeta_R(s) = \prod_{\gamma \,\text{primitive orbit}} \frac{1}{(1 - e^{-sl(\gamma)})^{\epsilon_\gamma}}, \quad \Re s \gg 1.$$

The induced action of $(\phi^s)^*$ on H^j_τ is of the form $e^{s\theta_j}$. Deninger's results (e.g. [De98], [De07a]) suggest to conjecture that for $0 \leq j \leq 2$, $s \to \det_\infty(s \,\mathrm{Id} - \theta_j : H^j_\tau)$ defines an entire holomorphic function and that

$$\zeta_R(s) = \prod_{j=0}^{2} \det{}_\infty(s \,\mathrm{Id} - \theta_j : H^j_\tau))^{(-1)^{j+1}}$$

where $\det_\infty$ denotes an infinite regularized determinant (see [De94] for definitions). If this last equality is true then (5) should constitute the explicit formula for ζ_R. Notice moreover that in Theorem 1 there is no term similar to

$W_\infty(\alpha)$ in (1) because the flow $(\phi^t)_{t\in\mathbb{R}}$ is assumed to have no fixed (e.g. stationary) point. Assume morever that there exists a leafwise metric g such that $\forall t \in \mathbb{R}$, $(\phi^t)^* g = g$. Then, by considering the associated bundlelike metric g_X one defines easily a scalar product $\langle ; \rangle$ on H^1_τ such that $e^{t\theta_1}$ becomes a unitary operator on the Hilbert completion of H^1_τ. Therefore, all the zeroes of ζ_R are on the real line $\Re z = 0$.

Recall that it is not always possible to find a leafwise metric g such that $\forall t \in \mathbb{R}$, $(\phi^t)^* g = g$. Here is an example communicated to me by Alvarez-Lopez. Let h be a diffeomorphism of S^2 fixing the two poles. Assume that for the corresponding $\mathbb{Z}$-action the north pole is attractive and the south pole is repulsive. Set $X = \frac{S^2\times\mathbb{R}}{\mathbb{Z}}$ where the action of $m \in \mathbb{Z}$ is defined by $m \cdot (l, x) = (h^m(l), m + x)$. The sets $S^2 \times \{x\}$ induce a foliation $\mathcal{F}$. Consider the flow ϕ^t defined by $\phi^t(l, x) = (l, x + t)$. Then there is no leafwise metric g on $(X, \mathcal{F})$ such that $\forall t \in \mathbb{R}$, $(\phi^t)^* g = g$.

Comment 5. Alvarez-Lopez and Kordyukov are working on a proof of a Lefschetz trace formula when the flow ϕ^t is allowed to have stationary points. They work with a notion of “adiabatic cohomology” and their programme is promising.

3.3 – Foliated spaces with a p–adic transversal

We shall now describe an example of foliated space where on can prove a Lefschetz trace formula exhibiting a dissymmetry quite similar to the one mentioned in Comment 1.

Let E_0 be an elliptic curve over a finite field $\mathbb{F}_q$ where $q = p^f$ and the prime number p is the characteristic of $\mathbb{F}_q$. Recall that the zeta function $\zeta_{E_0}(s)$ of E_0 is given by:

$$\zeta_{E_0}(s) = \prod_{w\in|E_0|} \frac{1}{1-(Nw)^{-s}} = \frac{(1-\xi q^{-s})(1-\overline{\xi} q^{-s})}{(1-q^{-s})(1-q^{1-s})} \tag{6}$$

where $|E_0|$ denotes the set of closed points of E_0 and ξ is a complex number which by Hasse’s theorem satisfies $|\xi| = \sqrt{q}$. The explicit formula for $\zeta_{E_0}(s)$ takes the following form. Let $\alpha \in C_c^\infty(\mathbb{R}, \mathbb{R})$ and set for any real s, $\Phi(s) = \int_{\mathbb{R}} e^{st}\alpha(t)\,dt$. Then, one has:

$$\begin{aligned} &\sum_{\nu\in\mathbb{Z}} \Phi\left(\frac{2\pi\nu i}{\log q}\right) - \sum_{\rho\in\zeta_{E_0}^{-1}\{0\}} \Phi(\rho) + \sum_{\nu\in\mathbb{Z}} \Phi\left(1+\frac{2\pi\nu i}{\log q}\right) = \\ &= \sum_{w\in|E_0|} \log Nw \left(\sum_{k\geq 1} \alpha(k\log Nw) + \sum_{k\leq -1} (Nw)^k \alpha(k\log Nw)\right). \end{aligned} \tag{7}$$

The idea of the proof is to apply the residue theorem to

$$s \to \left(\int_0^{+\infty} \sqrt{t}\, \alpha(\log t) t^s \frac{d\, t}{t} \right) \frac{\zeta'_{E_0}}{\zeta_{E_0}}(s)$$

and to use the functional equation $\zeta_{E_0}(s) = \zeta_{E_0}(1-s)$.

Let $\phi_0 : E_0 \to E_0$ be the $q-$th power Frobenius endomorphism of E_0 over $\mathbb{F}_q$. Deninger has used (see [De02]) the following result due to Oort [Oor73]:

LEMMA 1. *There exists*:

1) *a complete local integral domain R with field of fractions L a finite extension of $\mathbb{Q}_p$ $(q = p^f)$ such that $R/\mathcal{M} = \mathbb{F}_q$ where $\mathcal{M}$ is the maximal ideal of R.*
2) *an elliptic curve $\mathcal{E}$ over spec R together with an endomorphism $\phi : \mathcal{E} \to \mathcal{E}$ such that*:

$$(\mathcal{E}, \phi) \otimes \mathbb{F}_q = (E_0, \phi_0).$$

So $(\mathcal{E}, \phi)$ is a lift of (E_0, ϕ_0) in characteristic zero.

REMARK 1.

1) If the elliptic curve E_0 is ordinary, then one may take for R the ring of Witt vectors of $\mathbb{F}_q$, $W(\mathbb{F}_q)$, and then there is a canonical choice of the lifting $(\mathcal{E}, \phi)$. On the contrary, if E_0 is supersingular [Si92, page 137], then there is no canonical choice of $(\mathcal{E}, \phi)$.
2) It is possible to lift a curve of genus ≥ 2 (over $\mathbb{F}_q$) in characteristic zero, but Hurwitz's formula [Si92, page 41] shows that one cannot lift its Frobenius morphism.

Now (still following [De02]), we denote by $E = \mathcal{E} \otimes_R L$ the generic fibre. Then $\mathrm{End}_L(E) \otimes \mathbb{Q} = K$ is a field K which is either $\mathbb{Q}$ or an imaginary quadratic extension of $\mathbb{Q}$. We fix an embedding $L \subset \mathbb{C}$ and consider the complex analytic elliptic curve $E(\mathbb{C})$. Let ω be a non zero holomorphic one form on $E(\mathbb{C})$ and let Γ be its period lattice. Then the Abel-Jacobi map:

$$E(\mathbb{C}) \to \mathbb{C}/\Gamma,\ p \to \int_0^p \omega \mod \Gamma$$

induces an isomorphism. Next we choose the embedding $K \subset \mathbb{C}$ such that for any $\alpha \in K$, $\Theta(\alpha)$ induces the multiplication by α on the Lie algebra $\mathbb{C}$ of $\mathbb{C}/\Gamma$ where Θ is the natural homomorphism:

$$\Theta :\ K = \mathrm{End}_L(E) \otimes \mathbb{Q} \to \mathrm{End}(\mathbb{C}/\Gamma) \otimes \mathbb{Q}.$$

Next we consider the unique element $\xi \in \Theta^{-1}(\mathrm{End}_L(E)) \subset K$ such that $\Theta(\xi) = \phi \otimes L$. By construction one has $\xi\Gamma \subset \Gamma$ and the complex elliptic curve $\mathbb{C}/\Gamma$ endowed with the multiplication by ξ represents a lift of (E_0, ϕ_0). Now, we set

$$V = \cup_{n\in\mathbb{N}}\xi^{-n}\Gamma, \ \ \mathrm{T}\Gamma = \lim_{+\infty\leftarrow n} \frac{\Gamma}{\xi^n\Gamma}, \ \text{ and } \ V_\xi\Gamma = \mathrm{T}\Gamma \otimes_{\mathbb{Z}} \mathbb{Q}.$$

The set $\mathrm{T}\Gamma$ is a Tate module defined by a projective limit and $V_\xi\Gamma$ is a $\mathbb{Q}_p$-vector space of dimension 1 or 2. Any element v of V acts on $\mathbb{C} \times V_\xi\Gamma$ by $v.(z, \hat{v}) = (z + v, \hat{v} - v)$, we denote by $\frac{\mathbb{C}\times V_\xi\Gamma}{V}$ the quotient space.

Lemma 2. *The natural homomorphism*:

$$\frac{\mathbb{C} \times \mathrm{T}\Gamma}{\Gamma} \to \frac{\mathbb{C} \times V_\xi\Gamma}{V}$$

defines a $\{\xi^l,\ l \in \mathbb{Z}\}$*-equivariant isomorphism where the action of* ξ *is induced by the diagonal action on* $\mathbb{C} \times \mathrm{T}\Gamma$ *and* $\mathbb{C} \times V_\xi\Gamma$ *respectively.*

Now, any element $q^\nu \in q^{\mathbb{Z}}$ acts on $\frac{\mathbb{C}\times V_\xi\Gamma}{V} \times \mathbb{R}^{+*}$ by

$$q^\nu.([z, \hat{v}], x) = ([\xi^\nu z, \xi^\nu \hat{v}], xq^\nu).$$

In [De02], Deninger has introduced the (compact) laminated Riemannian foliated space $(S(E_0), \mathcal{F})$ where

$$S(E_0) \ = \ \frac{\mathbb{C} \times V_\xi\Gamma}{V} \times_{q^{\mathbb{Z}}} \mathbb{R}^{+*},$$

and the leaves of $\mathcal{F}$ are the images of the sets $\mathbb{C} \times \{\hat{v}\} \times \{x\}$ by the natural map $\pi : \mathbb{C} \times V_\xi\Gamma \times \mathbb{R}^{+*} \to S(E_0)$. Observe that the domain of a typical foliation chart is locally isomorphic to $D \times \Omega \times]1, 2[$ where D is an open disk of $\mathbb{C}$, Ω is an open subset of $\mathrm{T}\Gamma$ so that the leaves are given by $D \times \{\omega\} \times \{x\}$ for $(\omega, x) \in \Omega \times]1, 2[$; the term "laminated" refers to the fact that the local transversal to the foliation $\mathcal{F}$ is the disconnected space $\Omega \times]1, 2[$.

Remark 2. Using the fact that V (resp. $q^{\mathbb{Z}}$) acts freely on $V_\xi\Gamma$ (resp. $\mathbb{R}^{+*}$), the reader will check that $(S(E_0), \mathcal{F})$ has trivial holonomy.

One defines a flow ϕ^t acting on $(S(E_0), \mathcal{F})$ and sending each leaf into another leaf by: $\phi^t(z, \hat{v}, x) = (z, \hat{v}, xe^{-t})$. Let μ_ξ denote a Haar measure on the group $V_\xi\Gamma$ then, one has the following

Lemma 3 ([De02]).

1) *The measure*

$$dx_1 dx_2 \otimes \mu_\xi \otimes \frac{dx}{x}$$

on $\mathbb{C} \times V_\xi\Gamma \times \mathbb{R}^{+*}$ *induces a measure* μ *on* $S(E_0)$.

2) *The measure* μ *is invariant under the action of* ϕ^t.

Proof.

1) We just have to check that for any $\nu \in \mathbb{N}^*$ and any borel subset A of $V_\xi\Gamma$, one has
$$\mu_\xi(\xi^\nu A) = |\xi|^{-2\nu}\mu_\xi(A) = q^{-\nu}\mu_\xi(A).$$
Since $(\xi^\nu)_*\mu_\xi$ is also a Haar measure on $V_\xi\Gamma$ it suffices to check this equality for $A = \mathrm{T}\Gamma$. But this is an immediate consequence of the fact that
$$\mathrm{T}\Gamma/(\xi^\nu \mathrm{T}\Gamma) \simeq \Gamma/(\xi^\nu\Gamma)$$
has $|\xi|^{2\nu} = q^\nu$ elements.
2) This is obvious.

Using the fact that $|\xi| = \sqrt{q}$, one checks that the Riemannian metric on the bundle $\mathrm{T}\mathbb{C} \times V_\xi\Gamma \times \mathbb{R}^{+*}$ given by:
$$g_{z,\hat{v},x}(\eta_1,\eta_2) = x^{-1}\mathrm{Re}\,(\eta_1\overline{\eta_2})$$
induces a Riemannian metric g along the leaves of $(S(E_0),\mathcal{F})$ so that the following property is satisfied:

$$\forall \eta \in \mathrm{T}_{[z,\hat{v},x]}\mathcal{F},\ g(D_{[z,\hat{v},x]}\phi^t(\eta), D_{[z,\hat{v},x]}\phi^t(\eta)) = e^t g(\eta,\eta). \tag{8}$$

Compare with Comment 4.

Theorem 2 (Deninger [De02]). *There is a natural bijection between the set of valuations w of the function field $K(E_0)$ of E_0 and the set of primitive compact $\mathbb{R}-$orbits of ϕ^t on $S(E_0)$. It has the following property. If w corresponds to $\gamma = \gamma_w$, then*
$$l(\gamma_w) = \log N(w).$$

Deninger has also provided a nice spectral cohomological interpretation of the left hand side of (7).

Now we are going to recall briefly the definition of the leafwise Hodge cohomology that has allowed us to give in [Lei07] an Atiyah-Bott-Lefschetz proof à la Alvarez-Lopez Kordyukov of the explicit formula (7).

First we introduce carefully a natural transverse measure on $(S(E_0),\mathcal{F})$ and point out its important role. For more on these notions, the reader may have a look at: [Co94], [Sul93] and [Ghys99].

Set $\mathcal{L}_{E_0} = \frac{\mathbb{C}\times\mathrm{T}\Gamma}{\Gamma}$, this is a compact laminated space which is foliated by its path-connected components. Any element $q^\nu \in q^{\mathbb{Z}}$ acts on $[z,\widehat{v}] \in \mathcal{L}_{E_0}$ by $q^\nu \cdot [z,\widehat{v}] = [\xi^\nu z, \xi^\nu\widehat{v}]$. The Haar measure μ_ξ of $\mathrm{T}\Gamma$ induces a transverse measure, still denoted μ_ξ, of $\mathcal{L}_{E_0}$. For any Borel transversal T of $\mathcal{L}_{E_0}$ one has $\mu_\xi(q\cdot T) = q^{-1}\mu_\xi(T)$.

Moreover, the metric $\widetilde{g} = (dx_1)^2 + (dx_2)^2$ (where $z = x_1 + ix_2$) defines a leafwise metric on $\mathcal{L}_{E_0}$, let $\lambda_{\widetilde{g}}$ be the associated leafwise volume form. Then $\lambda_{\widetilde{g}}\mu_\xi$ defines a $q^{\mathbb{Z}}$−invariant measure of $\mathcal{L}_{E_0}$.

The leafwise metric g in (8) of $(S(E_0), \mathcal{F})$ is defined by $g = x^{-1}\widetilde{g}$ and its associated leafwise volume form is given by $\lambda_g = x^{-1}dx_1 \wedge dx_2$.

Definition 1.

1) Let $\mathcal{A}^j_{\mathcal{F}}(S(E_0))$ denote the vector space of leafwise differential forms which, in the local coordinates $(z, \widehat{v}, x)$, are of the form
$$u(z, \widehat{v}, x) d^a \Re z d^b \operatorname{Im} z$$
where $a + b = j \in \{0, 1, 2\}$ and $(z, \widehat{v}, x) \to D^\beta_{\Re z, \operatorname{Im} z, x} u(z, \widehat{v}, x)$ is continuous for any multiindex of differentiation $\beta \in \mathbb{N}^3$.
2) One defines the Sobolev space $H_{+\infty}(S(E_0)\,;\wedge^j \mathrm{T}^*\mathcal{F})$ in the same way but we simply require that the functions $(z, \widehat{v}, x) \to D^\beta_{\Re z, \operatorname{Im} z, x} u(z, \widehat{v}, x)$ are locally L^2.

Now it is clear that $\frac{\mu}{\lambda} = \mu_\xi dx$ defines a transverse measure on $(S(E_0), \mathcal{F})$ with associated Ruelle-Sullivan current $C(\frac{\mu}{\lambda})$. We can pair sections of $\mathcal{A}^2_{\mathcal{F}}(S(E_0))$ with $C(\frac{\mu}{\lambda})$, for instance the measure μ may be recovered by the formula:
$$\forall f \in C^0(S(E_0)),\ \left(f\lambda; C\left(\frac{\mu}{\lambda}\right)\right) = \int_{S(E_0)} f\, d\,\mu.$$
One defines a scalar product by the following formula:
$$\forall \omega, \omega' \in \mathcal{A}^j_{\mathcal{F}}(S(E_0)),\ \langle \omega; \omega' \rangle = \left(\omega \cup *\omega'; C\left(\frac{\mu}{\lambda}\right)\right)$$
where $*$ denotes the leafwise Hodge star operator associated to g.

Theorem 3 ([Lei07]).

1) *One has a Hodge decomposition* (*for* $0 \leq j \leq 2$):
$$H_{+\infty}(S(E_0), \wedge^j \mathrm{T}^*\mathcal{F}) = H^j_\tau \oplus^\perp \overline{\operatorname{Im} \Delta_\tau}.$$
Let π^j_τ *denote the associated projection onto the vector space of leafwise harmonic forms* H^j_τ.
2) *Let* $\alpha \in C^\infty_{\mathrm{compact}}(\mathbb{R}; \mathbb{R})$. *Then* $\int_{\mathbb{R}} \alpha(t)(\phi^t)^* \pi^j_\tau dt$ *is trace class and*
$$\sum_{j=0}^{2} (-1)^j TR \int_{\mathbb{R}} \alpha(t)(\phi^t)^* \pi^j_\tau dt =$$
$$= \sum_\gamma \sum_{k \geq 1} l(\gamma) \left(e^{-kl(\gamma)} \alpha(-kl(\gamma)) + \alpha(kl(\gamma)) \right)$$

where γ *runs over the set of primitive closed orbits of* $(S(E_0), \mathcal{F})$. *According to Theorem* 2, *we obtain in this way an Atiyah-Bott-Lefschetz proof* (*along the lines of* [A-K00]) *of the explicit formula* (7).

Comment 6. Actually, in [Lei07] we have defined in an abstract way (being motivated by the work of Deninger) a class of laminated foliated spaces for which the previous Theorem still holds true.

As noticed by Deninger, the dissymmetry in (7) of the coefficients of $\alpha(kl(\gamma))$ for $k \leq -1$ and $k \geq 1$ is due to property (8) (see the remark following Corollary 1 of [Lie03]). We are going to propose a dynamical explanation, à la Guillemin-Sternberg, of this dissymmetry. Consider a point $(z_0, \widehat{v}_0, 1) \in S(E_0)$, with $\widehat{v}_0 \in \mathrm{T}\Gamma$, such that $\phi^{-\log q}[z_0, \widehat{v}_0, 1] = [z_0, \widehat{v}_0, q] = [z_0, \widehat{v}_0, 1]$. Recall that by definition $(\xi^{-1} z_0, \xi^{-1}\widehat{v}_0, q^{-1}q) \sim (z_0, \widehat{v}_0, q)$. So $[\xi^{-1}z_0, \xi^{-1}\widehat{v}_0, 1] = [z_0, \widehat{v}_0, 1]$ and there exists $\gamma \in \xi^{-1}\Gamma$ such that

$$\xi^{-1} z_0 = z_0 + \gamma, \quad \xi^{-1}\widehat{v}_0 = \widehat{v}_0 - \gamma. \tag{9}$$

The operator $(\phi^t)^*$ acting on $\mathcal{A}^j_{\mathcal{F}}(S(E_0))$ admits a Schwartz kernel defined by the formula:

$$\forall \omega \in \mathcal{A}^j_{\mathcal{F}}(S(E_0)),\ (\phi^t)^*(\omega)(y) = \int_{S(E_0)} (D\phi^t)^* \delta_{\phi^t(y)=y'}\, \omega(y')\, d\mu(y').$$

Consider a point $y = [z, \widehat{v}, x]$ belonging to a small neighborhood of $\{\phi^t[z_0, \widehat{v}_0, 1], -\log q \leq t \leq 0\}$. Then, with the previous notations, one has:

$$\phi^t(y) = (\xi^{-1}z - \gamma, \xi^{-1}\widehat{v} + \gamma, q^{-1}xe^{-t}). \tag{10}$$

The following lemma shows basically that the graph of the flow $(\phi^t)_{t\in\mathbb{R}\setminus\{0\}}$ is transverse to the diagonal and computes $\delta_{\phi^t(y)=y}$.

Lemma 4.

1) *$z \in \mathbb{C} \to \xi^{-1}z - \gamma - z$ and $\widehat{v} \in V_\xi\Gamma \to \xi^{-1}\widehat{v} + \gamma - \widehat{v}$ are invertible and their jacobians are respectively given by*:

$$\mathrm{Jac}(\xi^{-1}z - \gamma - z) = |\xi^{-1} - 1|^2,\ \mathrm{Jac}(\xi^{-1}\widehat{v} + \gamma - \widehat{v}) = q.$$

2) *Let V be an open neighborhood of $(z_0, \widehat{v}_0)$, set*:

$$U = \{(z, \widehat{v}, e^{-s})/\ s \in]-\log q, 0],\ (s, \widehat{v}) \in V\}.$$

Consider $\epsilon > 0$ and V small enough so that $t \in [-\log q, 0] \to (z_0, \widehat{v}_0, e^{-t})$ is the only closed orbit of ϕ^t contained in U with length in $]-\epsilon - \log q, \epsilon - \log q[$. Then one has the following equality as a distribution on $U \times]-\epsilon - \log q, \epsilon - \log q[$:

$$\delta_{\phi^t(y)=y} = \frac{1}{|\xi^{-1} - 1|^2}\delta_{z-z_0} \otimes \frac{1}{q}\delta_{\widehat{v}-\widehat{v}_0} \otimes \delta_{t+\log q}.$$

PROOF.

1) We prove only the second equality. Recall that $\mathrm{T}\Gamma$ is an open compact subset of $V_\xi\Gamma$. Then, since $\widehat{v} \to \widehat{v} - \xi\widehat{v}$ defines an automorphism of $\mathrm{T}\Gamma$ whose inverse is $\widehat{v} \to \sum_{n\in\mathbb{N}} \xi^n \widehat{v}$, one has $\mathrm{Jac}\,(\widehat{v} - \xi\widehat{v}) = 1$. Now recall that the proof of Lemma 3 shows that $\mu_\xi(\xi\mathrm{T}\Gamma) = \frac{1}{q}\mu_\xi(\mathrm{T}\Gamma)$ so that $\mathrm{Jac}\,(\xi\widehat{v}) = \frac{1}{q}$. By combining the last two equalities for Jac, one gets:

$$\mathrm{Jac}(\xi^{-1}\widehat{v} + \gamma - \widehat{v}) = q.$$

2) Using the change of variable formula for $\int d\mu_\xi$ and the equality $\xi^{-1}\widehat{v}_0 + \gamma - \widehat{v}_0 = 0$, one sees that for $\widehat{v}$ close to $\widehat{v}_0$ one has

$$\delta_{\xi^{-1}\widehat{v}+\gamma-\widehat{v}} = \frac{1}{\mathrm{Jac}(\xi^{-1}\widehat{v} + \gamma - \widehat{v})}\delta_{\widehat{v}-\widehat{v}_0}.$$

Then a computation using (9) and (10) shows (see also [Co99, Section IV]), that for $y = [z, \widehat{v}, x] \in U$ and $t \in]-\epsilon - \log q, \epsilon - \log q[$ one has:

$$\delta_{\phi^t(y)=y} = \frac{1}{\mathrm{Jac}(\xi^{-1}z - \gamma - z)}\delta_{z-z_0} \otimes \frac{1}{\mathrm{Jac}(\xi^{-1}\widehat{v} + \gamma - \widehat{v})}\delta_{\widehat{v}-\widehat{v}_0} \otimes \delta_{t+\log q}.$$

By combining 1) with this equality one gets the result.

Recall now that $d\mu(y) = dx_1 dx_2 \otimes \mu_\xi \otimes \frac{dx}{x}$. The formula $|\int_0^{-\log q} \frac{d\,e^{-s}}{e^{-s}}| = \log q$ and Lemma 4. 2) show that for t close to $-\log q$ the distributional trace

$$\int_{S(E_0)} \mathrm{Tr}\,(D\phi^t)^* \delta_{\phi^t(y)=y}\, d\mu(y)$$

is well defined (near $-\log q$) and is equal to:

$$\log q \sum_{\gamma_w,\, l(\gamma_w)=\log q} \frac{1}{q}\delta_{-l(\gamma_w)}$$

where γ_w runs over the set of closed orbits of ϕ^t of length $l(\gamma_w) = \log q$.

Since $\mathrm{Jac}(\xi\widehat{v} + \gamma - \widehat{v}) = 1$ a similar argument shows that for t close to $\log q$ the distributional trace

$$\int_{S(E_0)} \mathrm{Tr}\,(D\phi^t)^* \delta_{\phi^t(y)=y}\, d\mu(y)$$

is well defined (near $\log q$) and equal to:

$$\log q \sum_{\gamma_w,\, l(\gamma_w)=\log q} \delta_{l(\gamma_w)}.$$

Therefore, we have given a dynamical explanation of the dissymmetry occuring in (7).

Now we come to another analogy. In [De07b], Deninger has suggested that, in the case of the Riemann zeta function, the distribution

$$\sum_{\rho\in\widehat{\zeta}^{-1}\{0\},\ \mathrm{Im}\,\rho>0} e^{\rho z}$$

might be interpreted as a trace involving a transversal wave operator. We refer to [De07b, Section 5] for a list of interesting open problems in this direction. In this Section, we simply check that, in the case of ζ_{E_0}, Deninger's intuition is right (see also the end of the next Section). Recall that local coordinates on $S(E_0) = \frac{\mathcal{L}_{E_0}\times\mathbb{R}^{+*}}{q^{\mathbb{Z}}}$ (where $\mathcal{L}_{E_0} = \frac{\mathbb{C}\times\mathrm{T}\Gamma}{\Gamma}$) are given by $(z, \widehat{v}, x)$. We endow $S(E_0)$ with the bundle like metric:

$$\frac{dzd\overline{z}}{x} \oplus \frac{dx^2}{x^2}.$$

We have a notion of transverse exterior derivative d_T:

$$d_T : \Gamma(S(E_0)\,;\wedge^1\mathrm{T}^*\mathcal{F}) \to \Gamma(S(E_0)\,;\wedge^1\mathrm{T}^*\mathcal{F}\otimes\wedge^1\mathrm{T}^*\mathcal{F}^{\perp}).$$

Let δ_T be its adjoint. Set $\Delta_T = \delta_T d_T$, it acts on $\Gamma(S(E_0)\,;\wedge^1\mathrm{T}^*\mathcal{F})$. We write, locally, an element of $\Gamma(S(E_0)\,;\wedge^{1,0}\mathrm{T}^*\mathcal{F})$ as $a(z,\widehat{v},x)dz$.

Lemma 5. *One has*:

$$\Delta_T(adz) = (-\partial_x x)(x\partial_x)(a)\,d\,z.$$

Consider the following transverse operator defined by:

$$\widetilde{\Delta}_T(adz) = \left(-\partial_x x + \frac{1}{2}\right)\left(x\partial_x + \frac{1}{2}\right)(a)\,d\,z.$$

Lemma 6. *Assume for simplicity that the zeta function $\zeta_{E_0}(s)$ of the elliptic curve E_0 does not vanish on $\mathbb{R}$. The following equality holds, between distributions of the variable $t\in\mathbb{R}$:*

$$\mathrm{Tr}\,\pi_\tau^1 e^{it\sqrt{\widetilde{\Delta_T}}} = 2\sum_{z\in\zeta_{E_0}^{-1}(0),\ \mathrm{Im}\,z>0} e^{it\,\mathrm{Im}\,z}.$$

4 – Foliations provide a simple analogue of Lichtenbaum conjecture for zeta functions

4.1 – Lichtenbaum's conjecture

Let K be a number field. Stephen Lichtenbaum has conjectured ([Licht]) the existence of certain Weil étale cohomology groups with and without compact support $H_c^j(K\,;\mathbb{Z})$, $H_c^j(K\,;\mathbb{R})$, and $H^j(K\,;\mathbb{Z})$, $H^j(K\,;\mathbb{R})$ (for $j \in \mathbb{N}$). These groups are additive (abelian) and should be related to the zeta function ζ_K of K as follows.

Conjecture 1 (Lichtenbaum). The groups $H_c^j(K\,;\mathbb{Z})$ are finitely generated and vanish for $j \geq 4$. Giving $\mathbb{R}$ the usual topology one has:

$$H_c^j(K\,;\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{R} = H_c^j(K\,;\mathbb{R}).$$

Moreover, there exists a canonical element $\psi \in H^1(K,\mathbb{R})$ which is functorial with respect to K and such that the following three properties hold.

1) The complex
$$\dots \xrightarrow{D} H_c^j(K,\mathbb{R}) \xrightarrow{D} H_c^{j+1}(K,\mathbb{R}) \to \dots$$
where $Dh = \psi \cup h$ is acyclic. Notice that $D^2 = 0$ because deg $\psi = 1$.

2) One has near $s = 0$,
$$\zeta_K(s) = s^{\left(\sum_{j=0}^{3}(-1)^j j \operatorname{rank} H_c^j(K,\mathbb{Z})\right)}\,\zeta_K^*(s).$$

3)
$$\zeta_K^*(0) = \frac{\prod\limits_{0\leq j\leq 3} |H_c^j(K,\mathbb{Z})_{\text{torsion}}|^{(-1)^j}}{det(H_c^\bullet(K,\mathbb{R}), D, f^\bullet)}$$
where for each $j \in \{0,1,2,3\}$, f^j is a basis of $\frac{H_c^j(K,\mathbb{Z})}{\text{torsion}}$.

We explain the meaning of 3). Since the complex in 1) is acyclic, we have a canonical isomorphism:

$$\otimes_{0\leq j\leq 3}(\det H_c^j(K,\mathbb{R}))^{(-1)^j} \simeq \mathbb{R}.$$

When j is odd we take the dual of this real line. Then each f^j induces a basis $\widehat{f^j}$ of $\det H_c^j(K,\mathbb{R})$ and $\otimes_{0\leq j\leq 3}(\widehat{f^j})^{(-1)^j}$ defines a real number denoted $\det(H_c^\bullet(K,\mathbb{R}), D, f^\bullet)$ which does not depend on the choice of $f^\bullet$.

Lichtenbaum [Licht05] has proven, in the function field case, the analogue of his conjecture.

4.2 – A dynamical foliation analogue

We follow Deninger [De07a]. Recall that in Section 3.2 we have defined the Ruelle zeta function ζ_R of $(X, \mathcal{F}, g, \phi^t)$. We assume now that all the hypothesis, conjectures and notations stated after Comment 4 at the end of Section 3.2. are satisfied by $(X, \mathcal{F}, g, \phi^t)$ and ζ_R. In particular the leafwise metric g is ϕ^t-invariant and the flow ϕ^t is transverse to the foliation $\mathcal{F}$ with no fixed point. Moreover for simplicity we assume that all the $\epsilon_\gamma = 1$.

Comment 7. By a structure theorem we could assume that $(X, \mathcal{F}, \phi^t)$ is of the form: $X = \frac{L \times \mathbb{R}}{\Gamma}$, where Γ is a subgroup of $(\mathbb{R}, +)$, L is a fixed (noncompact) Riemann surface, leaf = image of $L \times \{pt\}$, $\phi^s(l, t) = (l, t + s)$.

Set $Y_\phi = \frac{d\phi^t}{dt} = \partial_t$. Define $\omega \in C^\infty(X, \mathrm{T}^*X)$ by $\omega(Y_\phi) = 1$ and $\omega_{|\mathrm{T}\mathcal{F}} = 0$. Then ω is closed (one has $\omega = dt$ in the previous comment).

Theorem 4 (Deninger [De07a]).

1) *The complex*

$$\ldots \xrightarrow{D} H^j(X, \mathbb{R}) \xrightarrow{D} H^{j+1}(X, \mathbb{R}) \rightarrow \ldots$$

where $Dh = [\omega] \cup h$ *is acyclic and* $H^j(X, \mathbb{R})$ *denotes the standard singular cohomology.*

2) *One has near* $s = 0$,

$$\zeta_R(s) = s^{\left(\sum_{j=0}^{3} (-1)^j j \operatorname{rank} H^j(X, \mathbb{Z})\right)} \zeta_R^*(s).$$

3)

$$\zeta_R^*(0) = \frac{\prod_{0 \le j \le 3} |H^j(X, \mathbb{Z})_{\text{torsion}}|^{(-1)^j}}{\det(H^\bullet(X, \mathbb{R}), D, f^\bullet)}$$

where for each $j \in \{0, 1, 2, 3\}$, f^j *is a basis of* $\frac{H^j(X, \mathbb{Z})}{\text{torsion}}$.

– Analogy with Lichtenbaum conjecture

In this situation, the role of Lichtenbaum's Weil étale cohomology is played by the ordinary singular cohomology with $\mathbb{Z}$ or $\mathbb{R}$−coefficients. Since X is compact, we do not have to worry about compact supports. The Ruelle zeta function ζ_R is expressed in terms of Hodge leafwise cohomology $\ker \Delta_\tau$ (cf. (4)), which is related to $H^\bullet(X, \mathbb{R})$ via the decompositions (11) below. Recall that here ζ_R has no Gamma factors (since ϕ^t has no fixed point) and that the zeroes of ζ_R are located on $\Re s = 0$.

Proof. Sketch.

Define a metric g_X on X by: $g_X = g \oplus^\perp g_0$ on $TX = T\mathcal{F} \oplus \mathbb{R}Y_\phi$ where $g_0(Y_\phi) = 1$.

One has the bigrading:

$$\wedge^n T^*X = \oplus_{p+q=n} \wedge^p T^*\mathcal{F} \otimes \wedge^q (\mathbb{R}Y_\phi)^*.$$

We have $\Delta = \Delta_\tau \oplus -\theta^2$ where θ denotes the infinitesimal generator of $(\phi^s)^*(= e^{s\,\theta})$ acting on $C^\infty(X\,;\,\wedge^*\mathrm{T}^*\mathcal{F})$ and Δ_τ denotes the leafwise Laplacian. Moreover one has

$$\ker\Delta^n = \omega \wedge (\ker\Delta_\tau^{n-1})^{\theta=0} \oplus (\ker\Delta_\tau^n)^{\theta=0}. \tag{11}$$

Then, using techniques from the heat equation proof of the index theorem, Deninger proves that

$$\zeta_R^*(0) = \exp\left(-\sum_{0\le j\le 3}(-1)^j \frac{j}{2}\zeta'_{\Delta^j}(0)\right) = T(X, g_X)^{-1}$$

where $T(X, g_X)$ denotes the Ray-Singer analytic torsion. The metric g_X induces a scalar product on $\ker\Delta^j$ and on $H^j(X,\mathbb{R})$ via the Hodge isomorphism. Consider an orthonormal basis h^j on $H^j(X,\mathbb{R})$, denote by h_j the dual basis on $H_j(X,\mathbb{R})$. Consider also the basis f_j of $\frac{H_j(X,\mathbb{Z})}{\text{torsion}}$ which is dual to f^j $(0 \le j \le 3)$. Now recall that the Reidemeister torsion is defined by:

$$\tau(X) = \prod_{j=0}^{3} |H_j(X,\mathbb{Z})_{\text{torsion}}|^{(-1)^j}.$$

Using the Poincare duality isomorphism $H^j(X,\mathbb{Z}) \simeq H_{3-j}(X,\mathbb{Z})$ one then gets:

$$\tau(X) = \prod_{j=0}^{3} |H^j(X,\mathbb{Z})_{\text{torsion}}|^{(-1)^{j+1}}.$$

By the Cheeger-Mueller theorem one has:

$$T(X, g_X) = \tau(X)\prod_{j=0}^{3} |\det_{f_j} h_j|^{(-1)^j}.$$

Now, using the decompositions (11) Deninger shows that

$$1 = |\det(H^\bullet(X,\mathbb{R}), D, h^\bullet)|.$$

By construction, one has:

$$|\det(H^\bullet(X,\mathbb{R}), D, h^\bullet)| = |\det(H^\bullet(X,\mathbb{R}), D, f^\bullet)|\prod_{j=0}^{3} |\det_{f_j} h_j|^{(-1)^{j+1}}.$$

By combining the last six identities one then gets the theorem.

In the proof of Theorem 4, we have seen that the operator $|\theta|$ appears as a transverse square root Laplacian. Assume that the Ruelle zeta function does not vanish at 0. Then one checks that the following holds as distributions of the variable $t \in \mathbb{R}$.

$$2 \sum_{\rho \in \zeta_R^{-1}(0),\, \mathrm{Im}\,\rho > 0} e^{t\rho} = \mathrm{Tr}\, \pi_\tau^1 e^{it|\theta|}.$$

See the end of Section 3.3.

5 – Remarks about a conjectural dynamical foliated space $(S_{\mathbb{Q}}, \mathcal{F}, g, \phi^t)$ associated to the Riemann zeta function

The following Section is speculative in nature. It should be viewed as a working programme or a motivation for developing interesting mathematics.

5.1 – Structural Assumptions and their consequences

We assume, following Deninger (e.g. [De01b], [De01]), that to $\mathrm{Spec}\,\mathbb{Z} \cup \{\infty\}$, one can associate a Riemannian (laminated) foliated space $(S_{\mathbb{Q}}, \mathcal{F}, g, \phi^t)$ satisfying the following assumptions.

1. The leaves are Riemann surfaces and the path connected components of $S_{\mathbb{Q}}$ are three dimensional. Moreover, g denotes a leafwise riemannian metric, $(\phi^t)_{t\in\mathbb{R}}$ is a flow acting on $(S_{\mathbb{Q}}, \mathcal{F})$ and permuting the leaves.

2. To each prime $p \in \mathcal{P}$ there corresponds a unique primitive closed orbit γ_p of ϕ^t of length $\log p$. To the archimedean absolute value of $\mathbb{Q}$ there corresponds a unique fixed point $x_\infty = \phi^t(x_\infty), \forall t \in \mathbb{R}$, of the flow. The flow is transverse to all the leaves different from the one containing x_∞.

3. We assume that:

$$\forall t \in \mathbb{R},\ e^{-t/2} D_y \phi^t(x_\infty)_{|T_{x_\infty}\mathcal{F}} \in SO_2(T_{x_\infty}\mathcal{F}). \tag{12}$$

4. We have reduced real leafwise cohomology groups $\overline{H}^j_{\mathcal{F}}$ $(0 \le j \le 2)$ on which $(\phi^t)_{t\in\mathbb{R}}$ acts naturally such that $\overline{H}^0_{\mathcal{F}} \simeq \mathbb{R}$, $\overline{H}^2_{\mathcal{F}} \simeq \mathbb{R}$ and $\overline{H}^1_{\mathcal{F}}$ is infinite dimensional. Let $[\lambda_g]$ denote the class in $\overline{H}^2_{\mathcal{F}}$ of the leafwise kaehler metric λ_g associated to g. Then we assume that

$$\forall t \in \mathbb{R},\ (\phi^t)^*([\lambda_g]) = e^t [\lambda_g]. \tag{13}$$

5. The action of ϕ^t on $\overline{H}^1_{\mathcal{F}}$ commutes with the Hodge star $\star$ induced by g. Moreover there exists a transverse measure μ on $(S_{\mathbb{Q}}, \mathcal{F})$ such that $\int_{S_{\mathbb{Q}}} (\alpha \wedge \star\beta)\mu$ defines a scalar product on $\overline{H}^1_{\mathcal{F}}$.

6. For any $\alpha \in C^\infty_{\text{compact}}(\mathbb{R};\mathbb{R})$, $\int_{\mathbb{R}} \alpha(t)(\phi^t)^* dt$ acting on $\overline{H}^1_{\mathcal{F}}$ is trace class. The explicit formula (1) is interpreted as a Lefschetz trace formula for the riemannian foliated space $(S_{\mathbb{Q}}, \mathcal{F}, g, \phi^t)$ with respect to the leafwise cohomology groups $\overline{H}^j_{\mathcal{F}}$ $(0 \le j \le 2)$.

7. The fixed point $x_\infty \in S_{\mathbb{Q}}$ should be a limit point of a trajectory γ_∞ : $\lim_{t\to+\infty} \phi^t(y) = x_\infty$ for any $y \in \gamma_\infty$. Moreover, γ_∞ should have the following orbifold structure. Define an orbifold structure on $\mathbb{R}^{\ge 0}$ by requiring the following map to be an orbifold isomorphism:

$$Sq : \frac{\mathbb{R}}{\{1,-1\}} \to \mathbb{R}^{\ge 0},\ Sq(z) = z^2.$$

Notice that Sq transforms the flow $\phi^t_{\frac{\mathbb{R}}{\{1,-1\}}}(z) = ze^{-t}$ into the flow $\phi^t_{\mathbb{R}^{\ge 0}}(v) = ve^{-2t}$. Then we require that there exists an embedding $\Psi : \mathbb{R}^{\ge 0} \to \gamma_\infty$ such that $\Psi(0) = x_\infty$ and

$$\forall (t,v) \in \mathbb{R} \times \mathbb{R}^{\ge 0},\ \Psi(\phi^t_{\mathbb{R}^{\ge 0}}(v) = ve^{-2t}) = \phi^t(\Psi(v)). \tag{14}$$

Lastly we require that γ_∞ is transverse at x_∞ to $T_{x_\infty}\mathcal{F}$.

Comment 8. The stronger assumption $\forall t \in \mathbb{R}$, $(\phi^t)^*(g) = e^t g$ implies (12) (because $\phi^0 = \mathrm{Id}$), (13) and the fact that ϕ^t commutes with the Hodge star not only on $\overline{H}^1_{\mathcal{F}}$ but also on the vector space of leafwise differential $1-$forms. Deninger told us privately that this assumption $(\phi^t)^*(g) = e^t g$ might be too strong. Assumption 5 and (13) implies equation (3) in Deninger's formalism. Therefore, the first six Assumptions imply the Riemann hypothesis as explained in Section 2. Assumption 7 is stated here as a hint about a possible way to prove Assumption 6. See the next subsection.

Comment 9. The results that we have described in Section 3.3 (e.g. Lemma 4) suggest that the disymmetry mentionned in Comment 1 might be explained in the following way. For each prime $p \in \mathcal{P}$, $(S_{\mathbb{Q}}, \mathcal{F})$ should exhibit a transversal of the type $]0,1[\times \mathbb{Z}_p$ and possibly the ring of finite Adeles might enter into the picture.

5.2 – Remarks about the contribution of the archimedean place in (1)

Now we apply formally the Guillemin-Sternberg trace formula for

$$\sum_{j=0}^{2} (-1)^j \operatorname{Tr}((\phi^t)^* \,;\, \Gamma(S_{\mathbb{Q}} \,;\, \wedge^j T^*\mathcal{F}))$$

where $\Gamma(S_{\mathbb{Q}}; \wedge^j T^*\mathcal{F})$ denotes the set of "smooth" sections.

Lemma 7 (Deninger [De01]).

1) *The contribution of the fixed point* x_∞ *in the previous Guillemin-Sternberg trace formula is*:

$$\frac{1}{|\det(1 - D_y\phi^t(x_\infty)\,;\,T_{x_\infty}S_{\mathbb{Q}}/T_{x_\infty}\mathcal{F})|}.$$

2)

$$\forall t \in \mathbb{R}\setminus\{0\},\ \frac{1}{|\det(1 - D_y\phi^t(x_\infty)\,;\,T_{x_\infty}S_{\mathbb{Q}}/T_{x_\infty}\mathcal{F})|} = \frac{1}{|1-e^{-2t}|}.$$

Proof.

1) Using Proposition 1, one sees that the contribution of the fixed point x_∞ is equal to:

$$\frac{\displaystyle\sum_{j=0}^{2}(-1)^j \operatorname{Tr}((D_y\phi^t)^*(x_\infty);\wedge^j T^*_{x_\infty}\mathcal{F})}{|\det(1 - D_y\phi^t(x_\infty);T_{x_\infty}S_{\mathbb{Q}})|} =$$

$$= \frac{\det(1 - D_y\phi^t(x_\infty);T_{x_\infty}\mathcal{F})}{|\det(1 - D_y\phi^t(x_\infty);T_{x_\infty}\mathcal{F})|}\frac{1}{|\det(1 - D_y\phi^t(x_\infty);T_{x_\infty}S_{\mathbb{Q}}/T_{x_\infty}\mathcal{F})|}.$$

Using property (12) one checks easily that

$$\frac{\det(1 - D_y\phi^t(x_\infty);T_{x_\infty}\mathcal{F})}{|\det(1 - D_y\phi^t(x_\infty)\,;\,T_{x_\infty}\mathcal{F})|} = 1.$$

One then gets immediately 1).

2) Since $T_{x_\infty}S_{\mathbb{Q}}/T_{x_\infty}\mathcal{F}$ is a real line, there exists $\kappa \in \mathbb{R}$ such that:

$$\forall t \in \mathbb{R}, |\det(1 - D_y\phi^t(x_\infty)\,;\,T_{x_\infty}S_{\mathbb{Q}}/T_{x_\infty}\mathcal{F})| = |1-e^{\kappa t}|.$$

By Assumption 7, γ_∞ is transverse at x_∞ to $T_{x_\infty}\mathcal{F}$ and (14) shows that $D_y\phi^t(x_\infty)$ acts as e^{-2t} on the real line $T_{x_\infty}S_{\mathbb{Q}}/T_{x_\infty}\mathcal{F}$. One then gets 2) immediately.

Recall that we wish to interpret (1) as a Lefschetz trace formula via the Guillemin-Sternberg formula.

Proposition 2 (Deninger [De01]).

1) *The contribution of the real archimedean absolute value in* (1) *coincides for any real t positive with the contribution of the fixed point x_∞ in the Guillemin-Sternberg formula.*
2) *The contributions of the fixed point x_∞ for t real negative in the Guillemin-Sternberg formula and of the real archimedean absolute value in* (1) *do not coincide.*
3)

$$\forall t < 0,\ \frac{e^t}{1-e^{2t}} = \frac{\sum_{j=0}^{2}(-1)^j Tr((D_y\phi^t)^*(x_\infty)\,;\,\wedge^j T^*_{x_\infty}\mathcal{F})}{|\det(1-D_y\phi^{|t|}(x_\infty)\,;\,T_{x_\infty}S_{\mathbb{Q}})|} \tag{15}$$

where in the denominator we have written $\phi^{|t|}$.

Proof.

1) This is part 2) of the previous lemma.
2) Indeed, the Guillemin-Sternberg formula gives

$$\frac{1}{|1-e^{-2t}|} = \frac{e^{2t}}{1-e^{2t}},$$

whereas (1) gives $\frac{e^t}{1-e^{2t}}$ for $t < 0$.
3) This follows from (12) and a simple computation.

Comment 10. It was in order to explain the factor -2 (instead of -1) in $\frac{1}{1-e^{-2t}}$ for $t > 0$ in (1) that Deninger has proposed in [De01b, Section 3] the Assumption 7.

– **Open Question**

Find a conceptual explanation of the equation (15) by a suitable generalization of Guillemin-Sternberg's trace formula to a "suitable singular setting".

5.3 – Remarks about the contribution of the archimedean place in the explicit formula for $\zeta_{\mathbb{Q}[i]}$

We recall the explicit formula for the zeta function $\zeta_{\mathbb{Q}[i]}$ of $\mathbb{Q}[i]$ as an equality between two distributions in $\mathcal{D}'(\mathbb{R}\setminus\{0\})$ (t being the real variable).

$$\begin{aligned}1-\sum_{\rho\in\zeta^{-1}_{\mathbb{Q}[i]}\{0\},\Re\rho\geq 0} e^{t\rho}+e^t=&\sum_{\mathcal{Q}}\log N\mathcal{Q}\sum_{k\geq 1}(\delta_{k\log N\mathcal{Q}}+(N\mathcal{Q})^{-k}\delta_{-k\log N\mathcal{Q}})+\\ &+\frac{1}{1-e^{-t}}\,1_{\{t>0\}}+\frac{e^t}{1-e^t}\,1_{\{t<0\}},\end{aligned} \tag{16}$$

where $\mathcal{Q}$ runs over the set of non zero prime ideals of $\mathbb{Z}[i]$ and $N\mathcal{Q}$ denotes the norm of $\mathcal{Q}$.

Of course, one conjectures the existence of a Riemannian (laminated) foliated space $(S_{\mathbb{Q}[i]}, \mathcal{F}, g, \phi^t)$ satisfying a list of axioms quite similar to the ones stated in Section 5.1. We simply explain how Assumption 7 has to be modified for the pair of the two complex archimedean places $\{|\cdot|_{\mathbb{C}}, |\bar{\cdot}|_{\mathbb{C}}\}$ of $\zeta_{\mathbb{Q}[i]}$.

Substitute of 7 for $\mathbb{Q}[i]$ (cf. [De01b, Section 3]). There exists a stationary fixed point $z_\infty \in S_{\mathbb{Q}[i]}$ of ϕ^t and two trajectories $\gamma_\pm$ of the flow ϕ^t with end point z_∞. For any $z_\pm \in \gamma_\pm$, $\lim_{t\to+\infty} \phi^t(z_\pm) = z_\infty$. These two trajectories $\gamma_\pm$ are transverse to $\mathcal{F}$ at z_∞. Moreover there exists an embedding:

$$\Psi : \mathbb{R} \to \gamma_- \cup \gamma_+,$$

such that $\Psi(0) = z_\infty$, $\gamma_\pm \setminus \{0\} = \Psi(\mathbb{R}^\pm \setminus \{0\})$. Lastly, $\forall v, t \in \mathbb{R}$, $\Psi(ve^{-t}) = \phi^t(\Psi(v))$.

Therefore, the contribution of z_∞ in the Guillemin-Sternberg trace formula is:

$$\forall t \in \mathbb{R} \setminus \{0\},\ \frac{1}{|1-e^{-t}|} = \frac{1}{|\det(1 - D_y\phi^t(z_\infty)\,;\, T_{z_\infty}S_{\mathbb{Q}[i]}/T_{z_\infty}\mathcal{F})|}.$$

Part 1) of the following proposition shows that the contribution of a *complex* archimedean place in the explicit formula is better understood than the one of a *real* archimedean place (cf. Proposition 2. 2)). Part 2) suggests that the previous open question should admit a conceptual answer.

Proposition 3 ([De01, Section 5]).

1) *The contribution of z_∞ in the Guillemin-Sternberg trace formula coincides, for $t \in \mathbb{R} \setminus \{0\}$, with the contribution of the two complex archimedean places in* (16).
2) *For t real negative, one has*:

$$\frac{e^t}{1-e^t} = \frac{\displaystyle\sum_{j=0}^{2} (-1)^j Tr((D_y\phi^t)^*(z_\infty); \wedge^j T^*_{z_\infty}\mathcal{F})}{|\det(1 - D_y\phi^{|t|}(z_\infty)\,;\, T_{z_\infty}S_{\mathbb{Q}[i]})|}.$$

Acknowledgements

I am happy to thank Jesus Alvarez-Lopez and Christopher Deninger for helpful explanations and also Paolo Piazza and Yuri Kordyukov for comments on the first version.

REFERENCES

[A-K00] J. A. Alvarez Lopez – Y. Kordyukov: *Distributional Betti numbers of transitive foliations of codimension one*, Foliations: geometry and dynamics, (Warsaw, 2000), World Sci. Publishing, River Edge, NJ, 2002, 159-183.

[A-K01] J. A. Alvarez Lopez – Y. Kordyukov: *Long time behaviour of leafwise heat flow for Riemannian foliations*, Compositio Math., **125** (2001), 129-153.

[A-K05] J. A. Alvarez Lopez – Y. Kordyukov: *Distributional Betti numbers for Lie foliations*, To appear in the proceedings of Bedlewo conference 2005 on C^*-algebras.

[Co94] A. Connes: *Noncommutative Geometry*, Academic Press.

[Co99] A. Connes: *Trace formula in noncommutative geometry and the zeroes of the Riemann zeta function*, Selecta Math. (N.S.), **5** (1999), 29-106.

[De94] C. Deninger: *Motivic $L-$functions and regularized determinants*, Proc. Symp. Pure Math., **55** (1994), 707-743.

[De98] C. Deninger: *Some analogies between number theory and dynamical systems onf foliated spaces*, Doc. Math. J. DMV. Extra volume **ICM I** (1998), 23-46.

[De99] C. Deninger: *On dynamical systems and their possible significance for Arithmetic Geometry*, in: *Regulators in Analysis, Geometry and Number Theory*, A. Reznikov – N. Schappacher (eds.), Progress in Mathematics, **171** Birkhauser, 1999, 29-87.

[De-Si00] C. Deninger – W. Singhof: *A note on dynamical trace formulas*, Dynamical, spectral and arithmetic zeta functions (San Antonio, TX, 1999), Contemp. Math., Amer. Math. Soc., Providence, RI, **290** (2001) 41-55.

[De-Si02] C. Deninger – W. Singhof: *Real polarizable Hodge structures arising from foliations*, Annals of Global Analysis and Geometry. **21** (2002), 377-399.

[De01] C. Deninger: *Number theory and dynamical systems on foliated spaces*, Jahresberichte der DMV., **103** (2001), 79-100.

[De01b] C. Deninger: *A note on arithmetic topology and dynamical systems*, Algebraic number theory and algebraic geometry, Contemp. Math., Amer. Math. Soc., Providence, RI, **300** (2002), 99-114.

[De02] C. Deninger: *On the nature of explicit formulas in analytic number theory, a simple example*, Number theoretic methods (Iizuka, 2001), Dev. Math., Kluwer Acad. Publ., Dordrecht, **8** (2002), 97-118.

[De07a] C. Deninger: *A dynamical system analogue of Lichtenbaum's conjectures on special values of Hasse-Weil zeta functions*, Preprint 2007.

[De07b] C. Deninger: *Analogies between analysis on foliated spaces and arithmetic geometry*, to appear in the Proceedings of a conference in the honour of Hermann Weyl.

[Ghys99] E. Ghys: *Laminations par surfaces de Riemann*, Dynamique et géométrie complexes (Lyon, 1997), Panor. Synthèses, Soc. Math. France, Paris, **8** (1999), 49-95.

[G-S77] V. Guillemin – S. Sternberg: *Geometric Asymptotics*, Mathematical surveys and monographs, published by the A.M.S, Providence, R.I. n. 14, 1977.

[Lei03] E. Leichtnam: *An invitation to Deninger's work on arithmetic zeta functions*, Geometry, spectral theory, groups, and dynamics, Contemp. Math., Amer. Math. Soc., Providence, RI, **387** (2005), 201-236.

[Lei07] E. Leichtnam: *Scaling group flow and Lefschetz trace formula for laminated spaces with a $p-$adic transversal*, Bulletin des Sciences Mathematiques, **131** (2007), 638-669.

[Licht05] S. Lichtenbaum: *The Weil étale topology on schemes over finite fields*, Compositio Math. Journal, **141** (2005), 689-702.

[Licht] S. Lichtenbaum: *The Weil étale topology for number rings*, Preprint 2005.

[Meyer03] R. Meyer: *A spectral interpretation for the zeros of the Riemann zeta function*, Mathematisches Institut, Georg-August-Universitaet Goettingen: Seminars Winter Term 2004/2005, Universitaetsdrucke Goettingen, Goettingen, 2005, 117-137.

[Oor73] F. Oort: *Lifting an endomorphism of an elliptic curve to characteristic zero*, Indag. Math., **35** (1973), 466-470.

[Se60] J-P. Serre: *Analogues Kaehleriens de certaines conjectures de Weil*, Annals of Math., **65** (1960), 392-394.

[Si92] J. Silverman: *The arithmetic of elliptic curves*, Graduate text in Math., **106** (1992).

[Sul93] D. Sullivan: *Linking the universalities of Milnor-Thurston, Feigenbaum and Ahlfors-Bers*, Topological methods in modern mathematics (Stony Brook, NY, 1991), Publish or Perish, Houston, TX, 1993, 543-564.

Lavoro pervenuto alla redazione il 25 gennaio 2008
ed accettato per la pubblicazione il 8 marzo 2008.
Bozze licenziate il 30 settembre 2008

Indirizzo dell'Autore:

Eric Leichtnam – Institut de Jussieu et CNRS – Etage 7E – 175 rue du Chevaleret – 75013 Paris – France
E-mail: ericleichtnam@math.jussieu.fr

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 189–200

Uniqueness of renormalized solutions for a class of parabolic equations with unbounded nonlinearities

HICHAM REDWANE

ABSTRACT: *We prove uniqueness and a comparison principle of renormalized solutions for a class of doubly nonlinear parabolic equations* $\frac{\partial b(x,u)}{\partial t} - \mathrm{div}(A(t,x)Du + \Phi(u)) = f$, *where the right side belongs to* $L^1((0,T)\times\Omega)$ *and where* $b(x,u)$ *is unbounded function of* u *and where* $A(t,x)$ *is a bounded symmetric and coercive matrix, and* Φ *is continuous function but without any growth assumption on* u.

1 – Introduction

In the present paper we establish the uniqueness and comparison principle for a renormalized solutions for a class of doubly nonlinear parabolic equations of the type

$$\frac{\partial b(x,u)}{\partial t} - \mathrm{div}\Big(A(t,x)Du + \Phi(u)\Big) = f \quad \text{in } \Omega\times(0,T), \tag{1.1}$$

$$b(x,u)(t=0) = b(x,u_0) \quad \text{in } \Omega, \tag{1.2}$$

$$u = 0 \quad \text{on } \partial\Omega\times(0,T). \tag{1.3}$$

In Problem (1.1)-(1.3) the framework is the following: Ω is a bounded domain of $\mathbb{R}^N$, $(N \geq 1)$, T is a positive real number while the data f and $b(x,u_0)$ in $L^1(\Omega\times(0,T))$ and $L^1(\Omega)$. And where b is a Carathéodory function such that,

---

KEY WORDS AND PHRASES: *Nonlinear parabolic equations – Uniqueness – Renormalized solutions.*
A.M.S. CLASSIFICATION: 47A15 – 46A32 – 47D20

$b(x,s)$ is unbounded function of s. The matrix $A(t,x)$ is a bounded symmetric and coercive matrix. The function Φ is just assumed to be continuous on $\mathbb{R}$.

When Problem (1.1)-(1.3) is investigated one of the difficulties is due to the facts that the data f and $b(x,u_0)$ only belong to L^1 and the growths of $b(x,u)$ and $\Phi(u)$ are not controlled with respect to u (the function $\Phi(u)$ does not belong to $(L^1_{\mathrm{loc}}((0,T)\times\Omega))^N$ in general), so that proving existence of a weak solution (i.e. in the distribution meaning) seems to be an arduous task. To overcome this difficulty we use the framework of renormalized solutions. The existence of a renormalized solutions of (1.1)-(1.3) is proved in H. Redwane [15].

The notion of renormalized soluion is introduced by Lions and Di Perna [14] for the study of Boltzmann equation (see also P.-L. Lions [10] for a few applications to fluid mechanics models). This notion was then adapted to elliptic version of (1.1)-(1.3) in Boccardo, J.-L. Diaz, D. Giachetti, F. Murat [8], in P.-L. Lions and F. Murat [11] and F. Murat [12], [13] (see also [2], [3], [4], [5], [6], [7]). At the same the equivalent notion of entropy solutions have been developed independently by Bénilan and al. [1] for the study of nonlinear elliptic problems.

The paper is organized as follows: Section 2 is devoted to specify the assumptions on b, Φ, f and u_0 needed in the present study and gives the definition and the existence (Theorem 2.0.3) of a renormalized solution of (1.1)-(1.3). In Section 3 we establish uniqueness and a comparison principle of such a solution (Theorem 3.0.4)

2 – Assumptions on the data and definition of a renormalized solution

Throughout the paper, we assume that the following assumptions hold true: Ω is a bounded open set on $\mathbb{R}^N$ $(N \geq 1)$, $T > 0$ is given and we set $Q = \Omega\times(0,T)$.

$$b : \Omega \times \mathbb{R} \to \mathbb{R} \text{ is a Carathéodory function such that;} \tag{2.1}$$

for every $x \in \Omega$: $b(x,s)$ is a strictly increasing C^1-function, with $b(x,0) = 0$. For any $K > 0$, there exists $\lambda_K > 0$, a function A_K in $L^1(\Omega)$ and a function B_K in $L^2(\Omega)$ such that

$$\lambda_K \leq \frac{\partial b(x,s)}{\partial s} \leq A_K(x) \quad \text{and} \quad \left|\nabla_x\left(\frac{\partial b(x,s)}{\partial s}\right)\right| \leq B_K(x) \tag{2.2}$$

for almost every $x \in \Omega$, for every s such that $|s| \leq K$.

$$A(t,x) \text{ is a symmetric coercive matrix field with coefficients} \tag{2.3}$$

lying in $L^\infty(Q)$ i.e. $A(t,x) = (a_{ij}(t,x))_{1\le i,j\le N}$ with:

- $a_{ij}(t,x) \in L^\infty(Q)$ and $a_{ij}(t,x) = a_{ji}(t,x)$ a.e. in Q, $\forall\ i,\ j$
- $\exists\ \alpha > 0$ such that a.e. in Q, $\forall \xi \in \mathbb{R}^N \quad A(t,x)\xi\xi \ge \alpha \|\xi\|^2_{\mathbb{R}^N}$

(2.4) $\Phi\ :\ \mathbb{R} \to \mathbb{R}^N$ is a continuous function

(2.5) f is an element of $L^1(Q)$.

(2.6) u_0 is a measurable function defined on Ω such that $b(x,u_0) \in L^1(\Omega)$.

Remark 2.0.1. In (2.2), we denote by $\nabla_x\left(\frac{\partial b(x,s)}{\partial s}\right)$ the gradient of $\frac{\partial b(x,s)}{\partial s}$ defined in the sense of distributions.

As already mentioned in the introduction Problem (1.1), (1.2), (1.3) does not admit a weak solution under assumptions (2.1)-(2.6), since the growths of $b(x,u)$ and $\Phi(u)$ are not controlled with respect to u (so that these fields are not in general defined as distributions, even when u belongs $L^2(0,T;W_0^{1,2}(\Omega))$).

Throughout this paper and for any non negative real number K we denote by $T_K(r) = \min(K, \max(r,-K))$ the truncation function at height K. The definition of a renormalized solution for Problem (1.1), (1.2), (1.3) can be stated as follows.

Definition 2.0.2. A measurable function u defined on Q is a renormalized solution of Problem (1.1), (1.2), (1.3) if

(2.7) $T_K(u) \in L^2(0,T;W_0^{1,2}(\Omega))$ for any $K \ge 0$ and $b(x,u) \in L^\infty(0,T;L^1(\Omega))$;

$$\int_{\{(t,x)\in Q;\ n\le|u(x,t)|\le n+1\}} A(x,t)DuDu\,dx\,dt \longrightarrow 0 \quad \text{as } n\to+\infty; \tag{2.8}$$

and if, for every increasing function S in $W^{2,\infty}(\mathbb{R})$, which is piecewise C^1 and such that S' has a compact support, we have

$$\begin{aligned}\frac{\partial b_S(x,u)}{\partial t} &- \operatorname{div}(S'(u)A(t,x)Du) + S''(u)A(t,x)DuDu \\ &- \operatorname{div}(S'(u)\Phi(u)) + S''(u)\Phi(u)Du = fS'(u) \quad \text{in } D'(Q);\end{aligned} \tag{2.9}$$

$$b_S(x,u)(t=0) = b_S(x,u_0) \quad \text{in } \Omega; \tag{2.10}$$

where $b_S(x,r) = \int_0^r \frac{\partial b(x,s)}{\partial s} S'(s)\,ds$.

The existence theorem of renormalized solution of (1.1)-(1.3):

Theorem 2.0.3. *Under assumptions* (2.1)-(2.6) *there exists at least a renormalized solution u of Problem* (1.1)-(1.3).

Proof of Theorem 3.0.3. The existence theorem of renormalized solution of (1.1)-(1.3) is proved in H. Redwane [15]

3 – Comparison principle and uniqueness result

This section is concerned with a comparison principle (and an uniqueness result) for renormalized solutions. We establish the following theorem.

Theorem 3.0.4. *Assume that assumptions* (2.1), (2.2), (2.3), (2.4), (2.5) *and* (2.6) *hold true and moreover that.*

For any $K > 0$, there exists a positive real number $\beta_K > 0$, such that

$$\left|\frac{\partial b(x,z_1)}{\partial s} - \frac{\partial b(x,z_2)}{\partial s}\right| \leq \beta_K \left|z_1 - z_2\right| \tag{3.1}$$

for almost every x in Ω, and for every z_1 and every z_2 such that $|z_1| \leq K$ and $|z_2| \leq K$.

$$\Phi \text{ is a locally Lipschitz continuous function on } \mathbb{R}. \tag{3.2}$$

Let then u_1 and u_2 be renormalized solutions corresponding to the data (f_1, u_0^1) and (f_2, u_0^2). If these data satisfying $f_1 \leq f_2$ and $u_0^1 \leq u_0^2$ almost every where, we have

$$u_1 \leq u_2 \text{ almost every where.}$$

Proof of Theorem 3.0.4. The proof is divided into two steps. In Step 1, we define a smooth approximation S_n of T_n, and we consider two renormalized solutions u_1 and u_2 of (1.1), (1.2), (1.3) for the data (f_1, u_0^1) and (f_2, u_0^2) respectively. We plug the test function $\frac{1}{\sigma}{T_\sigma}^+\Big(b_{S_n}(x,u_1)-b_{S_n}(x,u_2))\Big)$ in the difference of equations (2.9) for u_1 and u_2 in which we have taken $S = S_n$.

In Step 2, we investigate the behaviour of the different terms in the estimate obtained in Step 1 (estimates (3.5)) as σ tends to 0 and when n tends to $+\infty$.

Step 1. Remark that when Φ is locally Lipschitz continuous on $\mathbb{R}$ the following derivation is licit for any function S and u satisfying the conditions mentioned in Definition 2.0.2.

$$\operatorname{div}\Big(S'(u)\Phi(u)\Big) - S''(u)\Phi(u)Du = S'(u)\Phi'(u)Du = \operatorname{div}(\Phi_S(u)). \tag{3.3}$$

Where $\Phi_S = (\Phi_{S,1}, \Phi_{S,2}, \cdots, \Phi_{S,N})$ with

$$\Phi_{S,i}(r) = \int_0^r \Phi'_{S,i}(t)S'(t)\,dt.$$

Let us now introduce a specific choice of function S in (2.9). For all $n > 0$, let $S_n \in C^1(\mathbb{R})$ be the function defined by $S_n(0) = 0$; $S'_n(r) = 1$ for $|r| \leq n$; $S'_n(r) = n + 1 - |r|$ for $n \leq |r| \leq n + 1$ and $S'_n(r) = 0$ for $|r| \geq n + 1$.

It yields, taking $S = S_n$ in (2.9)

$$\begin{aligned}\frac{\partial b_{S_n}(x,u_i)}{\partial t} - \operatorname{div}\Big(S'(u_i)A(t,x)Du_i\Big) + S''(u_i)A(t,x)Du_iDu_i + \\ - \operatorname{div}\Big(\Phi_{S_n}(u_i)\Big) = f_iS'_n(u_i) \ \text{ in } D'(Q);\end{aligned} \tag{3.4}$$

for $i = 1, 2$ and where $b_{S_n}(x,r) = \int_0^r \frac{\partial b(x,s)}{\partial s} S'_n(s)\, ds$.

We use $\frac{1}{\sigma}T^+_\sigma(b_{S_n}(x,u_1) - b_{S_n}(x,u_2))$ as a test function in the difference of equations (3.4) for u_1 and u_2.

$$\begin{aligned}&\frac{1}{\sigma}\int_0^T\int_0^t\Big\langle\frac{\partial(b_{S_n}(x,u_1)-b_{S_n}(x,u_2))}{\partial t}; T^+_\sigma(b_{S_n}(x,u_1)-b_{S_n}(x,u_2))\Big\rangle ds\, dt + A^\sigma_n = \\ &= B^\sigma_n + C^\sigma_n + D^\sigma_n,\end{aligned} \tag{3.5}$$

for any $\sigma > 0$, $n > 0$, and where

$$\begin{aligned}A^\sigma_n = \frac{1}{\sigma}\int_0^T\int_0^t\int_\Omega \Big[S'_n(u_1)A(t,x)Du_1 - S'_n(u_2)A(t,x)Du_2\Big]\cdot \\ \cdot DT^+_\sigma\Big(b_{S_n}(x,u_1) - b_{S_n}(x,u_2)\Big)\, dx\, ds\, dt\end{aligned} \tag{3.6}$$

$$\begin{aligned}B^\sigma_n = \frac{1}{\sigma}\int_0^T\int_0^t\int_\Omega S''_n(u_1)A(t,x)Du_1Du_1T^+_\sigma(b_{S_n}(x,u_1)-b_{S_n}(x,u_2)\, dx\, ds\, dt + \\ -\frac{1}{\sigma}\int_0^T\int_0^t\int_\Omega S''_n(u_2)A(t,x)Du_2Du_2T^+_\sigma(b_{S_n}(x,u_1)-b_{S_n}(x,u_2))\, dx\, ds\, dt\end{aligned} \tag{3.7}$$

$$C^\sigma_n = \frac{1}{\sigma}\int_0^T\int_0^t\int_\Omega[\Phi_{S_n}(u_1) - \Phi_{S_n}(u_2)]DT^+_\sigma(b_{S_n}(x,u_1) - b_{S_n}(x,u_2))\, dx\, ds\, dt \tag{3.8}$$

$$D^\sigma_n = \frac{1}{\sigma}\int_0^T\int_0^t\int_\Omega[f_1S'_n(u_1) - f_2S'_n(u_2)]T^+_\sigma(b_{S_n}(x,u_1) - b_{S_n}(x,u_2))\, dx\, ds\, dt. \tag{3.9}$$

In the sequel we pass to the limit in (3.5) when σ tends to 0 and then n tends to $+\infty$. Upon application of Lemma 2.4 of [9], the first term in the right hand side of (3.5) is derived as

$$\begin{aligned}&\frac{1}{\sigma}\int_0^T\int_0^t\Big\langle\frac{\partial(b_{S_n}(x,u_1)-b_{S_n}(x,u_2))}{\partial t};T_\sigma^+(b_{S_n}(x,u_1)-b_{S_n}(x,u_2))\Big\rangle ds\,dt=\\ &=\frac{1}{\sigma}\int_Q\tilde{T}_\sigma^+(b_{S_n}(x,u_1)-b_{S_n}(x,u_2))\,dx\,dt+\\ &\quad-\frac{T}{\sigma}\int_\Omega\tilde{T}_\sigma^+(b_{S_n}(x,u_0^1)-b_{S_n}(x,u_0^2))\,dx\end{aligned}\tag{3.10}$$

where $\tilde{T}_\sigma^+(r)=\int_0^r T_\sigma^+(s)\,ds$.

Due to the assumption $u_0^1\le u_0^2$ a.e. in Ω and the monotone character of $b_{S_n}(x,\ .)$ and $T_\sigma(.)$, we have

$$\int_\Omega\tilde{T}_\sigma^+\Big(b_{S_n}(x,u_0^1)-b_{S_n}(x,u_0^2)\Big)\,dx=0.\tag{3.11}$$

It follows from (3.5), (3.10) and (3.11) that

$$\frac{1}{\sigma}\int_Q\tilde{T}_\sigma^+\Big(b_{S_n}(x,u_1)-b_{S_n}(x,u_2)\Big)\,dx\,dt+A_n^\sigma=B_n^\sigma+C_n^\sigma+D_n^\sigma\tag{3.12}$$

for any $\sigma>0$ and any $n>0$.

Step 2. In this step, we study the behaviors of the terms A_n^σ, B_n^σ, C_n^σ and D_n^σ when σ tends to 0 and $n\to+\infty$. More precisely, we prove the following estimates,

$$\lim_{n\to+\infty}\underline{\lim}_{\sigma\to0}A_n^\sigma\ge0,\tag{3.13}$$

$$\underline{\lim}_{n\to+\infty}\underline{\lim}_{\sigma\to0}B_n^\sigma=0,\tag{3.14}$$

$$\underline{\lim}_{\sigma\to0}C_n^\sigma=0\quad\text{for all } n,\tag{3.15}$$

$$\underline{\lim}_{n\to+\infty}\overline{\lim}_{\sigma\to0}D_n^\sigma\le0.\tag{3.16}$$

Proof of (3.13)

$$\begin{aligned}A_n^\sigma=\frac{1}{\sigma}\int_0^T\int_0^t\int_\Omega\Big[S_n'(u_1)A(t,x)Du_1-S_n'(u_2)A(t,x)Du_2\Big]\\ DT_\sigma^+\Big(b_{S_n}(x,u_1)-b_{S_n}(x,u_2)\Big)\,dx\,ds\,dt.\end{aligned}$$

To establish (3.13) we first write A_n^σ, as follows

$$
\begin{aligned}
A_n^\sigma = & \int_Q \frac{(T-t)}{\sigma}\Big[S_n'(u_1)\Big(\frac{\partial b(x,u_1)}{\partial s}\Big)^{\frac{1}{2}} A(t,x)^{\frac{1}{2}} Du_1 + \\
& - S_n'(u_2)\Big(\frac{\partial b(x,u_2)}{\partial s}\Big)^{\frac{1}{2}} A(t,x)^{\frac{1}{2}} Du_2\Big]^2 (T_\sigma^+)'(b_{S_n}(x,u_1) - b_{S_n}(x,u_2))\,dx\,dt + \\
& - \int_Q \frac{(T-t)}{\sigma}\Big[\Big(\frac{\partial b(x,u_1)}{\partial s}\Big)^{\frac{1}{2}} - \Big(\frac{\partial b(x,u_2)}{\partial s}\Big)^{\frac{1}{2}}\Big]^2 A(t,x) DS_n(u_1) DS_n(u_2) \cdot \\
& \cdot (T_\sigma^+)'\Big(b_{S_n}(x,u_1) - b_{S_n}(x,u_2)\Big)\,dx\,dt + \int_Q \frac{(T-t)}{\sigma}\Big[S_n'(u_1)A(t,x)Du_1 + \\
& - S_n'(u_2)A(t,x)Du_2\Big]\Big[\int_{u_2}^{u_1} S_n'(s)\nabla_x\Big(\frac{\partial b(x,s)}{\partial s}\Big)\,ds\Big] \cdot \\
& \cdot (T_\sigma^+)'\Big(b_{S_n}(x,u_1) - b_{S_n}(x,u_2)\Big)\,dx\,dt.
\end{aligned} \tag{3.17}
$$

We denote by C_n the compact subset $[-n-1, n+1]$ of $\mathbb{R}$, and remark that due to (2.2) and (3.1), there exist a positive real numbers λ_n and β_n such that

$$
\begin{aligned}
& \Big|\Big(\frac{\partial b(x,z_1)}{\partial s}\Big)^{\frac{1}{2}} - \Big(\frac{\partial b(x,z_2)}{\partial s}\Big)^{\frac{1}{2}}\Big| \leq \\
& \leq \frac{\beta_n}{2\sqrt{\lambda_n}}\,|z_1 - z_2| \text{ for all } z_1,\ z_2 \text{ lying in } C_n,
\end{aligned} \tag{3.18}
$$

for almost every x in Ω.

Due to the definition of $b_{S_n}(x,r)$, we have

$$
\Big|b_{S_n}(x,s) - b_{S_n}(x,r)\Big| = \Big|\int_r^s S_n'(z)\frac{\partial b(x,z)}{\partial z}\,dz\Big| \geq \lambda_n\Big|S_n(s) - S_n(r)\Big| \tag{3.19}
$$

for almost every x in Ω, and $\forall\ s,\ r \in \mathbb{R}$.

As a consequence it follows that for $\sigma < n$ and if s and r are real numbers such that $\Big|S_n(s) - S_n(r)\Big| \leq \sigma$, then both $S_n(s)$ and $S_n(r)$ belong to concave or to convex branch of S_n. For $\sigma < n$, we then have:

$$
\min\Big(S_n'(s),\ S_n'(r)\Big)|r-s| \leq \Big|S_n(s) - S_n(r)\Big|
$$

for all real numbers such that $\Big|S_n(s) - S_n(r)\Big| \leq \sigma$.

From the above inequality and since $\|S_n'\|_{L^\infty(\mathbb{R})} = 1$ we deduce that

$$
\Big|S_n(s) - S_n(r)\Big| \leq \sigma < n \Longrightarrow S_n'(s)S_n'(r)|s-r| \leq \Big|S_n(s) - S_n(r)\Big|.
$$

Due to the definition of T_σ^+, it follows that

$$S'_n(s)S'_n(r)|s-r|(T_\sigma^+)'(S_n(s)-S_n(r)) \le \sigma\ (T_\sigma^+)'(S_n(s)-S_n(r)) \tag{3.20}$$

for all numbers s and r.

Recalling that $supp(S'_n) \subset [-(n+1),\ n+1]$, inequalities (3.18) and (3.19) lead to:

$$\begin{aligned}
&\Big|\int_Q \frac{(T-t)}{\sigma}\Big[\Big(\frac{\partial b(x,u_1)}{\partial s}\Big)^{\frac12} - \Big(\frac{\partial b(x,u_2)}{\partial s}\Big)^{\frac12}\Big]^2 \cdot \\
&\quad \cdot (T_\sigma^+)'\Big(b_{S_n}(x,u_1) - b_{S_n}(x,u_2)\Big)A(t,x)DS_n(u_1)DS_n(u_2)\,dx\,dt\Big| \le \\
&\le \frac{T\beta_n}{2\sqrt{\lambda_n}} \int_Q \Big|\Big(\frac{\partial b(x,u_1)}{\partial s}\Big)^{\frac12} - \Big(\frac{\partial b(x,u_2)}{\partial s}\Big)^{\frac12}\Big|\chi_{\{u_1\neq u_2\}}\cdot \\
&\quad \cdot (T^+_{\frac{\sigma}{\lambda_n}})'\Big(S_n(u_1) - S_n(u_2)\Big)\Big|A(t,x)DT_{n+1}(u_1)DT_{n+1}(u_2)\Big|\,dx\,dt.
\end{aligned}$$

The term just above is easily shown to converge to zero as σ goes to zero since the function

$$\begin{aligned}
&\Big|\Big(\frac{\partial b(x,u_1)}{\partial s}\Big)^{\frac12} - \Big(\frac{\partial b(x,u_2)}{\partial s}\Big)^{\frac12}\Big|\chi_{\{u_1\neq u_2\}}\cdot \\
&\quad \cdot (T^+_{\frac{\sigma}{\lambda_n}})'\Big(S_n(u_1) - S_n(u_2)\Big)\Big|A(t,x)DT_{n+1}(u_1)DT_{n+1}(u_2)\Big|
\end{aligned}$$

converges to zero almost everywhere in Q as σ goes to zero and (due to (3.1)) is bounded by the $L^1(Q)$-function $2\|\frac{\partial b(x,s)}{\partial s}\|_{L^\infty(\Omega\times C_n)}|A(t,x)DT_{n+1}(u_1)DT_{n+1}(u_2)|$. We remark that

$$\begin{aligned}
&\Big|\int_Q \frac{(T-t)}{\sigma}[S'_n(u_1)A(t,x)Du_1 - S'_n(u_2)A(t,x)Du_2](T_\sigma^+)'(b_{S_n}(x,u_1)+ \\
&\quad - b_{S_n}(x,u_2))\Big[\int_{u_2}^{u_1} S'_n(s)\nabla_x\Big(\frac{\partial b(x,s)}{\partial s}\Big)\,ds\Big]\,dx\,dt\Big| \le \\
&\le \Big|\int_Q \frac{(T-t)}{\sigma}\Big|S'_n(u_1)A(t,x)Du_1 - S'_n(u_2)A(t,x)Du_2\Big|(T_\sigma^+)'(b_{S_n}(x,u_1)+ \\
&\quad - b_{S_n}(x,u_2))\,\Big\|\nabla_x\Big(\frac{\partial b(x,s)}{\partial s}\Big)\Big\|_{L^\infty(C_n)}\Big|S_n(u_1)-S_n(u_2)\Big|\,dx\,dt \le \\
&\le T\int_Q \Big|S'(u_1)A(t,x)DT_{n+1}(u_1) - S'_n(u_2)A(t,x)DT_{n+1}(u_2)\Big|\chi_{\{u_1\neq u_2\}}\cdot \\
&\quad \cdot (T^+_{\frac{\sigma}{\lambda_n}})'\Big(S_n(u_1)-S_n(u_2)\Big)\|\nabla_x\Big(\frac{\partial b(x,s)}{\partial s}\Big)\|_{L^\infty(C_n)}\,dx\,dt.
\end{aligned} \tag{3.22}$$

The term just above is easily shown to converge to zero as σ goes to zero since the function

$$\Big|S'(u_1)A(t,x)DT_{n+1}(u_1) - S_n'(u_2)A(t,x)DT_{n+1}(u_2)\Big|\chi_{\{u_1\neq u_2\}}\cdot$$
$$\cdot (T^+_{\frac{\sigma}{\lambda_n}})'\Big(S_n(u_1) - S_n(u_2)\Big)\Big\|\nabla_x\Big(\frac{\partial b(x,s)}{\partial s}\Big)\Big\|_{L^\infty(C_n)}$$

converges to zero almost everywhere in Q as σ goes to zero and is bounded by the $L^1(Q)$-function

$$\Big|S'(u_1)A(t,x)DT_{n+1}(u_1) - S_n'(u_2)A(t,x)DT_{n+1}(u_2)\Big|B_n(x)$$

since $\|\nabla_x\Big(\frac{\partial b(x,s)}{\partial s}\Big)\|_{L^\infty(C_n)} \leq B_n(x) \in L^2(\Omega)$ (see (2.2)).

From the above analysis we conclude that (3.13) holds true.

Proof of (3.14). We have

$$|B_n^\sigma| \leq T\int_{\{n\leq|u_1|\leq n+1\}} A(t,x)Du_1Du_1\,dx\,dt + \\ + T\int_{\{n\leq|u_2|\leq n+1\}} A(t,x)Du_2Du_2\,dx\,dt. \tag{3.23}$$

As a consequence of (2.8), letting n go to infinity in the above estimates of B_n^σ shows that (3.14) holds true.

Proof of (3.15).

$$C_n^\sigma = \frac{1}{\sigma}\int_0^T\int_0^t\int_\Omega \Big[\Phi_{S_n}(u_1) - \Phi_{S_n}(u_2)\Big]DT_\sigma^+\Big(b_{S_n}(x,u_1) - b_{S_n}(x,u_2)\Big)\,dx\,ds\,dt.$$

To establish (3.15), let us remark that for all $s,\ r$ in $\mathbb{R}$, the following inequality holds true

$$\Big\|\Phi_{S_n}(s) - \Phi_{S_n}(r)\Big\|_{\mathbb{R}^N} \leq \|\Phi'\|_{L^\infty(C_n)}{}^N\Big|S_n(s) - S_n(r)\Big| \tag{3.24}$$

indeed, since $supp S_n' \subset [-n-1,\ n+1]$ and Φ' is assumed to be locally Lipschitz continuous, it follows that

$$\Big|\Phi_{S_n}(s) - \Phi_{S_n}(r)\Big| \leq \Big|\int_r^s S'(z)\Phi'(z)\,dz\Big| \leq \|\Phi'\|_{L^\infty(C_n)^N}\Big|S_n(s) - S_n(r)\Big|.$$

With the help of (3.24) the term C_n^σ may be estimated as follows

$$|C_n^\sigma| \le T\|\Phi'\|_{L^\infty(C_n)}{}^N \int_{\{0\le\left(b_{S_n}(x,u_1)-b_{S_n}(x,u_2)\right)\le\sigma\}} \frac{\left|S_n(u_1)-S_n(u_2)\right|}{\sigma}\cdot$$
$$\cdot\left|DT_\sigma^+\left(b_{S_n}(x,u_1)-b_{S_n}(x,u_2)\right)\right| dx\,dt \le$$
$$\le T\|\Phi'\|_{L^\infty(C_n)}{}^N \int_{\{0\le\left(S_n(u_1)-S_n(u_2)\right)\le\dfrac{\sigma}{\lambda_n}\}} \frac{\left|S_n(u_1)-S_n(u_2)\right|}{\sigma}\cdot$$
$$\cdot\left|DT_\sigma^+\left(b_{S_n}(x,u_1)-b_{S_n}(x,u_2)\right)\right| dx\,dt.$$

It yields

$$|C_n^\sigma| \le \frac{T}{\lambda_n}\|\Phi'\|_{L^\infty(C_n)^N}\int_Q\left|DT_\sigma^+\left(b_{S_n}(x,u_1)-b_{S_n}(x,u_2)\right)\right| dx\,dt,$$

which in turn implies (3.15) since $DT_\sigma^+\left(b_{S_n}(x,u_1)-b_{S_n}(x,u_2)\right)$ converges to zero in $L^1(Q)$ as σ goes to zero.

Proof of (3.16).

$$D_n^\sigma = \frac{1}{\sigma}\int_0^T\int_0^t\int_\Omega\left[f_1S_n'(u_1)-f_2S_n'(u_2)\right]T_\sigma^+\left(b_{S_n}(x,u_1)-b_{S_n}(x,u_2)\right) dx\,ds\,dt.$$

We have as n tends to $+\infty$,

$$f_1S_n'(u_1)-f_2S_n'(u_2)\to f_1-f_2 \ \text{ strongly in } L^1(Q). \tag{3.25}$$

Letting σ tends to 0, we have $\frac{1}{\sigma}T_\sigma^+(t)$ goes to $sg_0^+(t)$, for all $t\in\mathbb{R}$. For $n>0$ fixed, we have

$$\lim_{\sigma\to0} D_n^\sigma = \int_0^T\int_0^t\int_\Omega\left(f_1-f_2\right)sg_0^+\left(b_{S_n}(x,u_1))-b_{S_n}(x,u_2))\right) dx\,ds\,dt.$$

Since $f_1\le f_2$ a.e. in Q and $sg_0^+(t)\ge0$ for all t in $\mathbb{R}$, then shows that (3.16) holds true. In view of the definition of $\tilde{T}_\sigma^+$ and T_n^σ, we have

$$\begin{aligned}&\lim_{n\to+\infty}\lim_{\sigma\to0}\frac{1}{\sigma}\int_0^T\int_\Omega\tilde{T}_\sigma^+\left(b_{S_n}(x,u_1)-b_{S_n}(x,u_2)\right)dx\,dt=\\&=\int_Q\left(b(x,u_1)-b(x,u_2)\right)^+dx\,dt.\end{aligned} \tag{3.26}$$

In view of estimates (3.11), (3.12), (3.13), (3.14), (3.15), (3.16) and (3.26) we have

$$\int_Q \Big(b(x,u_1) - b(x,u_2)\Big)^+ \, dx\, dt \le 0,$$

so that $b(x,u_1) \le b(x,u_2)$ a.e. in Q which in turn implies that $u_1 \le u_2$ a.e. in Q, Theorem 3.0.4 is then established.

Acknowledgement

We would like to thank Dominique Blanchard for his attention to the work and Lucio Boccardo for his valuable suggestions.

REFERENCES

[1] P. Bénilan – L. Boccardo – T. Gallouët – R. Gariepy – M. Pierre – J. L. Vazquez: *An L^1-theory of existence and uniqueness of solutions of nonlinear elliptic equations*, Ann. Scuola Norm. Sup. Pisa Cl. Sci., **22** (1995), 240-273.

[2] D. Blanchard – F. Murat: *Renormalized solutions of nonlinear parabolic problems with L^1 data, Existence and uniqueness*, Proc. Roy. Soc. Edinburgh Sect., **A127** (1997), 1137-1152.

[3] D. Blanchard – F. Murat – H. Redwane: *Existence et unicité de la solution reormalisée d'un problème parabolique assez général*, C. R. Acad. Sci. Paris Sér., **I329** (1999). 575-580.

[4] D. Blanchard – F. Murat – H. Redwane: *Existence and Uniqueness of a Renormalized Solution for a Fairly General Class of Nonlinear Parabolic Problems*, J. Differential Equations, **177** (2001), 331-374.

[5] D. Blanchard – A. Porretta: *Stefan problems with nonlinear diffusion and convection*, J. Diff. Equations, **210** (2005), 383-428.

[6] D. Blanchard – H. Redwane: *Sur la résolution d'un problème quasi-linéaire à coefficients singuliers par rapport à l'inconnue*, C. R. Acad. Sci. Paris Sér., **I325** (1997), 993-998.

[7] D. Blanchard – H. Redwane: *Renormalized solutions of nonlinear parabolic evolution problems*, J. Math. Pure Appl., **77** (1998), 117-151.

[8] L. Boccardo – D. Giachetti – J. I. Diaz – F. Murat: *Existence and regularity of renormalized solutions for some elliptic problems involving derivation of nonlinear terms*, J. Differential Equations, **106** (1993), 215-237.

[9] L. Boccardo – F. Murat – J. P. Puel: *Existence of bounded solutions for nonlinear elliptic unilateral problems*, Ann. Mat. Pura Appl., **152** (1988), 183-196.

[10] P.-L. Lions: *Mathematical Topics in Fluid Mechanics*, Incompressible models, vol. 1, Oxford Univ. Press, 1996.

[11] P.-L. Lions – F. Murat: *Solutions renormalisées d'équations elliptiques*, in preparation.

[12] F. Murat: *Soluciones renormalizadas de EDP elipticas non lineales, Cours à l'Université de Séville*, Publication R93023, Laboratoire d'Analyse Numérique, Paris VI, 1993.

[13] F. Murat: *Equations elliptiques non linéaires avec second membre L^1 ou mesure*, Comptes Rendus du 26ème Congrès National d'Analyse Numérique Les Karellis, A12-A24, 1994.

[14] R.-J. Di Perna – P.-L. Lions: *On the Cauchy problem for Boltzmann equations: Global existence and weak stability*, Ann. Math., **130** (1989), 321-366.

[15] H. Redwane: *Existence of a solution for a class of parabolic equations with three unbounded nonlinearities*, Adv. Dyn. Syst. Appl., **2** (2007), 241-264.

Lavoro pervenuto alla redazione il 18 gennaio 2008
ed accettato per la pubblicazione il 15 aprile 2008.
Bozze licenziate il 30 settembre 2008

Indirizzo dell'Autore:

Hicham Redwane – Faculté des Sciences Juridiques – Economiques et Sociales – Université Hassan 1 – B.P. 784 Settat – Morocco
E-mail: redwane.hicham@mailcity.com

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 201–228

Sulla struttura combinatoria dei linguaggi bounded

FLAVIO D'ALESSANDRO – ALESSANDRA ZINNO PILO

Abstract: *This paper concerns the studying of some algebraic and algorithmical problems on the growth function of formal languages.*

Proporremo una rassegna sintetica di alcuni argomenti relativi alla teoria delle funzioni di crescita dei linguaggi formali. Come suggerisce il titolo della nota, la sua redazione è essenzialmente legata ad una classe significativa di linguaggi, detti *bounded*, di cui analizzeremo le principali proprietà strutturali ed i legami con altre tematiche di interesse per la matematica e per l'informatica. Il taglio editoriale di questa nota può essere così rapidamente descritto: gli aspetti tecnici sono presentati in modo contenuto, privilegiando, nella misura maggiore possibile, le idee soggiacenti ai teoremi e alle loro dimostrazioni, nonché alle relazioni con altri aspetti interessanti della teoria. Alcune definizioni sono volutamente omesse, altre, invece, sinteticamente presentate. La speranza è che chi abbia dimistichezza con la matematica approfitti senza difficoltà di questo breve excursus in una tematica di così grande interesse e attualità per l'informatica.

1 – Introduzione

Sia L un linguaggio. La *funzione di conteggio* di L, detta anche *funzione di struttura*, è la funzione f_L che associa, ad ogni intero non negativo n, il numero $f_L(n)$ delle parole di L di lunghezza uguale ad n. La *funzione di crescita* di L è

Key Words and Phrases: *Bounded language – Context-free language – Growth function.*
A.M.S. Classification: 68Q45

invece la funzione g_L che associa, ad ogni intero non negativo n, il numero $g_L(n)$ delle parole di L di lunghezza uguale o inferiore ad n.

Queste funzioni sono di grande interesse poiché lo studio del loro andamento asintotico fornisce preziose informazioni sulla struttura combinatoria del linguaggio considerato. Le nozioni di funzione di struttura e di funzione di crescita giocano un ruolo importante nello studio di problemi significativi formulati in altri ambiti distinti da quello della teoria dei linguaggi. Ad esempio, nella teoria matematica dell'informazione, due risultati fondamentali sui codici di parole a lunghezza variabile sono basati su tale nozione (*cfr.* [4]): la Disuguaglianza di Kraft-Mc Millan che fornisce una condizione necessaria affinché un insieme di parole sia codice oppure il teorema di Kraft relativo alla esistenza di codici prefissi ottimali per una data sorgente. Un altro notevole esempio ci è fornito dalla teoria dei gruppi dove un celeberrimo teorema dimostrato da M. Gromov in [20] permette di caratterizzare la struttura algebrica dei gruppi di crescita polinomiale: precisamente, un gruppo siffatto possiede un sottogruppo nilpotente di indice finito. Altri risultati profondi sono stati ottenuti nello studio della crescita delle algebre finitamente generate relativamente alla dimensione di Gelfand-Kirillov di tali algebre. Rimandiamo al testo [15] quale eccellente e completo riferimento bibliografico per questo argomento.

Nella teoria dei linguaggi formali, questa tematica è stata oggetto, anche recentemente, di un'attività di ricerca molto intensa con particolare riferimento ai linguaggi a crescita polimoniale (*cfr.* [6], [8], [9], [10], [13], [14], [16], [17], [19], [22], [23], [25], [26], [27], [29], [30], [32], [36]). Il punto di partenza dell'excursus che compiremo in questa nota è un quesito formulato da Flajolet in [16] relativo alla esistenza di linguaggi context-free di crescita intermedia. A tale quesito è stata data risposta negativa da Incitti in [30] e, qualche tempo dopo, indipendentemente, da Bridson e Gilman in [6], dimostrando la seguente proprietà di *gap*: *un linguaggio context-free ha crescita esponenziale oppure polinomiale.*

I linguaggi a crescita polinomiale, detti anche *sparsi*, svolgono un ruolo chiave sia nell'ambito della Teoria dei linguaggi formali, sia in quello della Complessità di calcolo. Questa circostanza è anche riconducibile al fatto che, nel caso context-free, la famiglia di tali linguaggi coincide con quella dei cosiddetti *linguaggi bounded* introdotti e studiati, per la prima volta, da S. Ginsburg nella seconda metà degli Anni 60 (*cfr.* [17]). Diversi risultati di questa ricca e ben sviluppata teoria sono stati riscoperti, anche in ambiti diversi da quello dei linguaggi, (e ripubblicati) più volte negli ultimi venti anni e uno degli scopi della nota è quello di mettere in evidenza alcuni aspetti, a nostro giudizio, rilevanti di questo affascinante argomento.

Eccoci infine, molto brevemente, a spiegare come si articolerà questa nota. Dopo aver presentato alcune definizioni preliminari, dimostreremo il teorema di gap e analizzeremo, sia dal punto di vista algebrico, sia da quello algoritmico, la struttura dei linguaggi context-free a crescita polinomiale. Ciò sarà fatto nelle Sezioni 2 e 3. Nella Sezione 4, invece, descriveremo un algoritmo il quale consente di calcolare in modo esatto la funzione di conteggio (e dunque anche quella di crescita) di un linguaggio context-free a crescita polinomiale. La Sezione 5 sarà invece dedicata alla descrizione del legame esistente fra queste tematiche ed un importante problema dell'Algebra: il *problema di Burnside per i semigruppi*. La Sezione 6 sarà dedicata ad un argomento "di frontiera": descriveremo una classe di linguaggi che pur avendo, nel senso che preciseremo poi, una struttura vicina a quella dei linguaggi context-free, sono a crescita intermedia. L'ultima sezione sarà invece dedicata alla presentazione di alcuni risultati ottenuti recentemente sulla funzione di conteggio delle relazioni razionali.

2 – Preliminari

Presenteremo molto brevemente, in questa sezione, un vocabolario minimo di definizioni soprattutto nell'intento di fissarne le corrispondenti notazioni. Quanto segue è in genere materia di un corso fondazionale di teoria dei linguaggi formali e può trovarsi nei testi, alcuni classici ed altri più recenti, di un'ampia rosa di scelta (*cfr.* [3], [12], [15], [17], [24], [33], [40]) alla quale rinviamo il lettore interessato.

2.1 – Monoidi liberi

Se A è un alfabeto, denoteremo con A^* il monoide libero delle parole, dette anche stringhe, sull'alfabeto A e con 1_{A^*}, o più semplicemente con 1, il suo elemento identità; tale elemento è detto *parola vuota*. L'insieme $A^* \setminus \{1\}$ è denotato A^+. Se u è una parola di A^*, la sua *lunghezza* è l'intero non negativo, denotato con il simbolo $|u|$, uguale a zero se $u = 1$ e, altrimenti, al numero dei simboli di cui si compone u. Fissata una lettera $a \in A$, denotiamo con $|u|_a$ il numero di occorrenze di a nella parola u. Siano u e w parole di A^*. La parola u si dice *fattore* di w se esistono λ, $\mu \in A^*$ tali che $\lambda u \mu = w$. L'insieme dei fattori di w è denotato con il simbolo $F(w)$. Similmente, se L è un insieme di parole di A^*, scriveremo $F(L) = \bigcup_{w \in L} F(w)$. u si dice poi *prefisso* di w se esiste $\mu \in A^*$ tale che $u\mu = w$. Infine u si dice *suffisso* di w se esiste $\lambda \in A^*$ tale che $\lambda u = w$. Una parola z è detta *primitiva* se $z = u^n$, con $n \geq 1$, implica $z = u$ e quindi $n = 1$. Ricordiamo infine che una congruenza di A^* è una relazione di equivalenza di A^* compatibile con l'operazione di prodotto di parole.

2.2 – Linguaggi context-free

Descriveremo ora in maggior dettaglio un'importante classe di linguaggi, quella dei *linguaggi context-free,* detti anche *liberi dal contesto*, introdotti in [11]. Questi linguaggi sono definiti tramite l'uso di certe strutture combinatorie dette *grammatiche context-free* di cui richiameremo la definizione anche al fine di fissare qualche utile notazione. Formalmente, una grammatica context-free è una quadrupla $G = (V, T, P, S)$, dove:

- V è un insieme finito di oggetti detti *variabili* della grammatica G;
- T è un insieme finito di simboli detti *terminali* della grammatica G, ed inoltre T è disgiunto da V;
- P è un sottoinsieme finito dell'insieme $V \times \{V \cup T\}^*$, detto delle *produzioni* della grammatica G; ogni produzione è usualmente denotata con la forma $A \to \alpha$ dove A è una variabile ed α è una stringa di simboli di $(V \cup T)^*$;
- S è una variabile detta *start symbol*, o *assioma della grammatica*.

Vogliamo ora definire il linguaggio generato da G. A tale scopo definiamo prima le seguenti due relazioni $\Rightarrow$ e $\overset{*}{\Rightarrow}$ tra stringhe in $(V \cup T)^*$. Se u, v sono parole in $(V \cup T)^*$, allora porremo

$$u \Rightarrow v$$

se esiste una produzione $A \to \beta$ della grammatica G tale che $u = \alpha A\gamma$ e $v = \alpha\beta\gamma$, dove $\alpha,\ \gamma \in (V \cup T)^*$. Le parole $\alpha,\ \gamma$ sono chiamate *contesti* di u e v e la relazione $\Rightarrow$ è detta di *derivazione atomica*. La relazione $\overset{*}{\Rightarrow}$ è poi definita quale chiusura transitiva e riflessiva di $\Rightarrow$, ovvero $\overset{*}{\Rightarrow} = \bigcup_{i \geq 0} \overset{i}{\Rightarrow}$. Questa relazione è detta di *derivazione* di G. In altre parole, se $\alpha,\ \beta \in (V \cup T)^*$ sono tali che $\alpha \Rightarrow^* \beta$, allora $\alpha = \beta$ oppure esistono parole $\alpha_0, \alpha_1, \ldots, \alpha_n, n \geq 1$ di $(V \cup T)^*$ tali che

$$\alpha = \alpha_0 \Rightarrow \alpha_1 \Rightarrow \cdots \alpha_{n-1} \Rightarrow \alpha_n = \beta.$$

Una qualsiasi stringa α tale che $S \overset{*}{\Rightarrow} \alpha$ si dirà *forma sentenziale* di G. Una variabile A si dice *utile* se esiste una parola α in T^* tale che $S \overset{*}{\Rightarrow} fAg \overset{*}{\Rightarrow} \alpha$, dove $f, g \in (V \cup T)^*$. Il motivo per cui queste grammatiche sono chiamate context-free, letteralmente "libere dal contesto", risiede nel fatto che, in un passo qualsiasi di una generica derivazione di G, l'applicazione di una produzione non dipende dalla scelta dei contesti della forma sentenziale a cui si applica. Infine il linguaggio generato da G, denotato con il simbolo $L(G)$, è l'insieme:

$$L(G) = \{w \in T^* \mid S \overset{*}{\Rightarrow} w\}.$$

Diremo che un linguaggio è *context-free* se esiste una grammatica context-free che lo genera. Diremo inoltre che due grammatiche context-free sono equivalenti se i linguaggi da esse generati coincidono. Ricordiamo infine che tali linguaggi sono anche detti *algebrici* poiché possono essere descritti quali componenti delle soluzioni di sistemi di equazioni polinomiali di parole, cioè di equazioni di parole in cui siano usati i soli operatori di unione e di concatenazione in A^*.

2.3 – Linguaggi regolari

Un'altra importante famiglia di linguaggi formali è quella dei cosiddetti *linguaggi regolari*. Tale famiglia di linguaggi corrisponde, rispetto alla classica gerarchia di Chomsky-Schützenberger, alla classe di linguaggi immediatamente precedente a quella dei linguaggi algebrici di cui ne costituisce una notevole sottoclasse. Un linguaggio su di un alfabeto A si dice *regolare* o *razionale* se può ottenersi a partire da linguaggi di cardinalità finita di A^*, tramite l'applicazione, in un numero finito di volte, delle operazioni razionali, cioè delle classiche operazioni di unione e di prodotto di linguaggi e di una terza operazione, questa più complessa, detta di *stella*, che associa ad ogni linguaggio L il sottomonoide L^* da esso generato. La famiglia di tali linguaggi è usualmente denotata con il simbolo $\mathrm{Rat}(A^*)$. In virtù della sua definizione, essa coincide con la più piccola famiglia di linguaggi di A^* contenente la famiglia dei linguaggi finiti e chiusa rispetto alle operazioni razionali. Due risultati fondamentali della teoria, entrambi dimostrati negli Anni '50, forniscono importanti caratterizzazioni dei linguaggi regolari. Il primo, dimostrato dal celebre logico americano Stephen Coole Kleene, è legato ad un modello di calcolo ben noto, quello di *automa a stati finiti*, e stabilisce che un linguaggio è regolare se e solo se esiste un automa a stati finiti che accetta (o riconosce) tale linguaggio. Il secondo teorema, dimostrato da John Myhill e Anil Nerode, costituisce una caratterizzazione algebrica di tali linguaggi basata sulle congruenze. A questo proposito, converrà ricordare che, dati una congruenza θ di A^* ed un linguaggio L, si dice che θ satura L, oppure che L è saturato da θ, se L è unione di classi di θ. Il teorema di Myhill e Nerode stabilisce che un linguaggio è regolare se e solo se esiste una relazione di congruenza di A^* di indice finito che satura il linguaggio. Come si è detto prima, i linguaggi regolari sono context-free: infatti essi sono generati da grammatiche context-free di tipo particolare, dette *lineari a destra* oppure *lineari a sinistra*. Ricordiamo che una grammatica context-free si dice lineare a destra (risp. lineare a sinistra) se ogni sua produzione $A \to \alpha$ è tale che α contiene, al più, una occorrenza di una variabile della grammatica e, se questo è il caso, tale variabile deve essere prefisso (risp. suffisso) di α. I teoremi ora ricordati mostrano dunque che la nozione di linguaggio regolare è robusta.

Ricordiamo infine un importante risultato, detto *Pumping Lemma*, che fornisce una condizione necessaria affinché un linguaggio sia regolare. Precisamente, se L è un linguaggio regolare su di un alfabeto A, esiste un intero non negativo n dipendente da L, tale che, per ogni parola $w \in L$ con $|w| \geq n$, w ammette una fattorizzazione del tipo $w = \alpha u \beta$ dove $\alpha, u, \beta \in A^*$, $u \neq \epsilon$ e, per ogni $i \geq 0$, $\alpha u^i \beta$ è una parola di L.

2.4 – Funzione di conteggio e funzione di crescita di un linguaggio

Siano A^* il monoide libero delle parole su di un dato alfabeto finito A ed L un sottoinsieme di A^*. La *funzione di struttura*, detta anche *funzione di conteggio* di L, è la funzione

$$f_L : \mathbb{N} \longrightarrow \mathbb{N},$$

così definita:

$$\forall\ n \in \mathbb{N}, \quad f_L(n) = \mathrm{Card}(L \cap A^n) = \mathrm{Card}(\{u \in L \mid |u| = n\}).$$

Si definisce poi *funzione di crescita* di L, la funzione

$$g_L : \mathbb{N} \longrightarrow \mathbb{N},$$

così definita:

$$\forall\ n \in \mathbb{N}, \quad g_L(n) = \mathrm{Card}(\{u \in L \mid |u| \leq n\}).$$

È facile osservare che, per ogni $n \in \mathbb{N}$, $g_L(n) = \sum_{i=0,\dots,n} f_L(i)$. È di grande utilità considerare una classificazione, riportata nella definizione seguente, dei linguaggi formali basata sull'andamento asintotico della funzione di crescita.

DEFINIZIONE 1. Sia $L \subseteq A^*$. Si dice che L è *fino* o *thin* se:

$$\forall\ n \geq 0, \qquad f_L(n) \leq 1.$$

Si dice invece che L è *snello* o *slender* se esiste un intero $C \geq 0$ tale che:

$$\forall\ n \geq 0, \qquad f_L(n) \leq C.$$

Si dice che L ha *crescita polinomiale* o è *poly-slender* se esiste $k > 0$ tale che:

$$\forall\ n \geq 0, \qquad g_L(n) \leq n^k.$$

Un linguaggio a crescita polinomiale è anche detto *sparso*.

Si dice che L è *quasi-polinomiale* se:

$$\lim_{n \to +\infty} \frac{f_L(n)}{(\mathrm{Card}(A))^n} = 0.$$

Si dice che L ha *crescita esponenziale* se esistono $k > 1$ ed un intero $n_0 \geq 0$ tali che:

$$\forall\ n \geq n_0, \qquad g_L(n) \geq k^n.$$

Semplici esempi di linguaggi a crescita polinomiale ed esponenziale sono i seguenti. Sia

$$L = \{a^n b^n \mid n \geq 1\}.$$

È facile osservare che la funzione di conteggio di L è definita come:

$$(1) \qquad \forall\ n \geq 0, \qquad f_L(n) = \begin{cases} 1 & \text{se n è pari} \\ 0 & \text{se n è dispari.} \end{cases}$$

Di conseguenza L è fino. In particolare, per ogni $n \geq 0$, $g_L(n) \leq n$, cioè L ha crescita lineare. Invece, come facilmente si verifica, i linguaggi

$$L = A^*, \qquad L' = \{ww \mid w \in A^*\}$$

sono tali che le rispettive funzioni di conteggio e di crescita sono esponenziali.

3 – Funzione di crescita dei linguaggi context–free

Intendiamo ora presentare alcuni notevoli risultati relativi alla struttura combinatoria dei linguaggi context-free: il primo riguarda una caratterizzazione dei linguaggi context-free a crescita polinomiale. Tale caratterizzazione è legata ad una importante definizione, quella di *linguaggio bounded*, che daremo nel seguito di questa sezione. Vedremo successivamente il teorema di *gap*, enunciato nell'Introduzione, per la funzione di crescita dei linguaggi context-free: la funzione di crescita di tali linguaggi è esponenziale oppure polinomiale. Mostreremo infine che il teorema di gap è effettivo: esiste un algoritmo, l'esecuzione del quale consente di decidere la natura, esponenziale oppure polinomiale, della funzione di crescita di un linguaggio siffatto. Nell'intento di dimostrare i risultati appena ricordati, riteniamo opportuno introdurre qualche definizione e presentare alcuni lemmi preliminari: cominceremo con la nozione, introdotta da Ginsburg in [17], di linguaggio dei cicli destri e sinistri associati ad una variabile della grammatica context-free.

Definizione 2. Siano $G = (V, T, P, S)$ una grammatica context-free, $L(G)$ il linguaggio da essa generato ed A una variabile di G. Denotiamo con L_A ed R_A i due linguaggi di parole sull'alfabeto T definiti come segue:

$$L_A = \{\alpha \in T^* \mid \exists\ \beta \in T^* \mid A \overset{*}{\Rightarrow} \alpha A \beta\},$$
$$R_A = \{\beta \in T^* \mid \exists\ \alpha \in T^* \mid A \overset{*}{\Rightarrow} \alpha A \beta\}.$$

L_A ed R_A prendono rispettivamente i nomi di *linguaggio dei cicli sinistri* e *linguaggio dei cicli destri della variabile A.*

È facile verificare che, per ogni variabile A di G, L_A ed R_A sono sottomonoidi di T^*. In effetti, se α, α' appartengono ad L_A, esistono $\beta, \beta' \in T^*$ tali che:

$$A \overset{*}{\Rightarrow} \alpha A \beta \text{ e } A \overset{*}{\Rightarrow} \alpha' A \beta'.$$

Poiché $A \overset{*}{\Rightarrow} \alpha A \beta \Rightarrow \alpha\alpha' A \beta\beta'$, si ha che $\alpha\alpha'$ appartiene ad L_A, cioè L_A è chiuso rispetto all'operazione di concatenazione di parole, e dunque L_A è un sottomonoide di T^*. Nello stesso modo, si dimostra che R_A è un sottomonoide di T^*. Il lemma seguente fornisce una proprietà combinatoria cruciale delle grammatiche, basata sulla definizione precedente, la quale assicura la crescita esponenziale dei linguaggi da esse generati.

Lemma 1. *Siano $G = (V, T, P, S)$ una grammatica context-free, $L = L(G)$ il linguaggio da essa generato ed A una variabile di G. Se $d_1,\ d_2 \in L_A$ (risp. $d_1, d_2 \in R_A$) sono parole tali che il sottomonoide $\{d_1, d_2\}^*$ è libero in T^*, allora L ha crescita esponenziale.*

Dim. Siano d_1, d_2 parole di L_A. Allora esistono parole $\alpha, \beta \in T^*$ tali che:

$$A \overset{*}{\Rightarrow} d_1 A \alpha \text{ e } A \overset{*}{\Rightarrow} d_2 A \beta.$$

Poiché la variabile A è utile, segue che esiste una derivazione della forma:

$$S \overset{*}{\Rightarrow} uAv \overset{*}{\Rightarrow} z, \text{ dove } u, v, z \in T^*.$$

Poniamo $|u| = N$, $|v| = M$, $|z| = Q$, $|d_1| = n_1$, $|d_2| = n_2$. Possiamo suppore che $n_1 = n_2 = n$ e $|\alpha| = |\beta| = m$, non essendo questa una restrizione. Infatti, nell'ipotesi in cui $n_1 \neq n_2$ (oppure $|\alpha| \neq |\beta|$), sarà sufficiente considerare il sottomonoide di $\{d_1, d_2\}^*$ generato dall'insieme $\{d_1 d_2, d_2 d_1\}$. Tale sottomonoide è ovviamente libero e le due parole che lo generano hanno la stessa lunghezza.

Si ha:

$$S \overset{*}{\Rightarrow} uAv \overset{*}{\Rightarrow} udA\gamma v \Rightarrow udz\gamma v, \tag{2}$$

dove $d \in \{d_1, d_2\}^*$ e $\gamma \in \{\alpha, \beta\}^*$.

Se $w = udz\gamma v$, si ha che $w \in L$ dato che $w \in T^*$ e $S \overset{*}{\Rightarrow} w$. Si ottiene:

$$|w| = |u| + |d| + |z| + |\gamma| + |v| = N + Q + M + |d| + |\gamma|.$$

Ponendo $N + Q + M = c$, possiamo scrivere:

$$|w| = c + |d| + |\gamma| = c + kn + km = c + k(n + m),$$

dove k è il numero di occorrenze delle parole d_1 e d_2 nella parola d ed è, ovviamente, anche il numero di occorrenze delle parole α e β nella parola γ. Indichiamo con L' l'insieme delle parole ottenute come nella Equazione (2) e osserviamo che $L' \subseteq L$. Sia g_L la funzione di crescita di L. Per ogni intero positivo i sufficientemente grande, si ha:

$$g_L(i) \geq g_{L'}(i) \geq 2^{\frac{i-c}{n+m}}.$$

Dalla condizione precedente segue che la funzione di crescita di L è esponenziale. La dimostrazione per R_A è simmetrica.

Nel seguito di queste pagine, dati due interi i e j, l'insieme degli interi compresi tra i e j sarà indicato col simbolo $[i, j]$.

Definizione 3. Sia $G = (V, T, P, S)$ una grammatica context-free e sia $V = \{S = S_1, S_2, \ldots, S_n\}$. Definiamo l'insieme V' ponendo $V' = \emptyset$ se $n = 1$, e $V' = \{S_2, \ldots, S_n\}$ se $n > 1$. Definiamo poi, per ogni $k = 1, \ldots, n$, la grammatica $G_k = (V', T, P_k, S_k)$ le cui produzioni sono quelle di G della forma:

$$A \to \alpha, \text{ con } A \in V' \cup \{S_k\} e\ \alpha \in (V' \cup T)^*.$$

Denotiamo con $L'(G)$ il sottoinsieme di $L(G)$ costituito dalle parole $w \in T^*$ che ammettono una derivazione:

$$S \Rightarrow w_1 \cdots \Rightarrow w_m = w,$$

tale che, per ogni $i = 1, \ldots, m$, $w_i \in (V' \cup T)^*$. È facile osservare che, per $n = 1$, $L'(G)$ è il linguaggio finito delle parole di $L(G)$ che possono essere ottenute, a partire da S_1, mediante un solo passo di derivazione. Enunciamo ora, senza dimostrarlo, un lemma che utilizzeremo nella dimostrazione del prossimo teorema.

Lemma 2 ([30] Lemma 3.6). *$L(G) \subseteq L_S L'(G) R_S$.*

Siamo ora in grado di dimostrare una significativa caratterizzazione dei linguaggi context-free a crescita polinomiale. Tale caratterizzazione è legata ad una importante definizione, quella di *linguaggio bounded*.

Definizione 4. Siano A un alfabeto, L un linguaggio di A^* ed n un intero positivo. Il linguaggio L è detto *n-bounded* se esistono parole $u_1, \ldots, u_n \in A^+$ tali che:

$$L \subseteq \{u_1\}^* \cdots \{u_n\}^*.$$

L è detto *bounded* se esiste un intero n tale che L è *n-bounded*.

È utile osservare che, dalla definizione precedente, si ha che la famiglia di tali linguaggi è chiusa rispetto alle operazioni di unione insiemistica e di prodotto di linguaggi.

La nozione di linguaggio bounded gioca un ruolo importante nello studio di diversi problemi di Matematica e di Informatica. In questo ultimo ambito l'interesse di tali linguaggi risiede, ad esempio, nel fatto che tutte le più significative proprietà dei linguaggi formali sono ricorsivamente decidibili rispetto alla famiglia dei linguaggi context-free bounded quando, in generale, non lo sono. La struttura combinatoria e le proprietà algoritmiche dei linguaggi context-free bounded sono state studiate approfonditamente da S. Ginsburg in [17]. Inoltre, in un lavoro del 1984 [38], A. Restivo e C. Reutenauer, nell'ambito dello studio del *problema di Burnside per i semigruppi*, hanno introdotto una condizione di finitezza, volta cioè ad assicurarne la finitezza del sostegno, per i semigruppi periodici, detta *proprietà di permutazione*, strettamente legata ai linguaggi bounded. Approfondiremo quest'ultimo argomento nella prossima sezione.

Teorema 1. *Sia G una grammatica context-free ed $L = L(G)$ il linguaggio da essa generato. Le seguenti condizioni sono equivalenti:*

1. *L ha crescita polinomiale*;
2. *Per ogni variabile A di G, L_A e R_A sono linguaggi* 1*-bounded*;
3. *L è un linguaggio bounded.*

Prima di dimostrare il Teorema 1, converrà ricordare che l'equivalenza delle condizioni (2) e (3) è stata dimostrata da S. Ginsburg nella prima metà degli anni sessanta (*cfr.* [17]), mentre l'equivalenza delle condizioni (1) e (3) è stata dimostrata da Latteux e Thierrin in [32] e ridimostrata successivamente in [27] e [36]. Per la dimostrazione dell'implicazione (2) $\Longrightarrow$ (3) utilizzeremo un argomento combinatorio descritto in [30].

DIM. ($1 \Rightarrow 2$).

Sia A una variabile di G e sia L_A il linguaggio dei cicli sinistri associato ad A.

Se $L_A = \{1_{T^*}\}$ allora L_A è banalmente un linguaggio 1-bounded. Siano ora u e z una parola non vuota di L_A e la sua radice primitiva. Supponiamo per assurdo che L_A non sia 1-bounded. Deve allora esistere una parola $v \in (L_A \setminus \{z\}^*)$. Consideriamo il sottomonoide $\{u, v\}^*$ di L_A. Per il *Teorema del difetto*, esso deve essere libero, e per il Lemma 1, L ha crescita esponenziale, così contraddicendo la condizione (1). Dunque L_A è un linguaggio 1-bounded. Lo stesso argomento consente di dimostrare che R_A è un linguaggio 1-bounded. L'implicazione ($1 \Rightarrow 2$) è dunque dimostrata.

($2 \Rightarrow 3$).

Dimostriamo il risultato per induzione sul numero n di variabili della grammatica G. Sia $n = 1$. Il linguaggio $L'(G)$ è finito e dunque bounded. Per ipotesi, L_S e R_S sono linguaggi 1-bounded e, dunque, si ha che $L_S L'(G) R_S$ è bounded. Per il Lemma 2, poiché $L \subseteq L_S L'(G) R_S$, ne segue infine che L è bounded. La base del procedimento induttivo è pertanto dimostrata.

Dimostriamo ora il passo induttivo. Osserviamo intanto che, per ipotesi induttiva, $L(G_k)$ è bounded, per ogni $k \in [2, n]$. In virtù del Lemma 2 e dal momento che, per ipotesi, L_S ed R_S sono linguaggi 1-bounded è sufficiente dimostrare che $L'(G)$ è un linguaggio bounded. Sia $W = \{\alpha_1, \dots, \alpha_m\}$ l'insieme di parole sull'alfabeto $(V' \cup T)$ tali che, per ogni $i = 1, \dots, m$, $S \to \alpha_i$ sono produzioni della grammatica G. Quindi, per ogni $i = 1, \dots, m$, si avrà

$$\alpha_i = r_{1,i} S_{k_{1,i}} \cdots r_{x_i,i} S_{k_{x_i},i} r_{x_i+1,i},$$

dove $r_{1,i}, \dots, r_{x_i+1,i} \in T^*$ e $S_{k_1,i}, \dots, S_{k_{x_i},i} \in V'$.

Sia ora $w \in L'(G)$. Allora w sarà ottenuta tramite una derivazione del tipo:

$$S \Rightarrow w_1 \dots \Rightarrow w_l = w,$$

dove, per ogni $j = 1, \dots, l$, $w_j \in (V' \cup T)^*$.

Osserviamo che $w_1 \in W$ e che nelle stringhe $w_1, \dots, w_l$ non occorre mai la variabile S. Di conseguenza si avrà:

$$w \in r_{1,i} L(G_{k_1,i}) \cdots r_{x_i,i} L(G_{k_{x_i},i}) r_{x_i+1,i},$$

e dunque:

$$L'(G) \subseteq \bigcup_{i=1}^{m} r_{1,i} L(G_{k_1,i}) \cdots r_{x_i,i} L(G_{k_{x_i},i}) r_{x_i+1,i}. \tag{3}$$

Osserviamo che, per ogni $k > 0$, l'insieme delle produzioni della grammatica G_k è un sottoinsieme dell'insieme delle produzioni della grammatica G; pertanto l'ipotesi definita dalla condizione (2) dell'enunciato del teorema vale ancora se in luogo di considerare la grammatica G si considera la grammatica G_k. Applicando l'ipotesi induttiva alla grammatica G_k si avrà allora che $L(G_k)$ è un linguaggio bounded. Dalla Equazione (3) segue dunque che $L'(G)$ è un linguaggio bounded. L'implicazione ($2 \Rightarrow 3$) è pertanto dimostrata.

($3 \Rightarrow 1$).

Dal momento che L è bounded, esiste un intero $q \geq 1$ e parole non vuote $u_1, \ldots, u_q$ tali che $L \subseteq \{u_1\}^* \cdots \{u_q\}^*$. Sia $w \in L$ tale che $|w| \leq n$.

Allora

$$w = {u_1}^{t_1} \cdots {u_q}^{t_q},$$

dove per ogni $i \in [1, q]$, $t_i \geq 0$ e $t_1 + \cdots + t_q \leq n$.

Indichiamo con $\phi(q, n) = \mathrm{Card}(\{(t_1, \ldots, t_q) \mid t_1 + \cdots + t_q \leq n\})$.

Osserviamo ora che, per ogni $n \geq 0$, $g_L(n) \leq \phi(q, n)$. Poiché la funzione $\phi(q, n)$ è un polinomio nella variabile n assumendo che q sia costante, ne segue che la funzione g_L è limitata superiormente da una funzione polinomiale, ovvero L ha crescita polinomiale.

Come conseguenza del Teorema 1 si ha il corollario seguente.

Teorema 2. *La funzione di crescita di un linguaggio context-free è esponenziale oppure polinomiale.*

Dim. Sia L un linguaggio context-free la cui funzione di crescita non è polinomiale. In virtù del teorema precedente, esiste una variabile A della grammatica che genera il linguaggio tale che L_A oppure R_A non sono linguaggi 1-bounded. Supponiamo, per semplicità che ciò accada per L_A. Esistono allora parole non vuote $u, v \in L_A$ le cui radici primitive sono distinte. Per il *Teorema del difetto*, il sottomonoide $\{u, v\}^*$ è libero in L_A e, per il Lemma 1, L ha crescita esponenziale.

Converrà ricordare che, benché il Teorema 2 sia un corollario immediato del Teorema 1, il suo enunciato è stato dimostrato esplicitamente da Incitti in [30] e poco tempo più tardi da Bridson e Gilman in [6]. Esso in particolare fornisce la risposta negativa ad un quesito posto da Flajolet in [16] circa l'esistenza di linguaggi context-free a crescita intermedia.

Intendiamo chiudere questa sezione presentando un risultato dimostrato da T. Ceccherini e W. Woess in [7]. Se L è un linguaggio sull'alfabeto A, il *tasso di crescita* di L è il numero

$$\gamma(L) = \limsup_{n\to\infty} \ \mathrm{Card}(\{w \in L \ \mid \ |w| = n\})^{\frac{1}{n}}.$$

Se L è infinito allora $\gamma(L)$ è tale che $0 \leq \gamma(L) \leq \mathrm{Card}(A)$. Sia F un insieme di parole tale che $F \subseteq F(L)$ e sia

$$L_F = \{u \in L \mid \ F(u) \cap F = \emptyset\}.$$

L'insieme F è detto delle *parole proibite* e, nel caso in cui $\gamma(L_F) < \gamma(L)$, diremo che il tasso di crescita di L è sensibile rispetto ad F. Il problema studiato concerne le condizioni che assicurano che $\gamma(L_F) < \gamma(L)$. Converrà osservare che questo problema è interessante soltanto nel caso in cui L ha crescita esponenziale, ovvero nel caso in cui $\gamma(L) > 1$. Infatti se $\gamma(L) = 1$ allora $\gamma(L_F) = 1$ oppure $\gamma(L_F) = 0$ ed in tal caso L_F è finito. Il risultato principale ottenuto in [7] concerne i linguaggi context-free ergodici. Sarà conveniente, a tale proposito, presentare questa nozione. Consideriamo una grammatica context-free ridotta in cui, cioè, ogni variabile è utile. Ad essa è possibile associare un grafo orientato definito nel modo seguente: l'insieme dei suoi vertici è l'insieme delle variabili della grammatica mentre, dati due vertici V e W, la coppia (V, W) è un arco del grafo se esiste una produzione $V \to \alpha$ della grammatica tale che W è fattore di α. Una grammatica si dice *ergodica* se il grafo ad essa associato ha almeno un arco ed è fortemente connesso. Un linguaggio si dice poi *ergodico* se esiste una grammatica ergodica che lo genera. Il risultato principale dimostrato in [7] è il seguente: *ogni linguaggio context-free, non ambiguo, non lineare ed ergodico è tale che il suo il tasso di crescita è sensibile rispetto ad un qualsiasi insieme di parole.*

3.1 – Un algoritmo di decisione

In questo paragrafo presenteremo un algoritmo che permette di decidere se la funzione di crescita di un linguaggio context-free è esponenziale o polinomiale. In virtù dei due teoremi precedenti, si tratta di dimostrare la ricorsiva decidibilità della proprietà di essere bounded rispetto alla famiglia dei linguaggi context-free. Questo problema è stato affrontato e risolto da Ginsburg in [17]; descriveremo pertanto l'algoritmo da lui proposto. A tale proposito, converrà ricordare che, dato un linguaggio context-free, esso può essere sempre generato da una grammatica di tipo opportuno, detta *in forma normale di Chomsky*, nella quale tutte le produzioni sono della forma $A \to BC$ oppure $A \to a$ dove A, B e C sono variabili della grammatica ed a è un suo simbolo terminale. La procedura proposta da Ginsburg è essenzialmente basata sul lemma seguente con il quale si dimostra che, per ogni variabile A di una grammatica context-free data, i linguaggi L_A ed R_A sono context-free.

LEMMA 3. *Sia $G = (V, T, P, S)$ una grammatica context-free in forma normale di Chomsky. Per ogni variabile A di G esistono grammatiche context-free G_A e G'_A tale che $L_A = L(G_A)$ e $R_A = L(G'_A)$.*

TEOREMA 3. *È possibile decidere se la funzione di crescita di un linguaggio context-free è polinomiale o esponenziale.*

DIM. Siano $G = (V, T, P, S)$ una grammatica context-free, n il numero delle variabili di G, $L = L(G)$ il linguaggio da essa generato e g_L la funzione di crescita di L. Dal momento che g_L è polinomiale o esponenziale, è sufficiente verificare se g_L è polinomiale o meno. Inoltre, dal teorema di caratterizzazione dei linguaggi context-free a crescita polinomiale, segue che g_L è polinomiale se e solo se, per ogni variabile A di G, L_A ed R_A sono linguaggi 1-bounded. Quindi è sufficiente costruire una procedura che permetta di decidere se L_A (rispettivamente R_A) sia un linguaggio 1-bounded. Costruiamo allora tale procedura per il linguaggio L_A, essendo simmetrico il procedimento di decisione per il linguaggio R_A. Dal lemma precedente segue che, a partire dalla grammatica G, possiamo costruire una grammatica context-free G_A tale che $L_A = L(G_A)$. Utilizzando un argomento standard è possibile verificare se L_A contenga o meno una parola diversa dalla parola vuota. Se $L_A \subseteq \{1_{T^*}\}$ allora L_A è banalmente un linguaggio 1-bounded e la procedura termina.

Supponiamo, invece, che esista una parola non vuota $u \in L_A$ e sia z la radice primitiva di u. Chiaramente L_A è un linguaggio 1-bounded se e solo se $L_A \subseteq \{z\}^*$, cioè, se e solo se $M \neq \emptyset$ dove $M = L_A \cap (T^* \setminus \{z\}^*)$. Ora, dal momento che L_A è un linguaggio context-free e $T^* \setminus \{z\}^*$ è un linguaggio regolare, in virtù di un risultato classico dei linguaggi context-free, ovvero che l'intersezione di un linguaggio regolare e di uno context-free è un linguaggio context-free (si osservi, a tale proposito, che l'intersezione di due linguaggi context-free non è, in generale, context-free), segue che M è un linguaggio context-free. Poiché il problema di decidere se una grammatica context-free generi l'insieme vuoto è decidibile ed ogni passo relativo alla costruzione del linguaggio M è effettivo, ne segue che, applicando tale procedura ad M, è possibile decidere se $M = \emptyset$ o meno, cioè se L_A è 1-bounded o meno.

È interessante ricordare che una procedura decisionale più semplice di quella presentata nel Teorema 3 è stata proposta in [8] relativamente alla famiglia dei linguaggi context-free lineari non ambigui.

4 – Calcolo della funzione di conteggio di un linguaggio bounded

In [13] D'Alessandro, Intrigila e Varricchio hanno proposto una tecnica che consente, dato un linguaggio bounded context-free, di descrivere, in modo esatto, la sua funzione di conteggio. Più precisamente, tale procedura di calcolo permette, a partire dalla grammatica context-free che genera il linguaggio bounded, di costruire una famiglia finita di polinomi a coefficienti razionali utile per il calcolo della funzione sopraSu. Vale infatti il risultato seguente.

Teorema 4. *Sia L un linguaggio bounded context-free. Esistono un intero non negativo* n_0 *ed un insieme finito di polinomi a cofficienti razionali* $p_0, \ldots, p_{m-1}$ *tali che, per ogni* $n \geq n_0$,

$$f_L(n) = p_l(n),$$

dove l *è tale che*:

$$l \equiv n \mod m.$$

Nel semplice caso del linguaggio $L = \{a^n b^n \mid n \geq 0\}$, la cui funzione di conteggio è stata definita nella Equazione (1), si ha che tale funzione può essere descritta nel senso del teorema precedente, ponendo $p_0(x) = 1$, $p_1(x) = 0$, $m = 2$, $n_0 = 0$. La dimostrazione del Teorema 4 è effettiva e fornisce un algoritmo per la costruzione, a partire da una grammatica che genera il linguaggio, della famiglia di polinomi atti al computo della sua funzione di conteggio. Presenteremo ora alcuni risultati che derivano in modo immediato dal teorema precedente. A tale proposito, sarà opportuno richiamare alcune definizioni relative all'andamento asintotico di funzioni. Date $f, g : \mathbb{N} \longrightarrow \mathbb{N}$ funzioni definite sull'insieme $\mathbb{N}$, scriviamo

$$f \in \mathcal{O}(g) \quad (\text{risp. } f \in \Omega(g))$$

se esiste un intero positivo n_0 ed un numero reale C tali che, per ogni $n \geq n_0$,

$$f(n) \leq Cg(n) \quad (\text{risp. } f(n) \geq Cg(n)).$$

Corollario 1 ([29]). *Dati un linguaggio context-free* L *ed un intero positivo* k, *è decidibile verificare se la funzione di conteggio* f_L *appartiene alla classe* $\mathcal{O}(n^k)$.

Dim. Per il teorema precedente, partendo dal linguaggio L, è possibile costruire la famiglia dei polinomi $\{p_i\}$ che consente di descrivere la funzione f_L, cosicché si potrà porre

$$\delta = \max\{\delta_i\},$$

dove, per ogni polinomio p_i, δ_i è il suo grado. L'asserto deriva allora dal fatto che

$$f_L \in \mathcal{O}(n^k) \iff \delta \leq k.$$

Analogamente al caso precedente, se γ è il minimo grado dei polinomi della famiglia $\{p_i\}$ prima considerata, dal momento che, per ogni $k \geq 0$,

$$f_L \in \Omega(n^k) \iff \gamma \geq k,$$

si otterrà il corollario seguente:

Corollario 2. *Dati un linguaggio context-free L ed un intero positivo k, è decidibile verificare se la funzione di conteggio f_L appartiene alla classe $\Omega(n^k)$.*

Un'altra conseguenza immediata del teorema è data dal risultato seguente simile al Corollario 13 del lavoro [27].

Corollario 3. *Sia ϵ un numero razionale tale che $0 < \epsilon < 1$. Non esistono linguaggi context-free L tali che*

$$f_L \in \mathcal{O}(n^{k-\epsilon}) \ \ e \ \ f_L \in \Omega(n^{k-1+\epsilon}),$$

dove k è un intero non negativo.

In particolare, il Corollario 3 estende un risultato dimostrato in [29] relativo alla non esistenza di linguaggi context-free la cui funzione sia sub-lineare senza essere limitata superiormente da una costante.

La tecnica combinatoria utilizzata per dimostrare il Teorema 4 è non banale per diverse ragioni: prima di tutto, come è ben noto, esistono linguaggi context-free bounded ambigui per cui una tecnica di conteggio legata alla struttura della grammatica si rivela, in questo contesto, di scarsa utilità. Ricordiamo che una grammatica context-free G si dice *ambigua* se qualche parola in $L(G)$ è il prodotto di due alberi di derivazione distinti. Un linguaggio è poi detto *ambiguo* se ogni grammatica context-free che lo genera è ambigua. Un esempio noto di linguaggio context-free ambiguo è il seguente (*cfr.* [24]):

$$L = \{a^n b^m c^m d^n\}_{n,m\geq 0} \cup \{a^m b^n c^n d^m\}_{n,m\geq 0},$$

e, come si vede facilmente, L è bounded.

Il secondo motivo che rende interessante la tecnica dimostrativa del Teorema 4 è che essa permette di ottenere il seguente risultato particolarmente significativo.

Teorema 5. *Sia L un linguaggio context-free bounded. È possibile costruire un linguaggio regolare bounded L' la cui funzione di conteggio coincide con quella di L.*

Recentemente, in [35] è stata data una significativa applicazione del Teorema 5 nello studio della proprietà di context-freeness per alcune famiglie notevoli di parole. Se consideriamo il linguaggio $L = \{a^n b^n \mid n \geq 0\}$, possiamo osservare che il linguaggio $L' = \{ab\}^* = \{(ab)^n \mid n \geq 0\}$ è regolare ed ha, come si vede immediatamente dalla sua definizione, la stessa funzione di conteggio di L. Questo esempio ci consente di fare un'osservazione. In [13], non è stato possibile dimostrare che l'alfabeto del linguaggio L' del Teorema 5 è lo stesso del linguaggio L. Tuttavia, gli esempi considerati dagli autori sembrano indicare che, non soltanto questo è il caso ma, addirittura, il linguaggio L' è *commutativamente equivalente* a L. Ricordiamo che due linguaggi L, L' si dicono commutativamente equivalenti se possiedono lo stesso alfabeto, ovvero $L, L' \subseteq A^*$, ed esiste una biezione

$$\alpha : L \longrightarrow L'$$

di L in L' tale che, per ogni $u \in L$, si abbia:

$$\forall\, a \in A, \quad |u|_a = |\alpha(u)|_a.$$

In questo contesto è allora naturale formulare il problema seguente:

Problema aperto: Dato un linguaggio context-free bounded, è sempre possibile ottenere un linguaggio regolare commutativamente equivalente ad esso?

Concludiamo infine questa sezione con un'ultima considerazione (*cfr.* [2]) di interesse nello studio della proprietà di ambiguità per i linguaggi context-free. Un risultato significativo dimostrato da Flajolet in [16] permette di dimostrare che se la funzione di crescita di un linguaggio context-free è trascendente, ovvero non è soluzione di alcuna equazione algebrica, allora il linguaggio è ambiguo. Questo criterio è dunque di aiuto nello studio dell'ambiguità di un linguaggio. Tuttavia, se il linguaggio è bounded context-free, per il Teorema 5, la sua funzione di conteggio e quella di crescita sono sempre razionali e quindi algebriche. Pertanto il criterio di Flajolet non è di utilità nello studio del problema predetto per i linguaggi bounded.

5 – Linguaggi bounded e problema di Burnside

Come si è detto nell'introduzione di questa nota, i linguaggi bounded giocano un ruolo significativo nello studio di un problema ben noto dell'Algebra: il *problema di Burnside per i semigruppi*. L'argomento è importante e vogliamo, per questa ragione, compiere un rapido *excursus* in questo ambito. Rimandiamo al testo [15] quale eccellente e completo riferimento bibliografico per questo argomento. Un semigruppo S si dice *periodico* o *di torsione* se, per ogni elemento s di S, esistono interi non negativi i, j tali che $i < j$ e $s^i = s^j$. Il semigruppo S si dirà poi *finitamente generato* se esiste un sottoinsieme finito di elementi di

S, detti *generatori*, tale che ogni elemento di S possa scriversi come prodotto di un numero finito di generatori. Come facilmente si verifica, un semigruppo di cardinalità finita è periodico e finitamente generato. È allora naturale porsi la domanda seguente: *È un semigruppo finitamente generato e periodico finito*?

Questo problema fu studiato per la prima volta da W. Burnside nel 1902 nel caso dei gruppi e tale studio fu successivamente esteso al caso dei semigruppi. La risposta al problema di Burnside è in generale negativa. Infatti, nel caso dei semigruppi, nel 1944, Hedlund e Morse fornirono in [21] un esempio di semigruppo infinito, costruito a partire dalla *parola di Thue*, che è 3-generato e periodico. Nel caso dei gruppi, invece, Golod, utilizzando un risultato di Shafarevitch relativo alla non finitezza della dimensione di un' opportuna algebra definita su un campo, nel 1964 in [18], dimostrò l'esistenza di un p-gruppo infinito 3-generato.

Se un semigruppo finitamente generato S è tale che, per ogni s di S, $s^m = s^n$, dove m ed n sono interi fissati tali che $0 \leq m < n$, allora il problema della finitezza di S è detto *problema di Burnside limitato*. Anche in questo caso, esistono semigruppi e gruppi che sono infiniti ed un risultato di grande interesse riguarda il gruppo libero $G(k, n)$ con k generatori nella varietà dei gruppi soddisfacenti alla identità $x^n = 1$. Un contributo di grande rilievo è infatti la dimostrazione data da Adjan e Novikov nel 1968 della infinitezza del gruppo $G(k, n)$ qualora $k > 1$ e n sia dispari con $n \geq 665$ (*cfr.* [1]). Gli autori dimostrarono inoltre che in tale caso il problema della parola è ricorsivamente decidibile, ovvero che esiste un algoritmo l'esecuzione del quale consente, date due parole sull'alfabeto dei generatori, di decidere se le due parole rappresentano il medesimo elemento del gruppo.

Un problema strettamente correlato al problema di Burnside per i semigruppi può essere formulato, nella teoria dei linguaggi formali, al modo seguente. Sia L un linguaggio su di un dato alfabeto A. Come si è detto in precedenza, in virtù del teorema di Myhill e Nerode, L è regolare se e solo se il suo monoide sintattico $M(L)$ è finito. Ricordiamo che il monoide $M(L)$ è il monoide quoziente $A^*/\equiv_L$ ove $\equiv_L$ è la relazione di congruenza definita come segue: per ogni $u, v \in A^*$, $u \equiv_L v$ se e solo se

$$\forall\, f,\ g \in A^*,\ (fug \in L \Longleftrightarrow fvg \in L).$$

La relazione $\equiv_L$ è detta congruenza sintattica associata ad L. Supponiamo ora di aver scelto L per modo che il suo monoide sintattico sia periodico; d'ora in avanti, chiameremo semplicemente *periodici* tali linguaggi. Poiché $M(L)$ è (ovviamente) finitamente generato, nell'ipotesi in cui L sia periodico, ogni condizione che assicura una risposta positiva al problema di Burnside per i semigruppi, implica la finitezza di $M(L)$ e, dunque, la regolarità di L. Lo studio delle condizioni di finitezza per i monoidi sintattici dei linguaggi periodici è chiamato *problema di Burnside per i linguaggi* (*si veda anche* [39]). Tale problema consiste quindi nello studio delle condizioni che assicurano che un linguaggio periodico sia regolare.

Queste condizioni sono anche dette di *regolarità*. Vale la pena di osservare che ogni condizione di regolarità è una condizione di finitezza per il monoide sintattico di un linguaggio periodico e tuttavia esistono monoidi finitamente generati e periodici che non sono monoidi sintattici di nessun linguaggio. Vediamo ora come i linguaggi bounded si inseriscono nel quadro concettuale dei problemi che abbiamo ora succintamente descritto. A tale proposito, converrà ricordare le definizioni di semigruppo permutabile e debolmente permutabile.

DEFINIZIONE 5. Siano M un monoide ed n un intero positivo tale che $n \geq 2$. M è detto n-permutabile se, per ogni sequenza $m_1, \ldots, m_n$ di n elementi di M, esiste una permutazione non banale ρ dell'insieme $\{1, 2, \ldots, n\}$ tale che:

$$m_1 m_2 \cdots m_n = m_{\rho(1)} m_{\rho(2)} \cdots m_{\rho(n)}.$$

Si dice che M è permutabile se esiste un intero $n \geq 2$ tale che M è n-permutabile.

La definizione precedente può essere generalizzata nel modo seguente.

DEFINIZIONE 6. Siano M un monoide ed n un intero positivo tale che $n \geq 2$. M è detto n-debolmente permutabile se, per ogni sequenza $m_1, \ldots, m_n$ di n elementi di M, esistono due permutazioni ϱ e τ dell'insieme $\{1, 2, \ldots, n\}$, $\varrho \neq \tau$, tali che:

$$m_{\varrho(1)} m_{\varrho(2)} \cdots m_{\varrho(n)} = m_{\tau(1)} m_{\tau(2)} \cdots m_{\tau(n)}.$$

Si dice che M è debolmente permutabile se esiste un intero $n \geq 2$ tale che M è n-debolmente permutabile.

La proprietà di permutazione costituisce una generalizzazione di quella commutativa. Nel 1984, A. Restivo e C. Reutenauer nel lavoro [38] introdussero questa proprietà nell'ambito del problema di Burnside dimostrando che essa costituisce una condizione di finitezza per i semigruppi finitamente generati e periodici. Vale infatti il seguente teorema.

TEOREMA 6. *Sia M un monoide finitamente generato. Allora M è finito se e solo se M è periodico e permutabile.*

Alla luce di questo teorema è oppurtuno osservare che, in generale, la proprietà di debole permutazione non è una condizione di finitezza per i semigruppi esistendo, come dimostrato da A. Restivo in [37], monoidi finitamente generati, periodici, debolmente permutabili, che non sono di cardinalità finita. Vogliamo, a questo punto, fornire una dimostrazione molto sintetica del Teorema 6 anche e soprattutto per mettere in luce il ruolo svolto dai linguaggi bounded nello studio del problema di Burnside. A tale proposito è opportuno ricordare alcune definizioni. Ogni monoide finitamente generato M è immagine di un opportuno epimorfismo

$$\psi : A^* \longrightarrow M,$$

dove la cardinalità di A è uguale a quella dell'insieme dei generatori di M. Questo morfismo è detto *epimorfismo canonico* di M. Supponiamo ora che A sia totalmente ordinato. Possiamo ordinare totalmente A^* definendo una nuova relazione di ordine $<_a$, detta di *ordine alfabetico*, come:

$$u <_a v \Longleftrightarrow (|u| < |v|) \text{ o } (|u| = |v| \text{ e } u < v),$$

dove $<$ è l'ordinamento lessicografico. Dalla sua definizione segue che $<_a$ è un buon ordinamento. Una parola v si dice *riducibile* se esiste $u \in A^*$ tale che

$$u <_a v \text{ e } \psi(u) = \psi(v).$$

Una parola che non è riducibile si dice *irriducibile*. Sia ora m un elemento di M. Nell'insieme $\psi^{-1}(m)$ esisterà una sola parola irriducibile che chiameremo *rappresentante canonico di* m. Se L è un sottoinsieme qualsiasi del sostegno di M, indicheremo con il simbolo C_L l'insieme dei rappresentanti canonici degli elementi di L. Ovviamente gli insiemi L e C_L sono equipotenti. Queste definizioni ci consentono di estendere la nozione di funzione di crescita ad un insieme di un qualsiasi monoide finitamente generato. Come prima, sia M un monoide e sia $\psi : A^* \longrightarrow M$ il morfismo canonico di A^* in M dove la cardinalità di A è uguale a quella dell'insieme dei generatori di M. Se $m \in M$ è un generico elemento di M, la *lunghezza di* m è definita come

$$|m| = \min\{n \geq 0 \mid \psi^{-1}(m) \cap A^n \neq \emptyset\}.$$

Ad ogni sottoinsieme $L \subseteq M$, associamo allora la sua *funzione di conteggio* $f_L : \mathbb{N} \to \mathbb{N}$ definita come

$$f_L(n) = \mathrm{Card}(\{m \in L \mid |m| = n\}).$$

La *funzione di crescita di* L $g_L : \mathbb{N} \to \mathbb{N}$ si definisce a partire da quella di conteggio nello stesso modo visto nel monoide libero. Siamo ora in grado di enunciare la seguente importante proposizione (*cfr.* [15], Capitolo 3).

Proposizione 1. *Sia M un monoide permutabile e finitamente generato. Allora l'insieme dei suoi rappresentanti canonici è un linguaggio bounded.*

Possiamo alla luce della Proposizione 1 dimostrare il teorema di Restivo e Reutenauer.

Dim. Per la Proposizione 1, esistono parole $u_1, \ldots, u_k$ tali che

$$C_M \subseteq u_1^* \cdots u_k^*.$$

Ogni elemento c di C_M è rappresentato nella forma

$$u_1^{r_1} \cdots u_k^{r_k}.$$

Dunque si ottiene

$$|c| = \sum_{i=1,\dots,k} r_i |u_i| \leq r \cdot \left(\sum_{i=1,\dots,k} |u_i| \right),$$

dove $r = \max\{r_i \mid i = 1, \dots, k\}$. Sia $m_i = \phi(u_i)$, con $i = 1, \dots, k$. Poiché M è periodico, esiste un intero p tale che, per ogni $i = 1, \dots, k$, $m_i^p = m_i^{q_i}$, con $q_i < p$. Se M è infinito lo è anche C_M. Quindi la lunghezza di c e l'intero r saranno arbitrariamente grandi. Qualora $r > p$, si ha che c è riducibile e ciò contraddice l'aver assunto c quale rappresentante canonico di un elemento di M.

Si è detto prima che un linguaggio è periodico se il suo monoide sintattico è un monoide periodico. Similmente, diremo che un linguaggio è *permutabile* (risp. *debolmente permutabile*) se il suo monoide sintattico è un monoide debolmente permutabile. Il teorema che segue, la cui dimostrazione è essenzialmente basata sul Teorema 6, fornisce una notevole condizione di regolarità per i linguaggi, proposta da A. De Luca e S. Varricchio in [15]:

TEOREMA 7. *Un linguaggio L è regolare se e solo se L è permutabile e periodico.*

I risultati che seguono, proposti in [14] da F. D'Alessandro e S. Varricchio, sono di particolare interesse in quanto forniscono una condizione di regolarità per i linguaggi context-free bounded.

LEMMA 4. *Il monoide sintattico di un linguaggio bounded è permutabile.*

DIM. Siano A^* il monoide libero delle parole su di un dato alfabeto finito A, L un sottoinsieme di A^*, $\equiv_L$ la congruenza sintattica di L ed M il monoide sintattico di L. Poiché L è un linguaggio bounded, esistono parole $u_1, \dots, u_n$ di A^+ tali che $L \subseteq \{u_1\}^* \cdots \{u_n\}^*$. È lecito supporre che le parole $u_1, \dots, u_n$ siano primitive. Sia γ la massima lunghezza di una parola dell'insieme $\{u_1, \dots, u_n\}$. Proviamo che M è k-permutabile per ogni intero k tale che:

$$k - 1 > n(1 + (\gamma + 1)^2) + (n - 1).$$

Dimostriamo che, per ogni sequenza $m_1, \dots, m_k$ di k elementi di M, esiste una permutazione non banale ρ dell'insieme $\{1, \dots, k\}$ tale che $m_1 m_2 \cdots m_k = m_{\rho(1)} m_{\rho(2)} \cdots m_{\rho(k)}$.

Supponiamo che ogni $m_i \neq 1_M$. Per ogni $i = 1, \ldots, k$, sia w_i la parola di A^* che rappresenta m_i, cioè, $m_i = [w_i]_{\equiv_L}$. Ovviamente, per ogni $i = 1, \ldots, k$, w_i è una parola non vuota. Sono possibili i seguenti due casi.

Caso 1. Supponiamo che la parola $w_1 \cdots w_{k-1}$ non sia un fattore di alcuna parola di L, cioè, per ogni $\lambda, \mu \in A^*$, $\lambda w_1 \cdots w_{k-1} \mu \notin L$. Da questo segue che, per ogni $\lambda, \mu \in A^*$, $\lambda w_k w_1 \cdots w_{k-1} \mu \notin L$ e $\lambda w_1 \cdots w_{k-1} w_k \mu \notin L$. Da cui si ha: $w_1 \cdots w_{k-1} w_k \equiv_L w_k w_1 \cdots w_{k-1}$, cioè, $m_1 m_2 \cdots m_{k-1} m_k = m_k m_1 \cdots m_{k-1}$. Il primo caso è pertanto dimostrato.

Caso 2. Supponiamo ora che la condizione del Caso 1 non sia verificata. Dal momento che $k-1 > n(1+(\gamma+1)^2)+(n-1)$, possiamo trovare nella parola $w_1 \cdots w_{k-1}$ una parola della forma:

$$w_i \cdots w_j, \; 1 \leq i < j \leq k-1,$$

tale che

$$j - i \geq 1 + (\gamma+1)^2, \text{ e } \alpha w_i \cdots w_j \beta \in \{u_l\}^+, \tag{4}$$

dove $1 \leq l \leq n$ e $\alpha, \beta \in A^*$. Consideriamo la sequenza di parole:

$$w_i, w_i w_{i+1}, \cdots, w_i \cdots w_{i+s}, \cdots, w_i \cdots w_j.$$

Sia v_l la coniugata di u_l. Dalla condizione (4), ogni parola della sequenza è un prefisso di una parola di $\{v_l\}^+$. Dal momento che $j-1 \geq 1 + (\gamma+1)^2$, e il numero dei prefissi distinti di v_l sono minori o uguali a γ, esistono almeno tre interi i_1, i_2, i_3, dove $1 \leq i_1 < i_2 < i_3 \leq j$, tali che:

$$w_i \cdots w_{i_1}, w_i \cdots w_{i_2}, \cdots, w_i \cdots w_{i_3} \in \{v_l\}^+ p,$$

dove p è un prefisso di v_l. Si ottiene:

$$w_{i_1+1} \cdots w_{i_2}, w_{i_2+1} \cdots w_{i_3} \in \{sp\}^+, v_l = ps,$$

e quindi

$$w_{i_1+1} \cdots w_{i_2} w_{i_2+1} \cdots w_{i_3} = w_{i_2+1} \cdots w_{i_3} w_{i_1+1} \cdots w_{i_2}.$$

Vale allora la seguente identità:

$$\begin{aligned} m_1 \cdots m_k &= m_1 \cdots m_{i_1}(m_{i_1+1} \cdots m_{i_2} m_{i_2+1} \cdots m_{i_3}) m_{i_3+1} \cdots m_k = \\ &= m_1 \cdots m_{i_1}(m_{i_2+1} \cdots m_{i_3} w_{i_1+1} \cdots m_{i_2}) m_{i_3+1} \cdots m_k. \end{aligned}$$

Il Caso 2 è così dimostrato.

Corollario 4. *Sia L un linguaggio context-free a crescita polinomiale. L è periodico se e solo se L è regolare.*

Dim. Sia M il monoide sintattico di L. Dal Teorema 1, L è un linguaggio bounded. Dal lemma precedente, il monoide M è permutabile, cosicché L è permutabile. Poiché, per ipotesi, L è periodico, dal Teorema 7 segue allora che L è regolare. Viceversa, sia L un linguaggio regolare. Per il teorema di Myhill e Nerode, il monoide M è finito e, di conseguenza permutabile e periodico. Il linguaggio L è pertanto permutabile e periodico.

Il Corollario 4 è effettivo poiché (si veda ancora [14]) esiste un algoritmo l'esecuzione del quale consente, a partire da una grammatica context-free che generi un linguaggio bounded, di decidere se il linguaggio è periodico o meno. Mostriamo infine che il corollario è non banale poiché esistono linguaggi context-free che sono periodici ma non regolari. A tale scopo introduciamo la definizione seguente.

Definizione 7. Siano w una parola di A^* e p un intero positivo tale che $p > 1$. w è detta *p-power-free* se, per ogni $u \in A^+$, u^p non è un fattore di w.

Vale il seguente seguente risultato (*cfr.* [5], [34]).

Lemma 5. *Sia f una parola infinita su di un alfabeto A generata da un morfismo $\phi : A^* \to A^*$ e sia* Pref(f) *l'insieme dei suoi prefissi finiti. Allora l'insieme complementare di* Pref(f) *in A^* è un linguaggio context-free.*

Utilizzando il Lemma 5, possiamo costruire un linguaggio context-free periodico non regolare. A tale proposito, siano $A = \{a, b, c\}$ un alfabeto di tre lettere, A^* il monoide libero delle parole sull'alfabeto A e sia f la parola infinita ottenuta iterando il morfismo $\phi : A^* \longrightarrow A^*$ definito come segue:

- $\phi(a) = abc$;
- $\phi(b) = ac$;
- $\phi(c) = b$.

Ad esempio, la parola abcacbabcbacabcacb è il prefisso di lunghezza 18 di f. Siano $F = \text{Pref}(f)$ ed $L = A^* \backslash F$. Per il Lemma 5, L è un linguaggio context-free. È possibile dimostrare che f è una parola 3-power-free, cioè, per ogni $u \in A^*$ e per ogni intero positivo n, con $n \geq 3$, u^n non è un fattore della parola f. Quindi $u^n \in L$. Da questo segue che, per ogni intero positivo n, con $n \geq 3$, e per ogni $u \in A^*$, $[u^n]_{\equiv_L} = [u^{n+1}]_{\equiv_L}$. Dunque il monoide sintattico $A^*/\equiv_L$ di L è periodico, ovvero L è periodico. Dimostriamo infine che L non è regolare. Per assurdo, supponiamo che L lo sia. Poiché la famiglia dei linguaggi regolari è chiusa rispetto alle usuali operazioni di intersezione e di unione insiemistica, F è un linguaggio regolare. Applicando il Pumping Lemma per i linguaggi regolari (*cfr.* Sezione 2.3) ad F, segue che esiste una costante n e una parola w di F, con $|w| \geq n$, tale che $w = \lambda u \mu$ ed inoltre, per ogni $i \geq 0$, $\lambda u^i \mu \in F$. Quindi, per ogni $i \geq 0$, $\lambda u^i \in F$. Di conseguenza f ammetterebbe parole della forma u^3 come suoi fattori, cioè f non sarebbe 3-power-free, e ciò è una contraddizione. Quindi L non è regolare.

6 – Linguaggi a crescita intermedia

In questa sezione ci interesseremo ad un risultato proposto da Grigorchuk e Machì in [19] relativo ai cosiddetti linguaggi di crescita intermedia. Diamone subito la definizione. Un linguaggio L si dice di *crescita intermedia* se la sua funzione di crescita è subesponenziale, ovvero se è definitivamente limitata superiormente da una funzione esponenziale k^n, con $k > 1$ ma non da alcuna funzione polinomiale di grado fissato. Abbiamo visto che un linguaggio context-free ha crescita polinomiale o esponenziale. Non esistono dunque linguaggi context-free a crescita intermedia. In [19], si fornisce un esempio di linguaggio a crescita intermedia la cui struttura combinatoria è, nel senso che preciseremo dopo, vicina a quella di un linguaggio context-free. Questa costruzione presuppone la conoscenza di un modello di calcolo, detto *automa stack*, il quale risulta, rispetto all'operazione di riconoscimento di linguaggi di parole, più potente di quello degli automi push-down, cioè delle macchine in grado di riconoscere i linguaggi context-free. Gli automi stack sono stati introdotti da Ginsburg, Greibach, ed Harrison (*cfr.* [24], Capitolo 14) alla fine degli Anni 60 nello studio di una possibile estensione dei linguaggi context-free. Non si ha intenzione di presentare in modo rigoroso la loro definizione ma, semplicemente, di fornire alcuni elementi descrittivi essenziali di questa struttura, al fine di illustrare l'esempio predetto. Essenzialmente, un automa stack è un automa push-down che, oltre a poter manipolare la pila con le stesse modalità di un automa push-down, può, in un qualsiasi passo della computazione, accedere, leggere e modificare il simbolo memorizzato in una locazione qualsiasi della pila, senza essere costretto (come accade negli automi push-down) a rimuovere tutti i caratteri memorizzati nelle locazioni che lo precedono. Quest' ultima modalità di accesso alla pila sarà detta *in modalità stack*. Gli automi stack possono essere deterministici, non deterministici, ad una via ("one way"), ovvero con la possibilità di leggere la parola di ingresso una sola volta (da sinistra verso destra) oppure a due vie ("two way"), ovvero con la possibilità di leggere, attraverso una oppurtuna testina di lettura, più volte la parola di input su di un apposito nastro, percorribile in entrambe le direzioni. In questo capitolo, siamo interessati ad automi 1-DNESA ("One-way, Deterministic, Non Erasing, Stack Automata"), cioè ad automi stack, deterministici, ad una via, in grado di leggere, ma non cancellare, il simbolo memorizzato in una locazione della pila.

Ricordiamo che se n un intero positivo, la sequenza di t interi positivi, $(n_1, n_2, \ldots, n_t)$ tali che $n_1 \geq n_2 \geq \ldots \geq n_t$, costituisce una partizione dell'intero

n se:

$$n = n_1 + n_2 + \cdots + n_t.$$

Indicando con $P(n)$ il numero di partizioni dell'intero n, si ha asintoticamente:

$$P(n) \approx \frac{e^{\pi\sqrt{\frac{2n}{3}}}}{4n\sqrt{3}}.$$

Teorema 8. *Sia $L \subseteq \{a,b\}^*$ il linguaggio sull'alfabeto $\{a,b\}$ costituito dall'insieme di tutte e sole le parole*:

$$ab^{i_1}ab^{i_2}\cdots ab^{i_k},$$

dove $0 \leq i_1 \leq i_2 \leq \ldots \leq i_k$ è una sequenza monotona non decrescente di $k \geq 1$ interi non negativi. Allora L è un linguaggio a crescita intermedia ed è accettato da un automa 1-DNESA.

Dim. Verifichiamo, prima di tutto, che L è un linguaggio a crescita intermedia. Se $n \in \mathbb{N}$ e $(l_1, l_2, \cdots, l_k)$ è una partizione di n, possiamo ad essa associare in modo univoco la parola di L $ab^{l_1}ab^{l_2}\ldots ab^{l_k}$. Di conseguenza, per ogni n, il numero delle parole di L di lunghezza n è uguale al numero delle partizioni distinte di n. Quindi, la funzione di struttura di L è subesponenziale ed in base alla definizione di g_L, un semplice argomento di conteggio permette di mostrare che anche g_L è subesponenziale. Quindi L ha crescita intermedia. Per il teorema di gap, L non è un linguaggio context-free. Mostriamo ora che L è accettato da un automa 1-DNESA. La struttura ed il funzionamento di questo automa possono essere descritti nel modo seguente. Se una parola w è una stringa (eventualmente vuota) di sole occorrenze del simbolo a, allora la parola w è accettata. Se la parola w comincia con una stringa di occorrenze del simbolo a seguita da una stringa s_1 di occorrenze del simbolo b allora tali occorrenze vengono inserite nella pila. A questo punto della computazione, i casi possibili sono i seguenti:

- Se, dopo aver letto la stringa s_1 di occorrenze del simbolo b, la lettura sul nastro d'ingresso termina, la parola è accettata;
- Supponiamo, invece, che la parola w sia della forma:

$$w = aa\cdots a\underbrace{bb\cdots b}_{s_1}a\underbrace{bb\cdots b}_{s_2}w',\ w' \in \{a,b\}^*$$

  Perché la parola w venga accettata deve necessariamente essere, per definizione del linguaggio L, $|s_1| \leq |s_2|$. Usando la testina della pila, l'automa, leggendo la pila in modalità stack, è in grado di confrontare il numero di b presenti nella pila, cioè $|s_1|$, con il numero di b nella stringa s_2, cioè $|s_2|$.

Più precisamente, leggendo in modo sincronizzato i simboli di s_1 ed s_2, uno per volta, se l'operazione di lettura del fattore s_2 (sul nastro d'ingresso) termina prima di quella di s_1 (nella pila), allora ciò significa che $|s_1| > |s_2|$. La computazione termina e la parola è rifiutata. Se ciò non accade, allora $|s_1| \leq |s_2|$. A questo punto, le eventuali occorrenze rimanenti del simbolo b vengono inserite nella pila e la computazione, nel caso in cui la parola w' sia non vuota, prosegue secondo la modalità prima descritta.

La descrizione sintetica ora data è quella di un modello di calcolo deterministico "one-way". Infine, poiché durante la computazione, l'operazione di cancellazione di simboli dalla pila non è mai effettuata, l'automa è "non erasing". Abbiamo quindi fornito la descrizione di un automa 1-DNESA che è in grado di accettare il linguaggio considerato.

7 – Sviluppi recenti

Chiudiamo questo articolo descrivendo molto sinteticamente il contenuto di un lavoro recente [10], relativo alla estensione dei teoremi presentati nelle Sezioni 3 e 4, al caso dei sottoinsiemi razionali dei monoidi di relazioni di parole. Rimandiamo il lettore alla consultazione dei testi [3] e [40], referenze ormai classiche sull'argomento, per una piana ed efficace introduzione ai concetti fondamentali della tematica delle relazioni razionali. Qui, ci limiteremo a richiamare un vocabolario minimo di concetti al solo fine di presentare i risultati a cui abbiamo accennato. Sia $M = A_1^* \times \cdots \times A_k^*$ il prodotto diretto di monoidi liberi generati da alfabeti $A_1, \ldots, A_k$. I sottoinsiemi di M sono chiamati k-relazioni.

Una k-relazione si dice *razionale* se si ottiene, a partire da relazioni finite, tramite l'applicazione, in un numero finito di volte, delle operazioni razionali di M, ovvero delle operazioni di unione insiemistica e di prodotto di due relazioni e della operazione di stella che associa ad ogni relazione il sottomonoide di M da essa generato. In virtù di un ben noto teorema di caratterizzazione, una relazione è k-razionale se i suoi elementi, cioè k-uple di parole su alfabeti fissati, sono accettati da uno specifico modello di calcolo detto *automa a k nastri*. Un automa a k nastri è in essenza un automa a stati finiti, non deterministico, dotato di k nastri, ognuno dei quali in grado di memorizzare una parola data. Al generico istante di computazione, l'automa è in grado di leggere, e di elaborare, uno per volta, i caratteri di una qualsiasi delle k parole registrate sui nastri. La k-upla è poi accettata se, una volta completata la lettura di tutte le k parole, l'automa si trovi in uno stato scelto nell'ambito di un insieme particolare di stati detti *accettanti*. Un automa siffatto costituisce una estensione del modello classico di automa a stati finiti. Come si è visto nella Sezione 5, i concetti di funzione di conteggio, di funzione di crescita (e la relativa classificazione dei linguaggi vista nella Definizione 1) e di insieme bounded possono essere definiti in un monoide finitamente generato qualsiasi in modo simile a quanto visto nei monoidi di parole. Vale allora il teorema seguente.

TEOREMA 9 ([10]). *Sia* $M = A_1^* \times \cdots A_k^*$ *il prodotto diretto di monoidi liberi generati da alfabeti* $A_1, \ldots, A_k$. *Valgono le condizioni seguenti*:

- *Una* k*-relazione razionale di* M *ha crescita esponenziale oppure polinomiale.*
- *Una* k*-relazione razionale di* M *ha crescita polinomiale se e solo se è bounded in* M.
- *È possibile decidere se una* k*-relazione razionale ha crescita esponenziale oppure polinomiale.*

È infine interessante ricordare che in [10] il teorema precedente è stato dimostrato nel caso più generale dei monoidi parzialmente commutativi di cui il prodotto di monoidi liberi costituisce una istanza particolare.

Ringraziamenti

Un "grand merci" da parte degli autori a Arturo Carpi e Tullio Ceccherini-Silberstein per le interessanti discussioni ed i suggerimenti numerosi che hanno preceduto e accompagnato la redazione di questa nota.

REFERENCES

[1] S. I. ADJAN: *The Burnside problem and identities in groups*, Springer, Berlin, 1985.

[2] A. BERTONI: *comunicazione privata agli autori*, 2006.

[3] J. BERSTEL: *Transduction and Context-Free Languages*, Teubner, Stuttgart, 1979.

[4] J. BERSTEL – D. PERRIN: *Theory of codes*, Academic Press, New York, 1985.

[5] J. BERSTEL: *Every iterated morphism yields a co-CFL, Information Processing Letters*, **22** (1986), 7-9.

[6] M. R. BRIDSON – R. H. GILMAN: *Context-free languages of sub-exponential growth*, J. Comput. System Sci., **64** (2) (1999), 308-310.

[7] T. CECCHERINI-SILBERSTEIN – W. WOESS: *Growth and ergodicity of context-free languages*, Trans. Amer. Math. Soc., **354** (11) (2002), 4597-4625.

[8] T. CECCHERINI-SILBERSTEIN: *On the growth of linear languages*, Advances in Appl. Math., **35** (2005), 243-253.

[9] C. CHOFFRUT – F. D'ALESSANDRO – S. VARRICCHIO: *On the separability of sparse context-free languages and of bounded rational relations*, Theoret. Comput. Sci., **381** (1-3) (2007), 274-279.

[10] C. CHOFFRUT – F. D'ALESSANDRO – S. VARRICCHIO: *On bounded rational trace languages*, in corso di pubblicazione su Theory of Computing Systems.

[11] N. Chomsky – M. P. Schützenberger: *The algebraic theory of context-free languages*, in: "Computer Programming and Formal System", P. Braffort, D. Hirschberg (eds.), North-Holland, Amsterdam,1963, 118-161.

[12] F. D'Alessandro – A. de Luca: *Teoria degli automi*, in preparazione.

[13] F. D'Alessandro – B. Intrigila – S. Varricchio: *On the structure of the counting function of context-free languages*, Theoret. Comput. Sci., **356** (2006), 104-117.

[14] F. D'Alessandro – S. Varricchio: *On the growth of context-free languages*, to appear on Journal of Automata, Languages and Combinatorics, **13** (2) (2008).

[15] A. de Luca – S. Varricchio: *Finiteness and Regularity in Semigroups and Formal Languages*, Springer, Berlin, 1999.

[16] P. Flajolet: *Analytic models and ambiguity of context-free languages*, Theoret. Comput. Sci., **49** (1987), 283-309.

[17] S. Ginsburg: *The mathematical theory of context-free languages*, McGraw-Hill Book Co., 1966.

[18] E. S. Golod: *On nil-algebras and finitely approximable p-groups*, Izv. Akad. Nauk. SSSR, Ser. Mat., **28** (1964), 273-276.

[19] R. I. Grigorchuk – A. Machì: *An example of an indexed language of intermediate growth*, Theoret. Comput. Sci., **215** (1999), 325-327.

[20] M. Gromov: *Groups of polynomial growth and expanding maps*, Inst. Hautes Études Sci., Publ. Math., **53** (1981), 53-73.

[21] G. Hedlund – M. Morse: *Unending chess, symbolic dynamics and a problem in semigroups*, Duke Math. J., **11** (1944), 1-7.

[22] J. Honkala: *Decision problems concerning thinness and slenderness of formal languages*, Acta Informatica, **35** (1998), 625-636.

[23] J. Honkala: *On Parikh slender context-free languages*, Theoret. Comput. Sci., **255** (2001), 667-677.

[24] J. Hopcroft – J. Ullman: *Introduction to Automata Theory, Languages and Computation*, Addison-Wesley Pub. Co., 1979.

[25] O. Ibarra – B. Ravikumar: *On sparseness, ambiguity and other decision problems for acceptors and transducers*, Lecture Notes in Computer Science, **210** Springer-Verlag, Berlin, 1986, 171-179.

[26] L. Ilie: *On a conjecture about slender context-free languages*, Theoret. Comput. Sci., **132**, (1994), 427-434.

[27] L. Ilie – G. Rozenberg – A. Salomaa: *A characterization of poly-slender context-free languages*, Theoretical Informatics and Applications, **34** (2000), 77-86.

[28] L. Ilie: *On length of words in context-free languages*, Theoret. Comput. Sci., **242** (2000), 327-359.

[29] L. Ilie: *On generalized slenderness of context-free languages*, in: "Words, semigroups, and transductions", World Sci. Publishing, River Edge, NJ, 2001, 189-202.

[30] R. Incitti: *The growth function of context-free languages*, Theoret. Comput. Sci., **255** (2001), 601-605.

[31] G. Krause – T. H. Lenagan: *Growth of Algebras and Gelfand-Kirillov Dimension*, Research Notes in Math. **116** Pitman, London, 1985.

[32] M. Latteux – G. Thierrin: *On bounded context-free languages*, Elektron. Informationsverarb. Kybernet., **20** (1984), 3-8.

[33] Lothaire: *Combinatorics on words*, World Science Division, Addison-Wesley, Reading, MA, 1983.

[34] M. G. Main – W. Bucher – D. Haussler: *Applications of an infinite co-CFL*, Proc. 12th ICALP 1985, Lecture Notes in Computer Science, Vol. 194, 1985, 404-412.

[35] N. Rampersad: *On the context-freeness of the set of words containing overlaps*, Information Processing Letters, **102** (2007), 74-78.

[36] D. Raz: *Length considerations in context-free languages*, Theoret. Comput. Sci., **183** (1997), 21-32.

[37] A. Restivo: *Permutation properties and the Fibonacci Semigroup*, Semigroup Forum, **38** (1989), 337-345.

[38] A. Restivo – C. Reutenauer: *On the Burnside problem for semigroups*, J. of Algebra, **89** (1984), 102-104.

[39] A. Restivo – C. Reutenauer: *Rational languages and the Burnside problem*, Theoret. Comput. Sci., **40** (1985), 13-30.

[40] J. Sakarovitch: *Éléments de théorie des automates*, Vuibert, Paris, 2003.

Lavoro pervenuto alla redazione il 25 giugno 2008
ed accettato per la pubblicazione il 30 giugno 2008.
Bozze licenziate il 30 settembre 2008

Indirizzo dell'Autore:

Flavio D'Alessandro – Alessandra Zinno Pilo – Dipartimento di Matematica – Università di Roma "La Sapienza" – Piazzale Aldo Moro 2, 00185 Roma – Italy.
E-mail: dalessan@mat.uniroma1.it
http: www.mat.uniroma1.it/people/dalessandro