

CONTENTS

H.S.M. Coxeter, 1907-, Dedication	1
T.-Y. LEE and B. REMILLARD Occupation times in systems of null recurrent Markov processes	2
K.-Q. LIU The quantum Witt algebra and quantizations of some Witt-modules	7
S.D. MORGERA On noisy pattern matching under geometrical constraints	13
V. DRENSKY Tame primitivity for free nilpotent algebras	19
P.R. HEATH Nielsen type numbers for fibre preserving maps, coincidences and periodic points	25
B. SZYSZKOWICZ $L_p$ -functionals of weighted partial sum processes	31
S. FENG Large deviations for Markov processes with mean field interaction and unbounded jumps	37
Y. WU Extinction of the multilevel $M(M(\mathbb{R}^d))$ -valued branching diffusion process	43
S.A. HASSANI Construction d'un exemple d'anneau de Jaffard local factoriel non noetherienne de dimension 2 et de caractéristique 0	49
N. TERAJ and H. OSADA The Diophantine equation $x^4 - dy^4 = zp$	55
Mailing Addresses	59

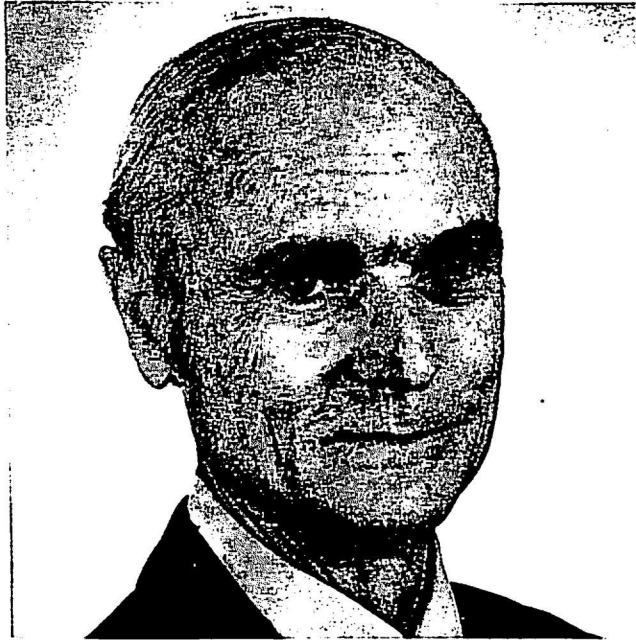
This issue

COMPTES RENDUS MATHÉMATIQUES  
MATHEMATICAL REPORTS

Vol. XIV, No. 1, February 1992

is dedicated by his admirers, colleagues,  
disciples, friends, readers, students and  
well-wishers to

Professor Harold Scott Macdonald Coxeter  
B.A., Ph.D., D.Sc., LL.D., D.Math., F.R.S., F.R.S.C.  
1907-  
Emeritus Professor of Mathematics, University of Toronto



Academician, artist, bestselling writer, dean of geometers, first editor of the *Canadian Journal of Mathematics* 1948-57, friend to all, geometer of purest ray serene, lecturer of spellbinding magnetism, mathematician of  $n$ -dimensional insight, power, range and strength, musician, President of the Canadian Mathematical Society 1965-67, President of the International Mathematical Congress, 1974, teacher of worldwide fame and many-sided scholar,

On the occasion of his retirement from the editorial board, and with grateful thanks for his valuable contributions as author, editor, and inspiration during its formative period 1979-1991.

# OCCUPATION TIMES IN SYSTEMS OF NULL RECURRENT MARKOV PROCESSES

Tzong-Yow Lee<sup>1</sup> and Bruno Remillard<sup>2</sup>  
University of Maryland, College Park  
and Université du Québec, Trois-Rivières

*Presented by Donald A. Dawson, F.R.S.C.*

## Abstract

We study asymptotic properties of differences of occupation times for infinite systems of noninteracting Markovian particles. Under a suitable normalisation we prove convergence in law to a nondegenerate Gaussian field. We also obtain large deviations properties.

## 1 Introduction

In this article we will consider infinite particle systems of noninteracting Markov chains i.e. each particle executes a Markov chain independently of all other particles. We restrict our study to Markov chains on a countable state space  $X$  satisfying the following assumptions :

(A1) The Markov chain is irreducible and null recurrent.

Denote by  $\{\pi_{xy}, x, y \in X\}$  the transition probability matrix. Assumption (A1) ensures existence and uniqueness of a positive function  $\alpha(\cdot)$  on  $X$  such that

$$\sum_{x \in X} \alpha(x) \pi_{xy} = \alpha(y) \text{ and } \alpha(0) = 1$$

where  $0 \in X$  is fixed throughout this article.

<sup>1</sup>Supported in part by the Office of Graduate Studies and Research (1990), University of Maryland and by NSF Grant No. DMS 9106562.

<sup>2</sup>Supported in part by the Fonds Institutionnel de Recherche, Université du Québec à Trois-Rivières and by the Natural Sciences and Engineering Research Council of Canada, Grant No. OGP0042137.

Key words and phrases : Large deviations, Occupation times, Infinite particle systems, Markov chains.

AMS 1985 subject classifications: Primary 60F10; secondary 60J05.

(A2) The resolvent  $G_s(x, y)$  defined as

$$G_s(x, y) = \sum_{n=0}^{\infty} e^{-sn} \pi_{xy}^{(n)} / \alpha(y), \quad s > 0$$

satisfies for each  $y \in X$

$$\sup_{x \in X} |G_s(x, y) - G_s(x, 0)| < \infty \text{ uniformly as } s \rightarrow 0$$

(A3) The resolvent admits the representation

$$G_s(x, y) = \Gamma(1 + \beta) s^{-\beta} l(s^{-1}) + U(x, y) + \epsilon(x, y, s)$$

where  $l: (0, \infty) \rightarrow (0, \infty)$  is slowly varying at infinity,  $0 \leq \beta < 1$ , and (i), (ii) hold, where

(i)

$$\|V\|^2 = 2 \cdot \sum_{x, y \in X} \alpha(x)\alpha(y)V(x)V(y)U(x, y) - \sum_{x \in X} \alpha(x)V(x)^2 > 0$$

for all  $V \neq 0$ ,  $V \in \mathcal{V}_0 = \{W: X \rightarrow \mathbb{R}, W \text{ has finite support and } \sum_x \alpha(x)W(x) = 0\}$

(ii) There exists a function  $\rho: X \rightarrow [1, \infty)$  such that for every  $y \in X$ ,

$$\sup_x U(x, y) / \rho(x) < \infty \text{ and } \lim_{s \downarrow 0} \sup_x \epsilon(x, y, s) / \rho(x) = 0$$

(A4) There exists a constant  $\eta$  such that

$$\sup_{x \in X} G_s(x, x) \leq \eta s^{-\beta} l(s^{-1}), \quad \forall s \in (0, 1]$$

**Remark 1** Assumptions (A2) and (A3) are discrete time analogues of Kasahara's assumptions (see Kasahara (1977) for more details). By adding (A4) we are able to strengthen his convergence in distribution result to convergence of the moment generating function in a neighborhood of 0. The later is essential to prove our large deviations principles for particle systems. It is easy to prove that (A1) to (A4) are satisfied by random walks in the domain of attraction of a stable law of index  $\alpha$  ( $\alpha \in [1, 2]$  in dimension 1 and  $\alpha = 2$  in dimension 2).

Let us now go back to our infinite particle system. We suppose that the number of particles at time zero at site  $x$  is a Poisson random variable with mean  $\alpha(x)$  and is independent of all other sites. It can be checked that the process thus constructed is an ergodic, stationary Markov process. Let  $P$  denote this process,  $E$  denote the corresponding expectation and  $(\omega_j(n), j = 1, 2, \dots, n = 0, 1, \dots)$  denote a labeling of the sample trajectories. Let  $1_z$  denote the index function of the site  $z \in X$ . Further let

$$\xi_N(x) = \sum_{j=1}^{\infty} \sum_{n=0}^{N-1} 1_x(\omega_j(n)) \quad x \in X$$

and define  $\delta_N(x) = (\xi_N(x) - \alpha(x)\xi_N(0))/N$ . We expect that the normalized differences of occupation times  $\{\delta_N(x), x \in X \setminus \{0\}\}$  converge in distribution to a Gaussian field of mean zero and nondegenerate covariance.

This implies that  $\log P \{\delta_N(x) > a\}$  tends to a negative number depending on  $a$ , as  $N \rightarrow \infty$ . On the other hand, Lee (1989) proved a large deviation property for  $\{\xi_N(x)/N, x \in X\}$ . Using his results we can prove that

$$\lim_{N \rightarrow \infty} \left( N^{1-\beta}/l(N) \right)^{-1} \log P \{ \delta_N(x) > aN^{1/2} \} = -\infty$$

for  $a > 0$ .

Our objective is to calculate the leading term of the asymptotic probabilities

$$\log P \{ \delta_N(x) > a\lambda(N) \}, \quad a > 0,$$

for  $\lambda(N) \equiv 1$  and for  $\lambda(N) \rightarrow \infty$ . In this paper we present the case when  $\lambda(N)$  grows no faster than  $(N^{1-\beta}/l(N))^{1/2}$ . More precisely we are interested in the following cases :

$$\left\{ \begin{array}{l} \lambda(N) \equiv 1, \\ \lambda(N)/(N^{1-\beta}/l(N))^{1/2} \rightarrow 0 \text{ as } N \rightarrow \infty, \\ \lambda(N) = (N^{1-\beta}/l(N))^{1/2} \end{array} \right.$$

For a simple random walk on  $Z$ , Cox and Durrett (1990) calculated the leading term in the third case, i.e.  $\lambda(N) = N^{1/4}$  and Remillard (1990) extended their results to its function space generality as Lee (1989) did.

For  $\gamma : X \setminus \{0\} \rightarrow \mathbb{R}$ , i.e.  $\gamma \in \mathbb{R}^{X \setminus \{0\}}$ , let

$$\|\gamma\|_0 = \sup_{\|V\|=1, V \in \mathcal{V}_0} \sum_{x \in X} \gamma(x)V(x) \in [0, \infty]$$

It is easy to see that if  $\gamma$  has finite support, then  $\|\gamma\|_0 < \infty$ .

## 2 Asymptotic properties of occupation times

We are now in a position to state our main results, all of which assume (A1) to (A4).

**Theorem 1** *If  $\beta > 0$  and  $V \in \mathcal{V}_0$ , then*

$$\lim_{N \rightarrow \infty} E \left\{ \exp \sum_{x \in X} V(x)\delta_N(x) \right\} = \exp (\|V\|^2/2)$$

*This implies that  $\delta_N := \{\delta_N(x), x \in X \setminus \{0\}\}$  converges in distribution to a nondegenerate Gaussian field  $\delta := \{\delta(x), x \in X \setminus \{0\}\}$  with mean zero and covariance*

$$\text{Cov}(\delta(x), \delta(y)) = \langle 1_x - \alpha(x)1_0, 1_y - \alpha(y)1_0 \rangle$$

where  $\langle \cdot, \cdot \rangle$  denotes the scalar product over  $\mathcal{V}_0$  defined by

$$\langle V_1, V_2 \rangle = (\|V_1 + V_2\|^2 - \|V_1 - V_2\|^2) / 4, \quad V_1, V_2 \in \mathcal{V}_0$$

**Theorem 2** Suppose that  $\beta > 0$  and  $\lambda(N)$  tends to infinity in such a way that

$$\lim_{N \rightarrow \infty} \frac{\lambda(N)}{[N^{1-\beta}/l(N)]^{1/2}} = 0$$

Let  $C$  be a closed subset and let  $G$  be an open subset of the function space  $\mathbb{R}^{\mathcal{X} \setminus \{0\}}$  equipped with the product topology. Then

$$\limsup_{N \rightarrow \infty} \lambda(N)^{-2} \log P \{ \delta_N \in \lambda(N)C \} \leq - \inf_{\gamma \in C} \|\gamma\|_*^2 / 2$$

$$\liminf_{N \rightarrow \infty} \lambda(N)^{-2} \log P \{ \delta_N \in \lambda(N)G \} \geq - \inf_{\gamma \in G} \|\gamma\|_*^2 / 2$$

where  $\lambda(N)A = \{ \sigma \lambda(N); \sigma \in A \}$ ,  $A \subset \mathbb{R}^{\mathcal{X} \setminus \{0\}}$ .

**Remark 2** It follows easily from the definition of  $\|\cdot\|_*$  that the mapping  $\gamma \mapsto \|\gamma\|_*$  is lower semicontinuous under the product topology of  $\mathbb{R}^{\mathcal{X} \setminus \{0\}}$ .

Denote by  $f_\beta$  the moment generating function of the Mittag-Leffler distribution of index  $\beta \in (0, 1)$

$$f_\beta(a) = \begin{cases} \sum_{k=0}^{\infty} a^k / \Gamma(1+k\beta) & \text{for } a \in \mathbb{R} \text{ when } \beta \in (0, 1) \\ 1/(1-a) & \text{for } a < 1 \text{ when } \beta = 0 \\ +\infty & \text{for } a \geq 1 \text{ when } \beta = 0 \end{cases}$$

and let

$$A_\beta(a) = \frac{\sin(\beta\pi)}{\beta\pi} \int_0^1 b^{-\beta} (-1 + f_\beta(\Gamma(1+\beta)(1-b)^\beta a^2/2)) db$$

$a \in \mathbb{R}$ , when  $\beta \in (0, 1)$ , and

$$A_0(a) = \begin{cases} -1 + 1/(1-a^2/2) & \text{if } |a| < \sqrt{2} \\ +\infty & \text{if } |a| \geq \sqrt{2} \end{cases}$$

Further let

$$B_\beta(b) = \sup_{a \in \mathbb{R}} (ba - A_\beta(a))$$

It is readily checked that both functions  $A_\beta$  and  $B_\beta$  are even, strictly convex, differentiable (when finite), and are equal to zero when the argument is equal to zero.

**Theorem 3** Suppose that  $\beta > 0$ . If  $C$  is a closed subset and  $G$  is an open subset of  $\mathbb{R}^{\mathcal{X} \setminus \{0\}}$  under product topology, then

$$\limsup_{N \rightarrow \infty} (N^{1-\beta}/l(N))^{-1} \log P \left\{ \delta_N \in (N^{1-\beta}/l(N))^{1/2} C \right\} \leq - \inf_{\gamma \in C} B_\beta(\|\gamma\|_*)$$

$$\liminf_{N \rightarrow \infty} (N^{1-\beta}/l(N))^{-1} \log P \left\{ \delta_N \in (N^{1-\beta}/l(N))^{1/2} G \right\} \geq - \inf_{\gamma \in G} B_\beta(\|\gamma\|_*)$$

In case  $\beta = 0$  (see assumption (A3)), we need to add one more assumption. For  $f : \mathcal{X} \mapsto \mathbb{R}$  satisfying  $\sum_{y \in \mathcal{X}} \pi_{xy} |f(y)| < \infty$  for all  $x \in \mathcal{X}$ , we define

$$(\pi f)(x) = \sum_{y \in \mathcal{X}} \pi_{xy} f(y)$$

(A5) The number  $c_\pi$  defined by

$$c_\pi := \sup_{V \in \mathcal{V}_0} \sup_{x \in X} \frac{\log(\pi(\exp f_x))(x)}{(\pi f_x^2)(x)}$$

is finite, where  $f_x(\cdot) = GV(\cdot) - (\pi GV)(x)$  and  $GV(x) = \sum_y U(x, y)V(y)\alpha(y)$ .

**Theorem 4** *Suppose that  $\beta = 0$ . If in addition (A5) is fulfilled, then Theorem 1 and Theorem 2 hold. Moreover there exists a number  $\bar{\gamma}$  such that Theorem 3 holds if the function space  $\mathbb{R}^X \setminus \{0\}$  is replaced by the subspace  $\{\gamma; \|\gamma\|_\infty < \bar{\gamma}\}$ ; we can choose  $\bar{\gamma} = A'_0(1/2(c_\pi\eta))^{1/2}$ .*

**Remark 3** *It is readily checked that (A5) holds when  $\inf\{\pi_{xy}; \pi_{xy} > 0\} > 0$ . It is easy to see that Theorem 4 applies (with  $c_\pi = 1/2$ ) to the simple random walk in two dimensions.*

**Conjecture 1** *Theorem 4 holds with  $\bar{\gamma} = +\infty$ .*

These theorems are proved by estimating the cumulant generating function of  $\{\delta_N(x), x \in X \setminus \{0\}\}$  as  $N \rightarrow \infty$ .

## References

- [1] Cox, J.T. and Durrett, R. (1990) Large deviations for independent random walks. *Prob. Theory Rel. Fields* **84**, 67-82.
- [2] Dawson, D. and Gärtner, J. (1987) Large deviations for McKean-Vlasov limit of weakly interacting diffusions. *Stochastics* **20**, 247-308.
- [3] Feller, W. (1971) *An introduction to probability theory and its applications*, Vol. 2. John Wiley & Sons, New York.
- [4] Kasahara, Y. (1977) Limit theorems of occupation times for Markov processes. *Publ. RIMS. Kyoto Univ.* **12**, 801-818.
- [5] Lee, T.-Y. (1989) Large deviations for systems of noninteracting recurrent particles. *Ann. Probab.* **17**, 46-57.
- [6] Remillard, B. (1990) Asymptotic behaviour of the Laplace transform of weighted occupation times of random walks and applications. *Diffusion Processes and Related Problems in Analysis*, Volume 1, Mark A. Pinski ed. *Progress in Probability*, **22**, Birkhäuser, Boston, 497-519.

Tzong-Yow Lee  
 Department of Mathematics  
 University of Maryland  
 College Park MD 20742  
 U.S.A.  
 e-mail : tyl@hilda.umd.edu

Bruno Remillard  
 Département de mathématiques  
 Université du Québec à Trois-Rivières  
 Trois-Rivières, Qc  
 Canada G1K 7P4  
 e-mail : bruno.remillard@uqtr.quebec.ca

---

Received February 27, 1992

# THE QUANTUM WITT ALGEBRA AND QUANTIZATIONS OF SOME WITT-MODULES

KE-QIN LIU

*Presented by R.V.Moody, F.R.S.C.*

**Abstract.** The quantum universal central extension of the quantum Witt algebra at roots of unity is determined, two characterizations are given for the quantum Witt algebra, a  $q$ -analogue of the enveloping algebra of the Virasoro algebra is introduced and quantizations of the module of tensor fields over the Witt algebra are obtained.

**§0. Notations and Definitions.** Throughout the paper, all vector spaces are vector spaces over the complex number field  $\mathbb{C}$  and the following notations are used:

$$\mathbb{C}^* := \{x \in \mathbb{C} \mid x \neq 0\};$$

$q$  is a complex number satisfying  $q^2 \neq 0, 1$ ;

If  $q$  is a  $t$ -th primitive root of unity, then

$$T := \begin{cases} \alpha(q), & \text{if } \alpha(q) \text{ is odd,} \\ \frac{\alpha(q)}{2}, & \text{if } \alpha(q) \text{ is even;} \end{cases}$$

$\ell n(z)$  is the principal value of the function  $\ell n(z)$ ;

$$q^\alpha := e^{\alpha \ell n(q)}, \quad [\alpha] := \frac{q^\alpha - q^{-\alpha}}{q - q^{-1}}, \quad \langle \alpha \rangle := \frac{q^\alpha + q^{-\alpha}}{2}, \quad \text{where } \alpha \in \mathbb{C};$$

For any  $\mathbb{Z}$ -graded vector space  $V = \bigoplus_{n \in \mathbb{Z}} V_n$ ,  $J \in \text{Hom}(V, V)$  is defined by  $J(v_n) := q^n v_n$  for  $n \in \mathbb{Z}$  and  $v_n \in V_n$ ;

$\sigma \in \text{Hom}(V, V)$  is defined by  $\sigma := \frac{J + J^{-1}}{2}$ ;

Let  $A = \bigoplus_{n \in \mathbb{Z}} A_n$  be a  $\mathbb{Z}$ -graded algebra over  $\mathbb{C}$  with multiplication denoted by  $xy$  for  $x, y \in A$ . We define  $(x, y, z)_q$ ,  $J_q(x, y, z)$ , the plus algebra  $(A^+, \circ)$ , and the minus algebra  $(A^-, [ , ]) by$

$$(x, y, z)_q := (xy)\sigma(z) - \sigma(x)(yz),$$

$$J_q(x, y, z) := (xy)\sigma(z) + (yz)\sigma(x) + (zx)\sigma(y),$$

$$A^\pm := A \quad (\text{as } \mathbb{Z}\text{-graded vector space}),$$

$$x \circ y := \frac{1}{2}(xy + yx), \quad [x, y] := xy - yx, \quad \text{where } x, y, z \in \mathbb{Z}.$$

**Definition 1.** Let  $A$  be a  $\mathbb{Z}$ -graded algebra over  $\mathbb{C}$  with multiplication denoted by  $xy$ .

(1)  $A$  is called quantum flexible if  $(x, y, x)_q = 0$  for  $x, y \in A$ .

(2)  $A$  is called a quantum Lie algebra if  $xy = -yx$  and the following quantum Jacobi identity holds:

$$J_q(x, y, z) = 0 \quad \text{for } x, y \text{ and } z \in A.$$

(3)  $A$  is called quantum associative if  $(x, y, z)_q = 0$  for  $x, y, z \in A$ .

(4)  $A$  is called a quantum Jordan algebra if  $(x, y, x^2)_q = 0$  for  $x, y \in A$ .

The quantum Witt algebra  $W_q$ , defined by

$$W_q := \bigoplus_{n \in \mathbb{Z}} \mathbb{C}d_n, \quad d_m d_n := [m - n]d_{m+n} \quad \text{for } m, n \in \mathbb{Z},$$

is a quantum Lie algebra.

If  $A$  is a quantum associative algebra, then  $(A^-, [ , ]) is a quantum Lie algebra. If  $A$  is a quantum associative algebra satisfying  $[x^2, x] = 0$  for  $x \in A$ , then  $(A^+, \circ)$  is a quantum Jordan algebra.$

**Definition 2.** The  $q$ -analogue  $U(W_q)$  of the enveloping algebra of the Witt algebra is defined as the associative algebra generated by  $\{J^{\pm 1}, d_m \mid m \in \mathbb{Z}\}$  with relations:

$$JJ^{-1} = J^{-1}J = 1, \quad Jd_m J^{-1} = q^m d_m,$$

$$q^m d_m d_n J - q^n d_n d_m J = [m - n]d_{m+n} \quad \text{for } m, n \in \mathbb{Z}.$$

**Definition 3.** The  $q$ -analogue  $U(V_q)$  of the enveloping algebra of the Virasoro algebra is defined as the associative algebra generated by  $\{J^{\pm 1}, c, d_m \mid m \in \mathbb{Z}\}$  with relations:

$$JJ^{-1} = J^{-1}J = 1, \quad Jd_m J^{-1} = q^m d_m, \quad cJ = Jc, \quad cd_m = q^m d_m c,$$

$$q^m d_m d_n J - q^n d_n d_m J = [m - n]d_{m+n} + \frac{[m - 1][m][m + 1]}{[2][3]} \delta_{m+n, 0} c,$$

where  $m, n \in \mathbb{Z}$  and  $q$  is not a root of unity.

Let  $U_q(W)$  be the associative algebra generated by  $\{d_m \mid m \in \mathbb{Z}\}$  with the following relations:

$$q^{m-n}d_m d_n - q^{n-m}d_n d_m = [m-n]d_{m+n} \quad \text{for } m, n \in \mathbb{Z}.$$

This algebra was constructed by T.L.Curtright and C.K.Zachos in [2].

**Definition 4.** A  $U(W_q)$ -module ( or  $U_q(W)$ -module )  $V$  is called a  $\mathbb{Z}$ -graded module if  $V = \bigoplus_{n \in \mathbb{Z}} V_n$  and  $d_m(v_n) \in V_{m+n}$  for  $m, n \in \mathbb{Z}$ .

**Remark.**  $V = \bigoplus_{n \in \mathbb{Z}} V_n$  is a  $\mathbb{Z}$ -graded  $U_q(W)$ -module if and only if  $V$  is a  $\mathbb{Z}$ -graded  $U(W_q)$ -module with  $J \mid V_n = q^n \cdot id$ .

§1. The quantum universal central extension of  $W_q$  at roots of unity. In [6], we have seen that if  $q$  is not a root of unity, then the quantum universal central extension of  $W_q$  is a 1-dimensional quantum central extension of  $W_q$ . The first theorem in this paper describes the quantum universal central extension of  $W_q$  at roots of unity. It turns out that if  $q$  is a root of unity, then the quantum universal central extension of  $W_q$  is an infinite-dimensional quantum central extension of  $W_q$ .

**THEOREM 1.** Let  $q$  be a  $t$ -th primitive root of unity and  $L = \bigoplus_{n \in \mathbb{Z}} L_n$  the quantum universal central extension of  $W_q$ . We have

(1) If  $T$  is odd and  $T \geq 5$ , then

$$L_{mT+i} = \begin{cases} Cl_{mT+i}, & \text{if } 1 \leq i \leq T-1, \\ Cl_{mT} \oplus \left( \bigoplus_{\substack{k+r=m \\ k,r \in \mathbb{Z}}} Cc_{kT,rT} \right), & \text{if } i = 0, \end{cases}$$

$$c_{kT,rT}L = 0, \quad \ell_{kT+i}\ell_{rT-j} = q^{(k+r)T}[i+j]\ell_{(k+r)T+(i-j)} + \delta_{ij} \frac{[i-1][i+1]}{[2][3] < i >} c_{kT,rT},$$

where  $m, k, r, i, j \in \mathbb{Z}$  and  $0 \leq i, j \leq T-1$ .

(2) If  $T$  is even and  $T \geq 4$ , then

$$L_{mT+i} = \begin{cases} Cl_{mT+i}, & \text{if } 1 \leq i \leq T-1, \\ Cl_{mT} \oplus \left( \bigoplus_{\substack{k+r=m \\ k,r \in \mathbb{Z} \\ k-r > -1}} Cc_{kT,rT} \right), & \text{if } i = 0, \end{cases}$$

$$c_{kT,rT}L = 0, \quad \ell_{kT+i}\ell_{rT-j} = q^{(k+r)T}[i+j]\ell_{(k+r)T+(i-j)} + \delta_{ij}\delta_{i+j,T}c_{kT,rT},$$

where  $m, k, r, i, j \in \mathbb{Z}$ ,  $0 \leq i, j \leq T-1$  and  $c_{kT,rT} := -c_{(r-1)T,(k+1)T}$  for  $k-r \leq -1$ .

(3) If  $T = 3$ , then

$$L_{3m+i} = \begin{cases} Cl_{3m+i}, & \text{if } i = 1, 2, \\ Cl_{3m} \oplus \left( \bigoplus_{\substack{k+r=3m \\ k > r, k,r \neq 0 \\ k,r \in \mathbb{Z}}} Cc_{3k,3r} \right), & \text{if } i = 0, \end{cases}$$

$$c_{3k,3r}L = 0,$$

$$\ell_{3k+i}\ell_{3r-j} = q^{3(k+r)}[i+j]\ell_{3(k+r)+(i-j)} + \delta_{ij}(\delta_{k0}-1)(\delta_{r0}-1)(3\delta_{i0}-2)c_{3k,3r},$$

where  $m, k, r \in \mathbb{Z}$ ,  $i, j = 0, 1$  and  $c_{3k,3r} := -c_{3r,3k}$  for  $k \leq r$ .

(4) If  $T = 2$ , then

$$L_n = \begin{cases} Cl_n, & \text{if } n \text{ is odd,} \\ Cl_n \oplus \left( \bigoplus_{\substack{k+r=n \\ k > r \\ k,r \in \mathbb{Z}}} Cc_{kr} \right), & \text{if } n \text{ is even,} \end{cases}$$

$$c_{kr}L = 0,$$

$$\ell_m \ell_n = \begin{cases} [m-n]\ell_{m+n}, & \text{if } m+n \text{ is odd,} \\ c_{mn}, & \text{if } m+n \text{ is even,} \end{cases}$$

where  $m, n, k, r \in \mathbb{Z}$  and  $c_{mn} := -c_{nm}$  if  $m \leq n$ .

**Remark.** For a fixed positive integer  $n$ , if  $q^{4n} = 1$ , then the quantum universal central extension of  $C[t, t^{-1}] \otimes sl_q^n(2)$  is also a quantum central extension of  $C[t, t^{-1}] \otimes sl_q^n(2)$  by an infinite dimensional  $\mathbb{Z}$ -graded vector space.

**§2. Characterizations of the quantum Witt algebra.** From [5], [8] and [9], we know that the usual Witt algebra has two characterizations. The second theorem in this paper asserts that if  $q$  is not a root of unity, then the  $q$ -deformations of the two characterizations hold for the quantum Witt algebra  $W_q$ .

**THEOREM 2.** Let  $A = \bigoplus_{n \in \mathbb{Z}} A_n$  be a  $\mathbb{Z}$ -graded algebra over  $\mathbb{C}$ . If  $q$  is not a root of unity, then the following are equivalent:

- (1)  $A$  is the quantum Witt algebra.
- (2)  $A$  is a quantum flexible algebra and  $(A^-, [ , ])$  is the quantum Witt algebra.
- (3)  $A$  is a quantum Lie algebra,  $\dim(A_n) \leq 1$  for all  $n \in \mathbb{Z}$  and

$$(A_0 A_1) \neq 0, \quad (A_1 A_{-1}) \neq 0, \quad (A_2 A_{-1}) \neq 0, \quad (A_{-2} A_1) \neq 0.$$

**§3. The  $q$ -analogue of one of Kaplanaky's theorems.** For any  $(\lambda, \alpha, \beta) \in \mathbb{C}^* \times \mathbb{C} \times \mathbb{C}$ , we have five kinds of  $\mathbb{Z}$ -graded  $U(W_q)$ -module  $A(\lambda, \alpha, \beta)$ ,  $B(\lambda, \alpha, \beta)$ ,  $C(\lambda, \alpha)$ ,  $D(\lambda, \alpha)$  and  $E(\lambda)$  defined by

$$\begin{aligned} A(\lambda, \alpha, \beta) &:= \bigoplus_{k \in \mathbb{Z}} \mathbb{C} v_k, & J(v_k) &:= \lambda q^k v_k, \\ d_n(v_k) &:= -\lambda^{-1} ([k + \alpha] q^\alpha + [1 + n] [\beta] q^{n+k}) v_{n+k}; \end{aligned}$$

$$\begin{aligned} B(\lambda, \alpha, \beta) &:= \bigoplus_{k \in \mathbb{Z}} \mathbb{C} v_k, & J(v_k) &:= \lambda q^k v_k, \\ d_n(v_k) &:= -\lambda^{-1} ([n + k + \alpha] q^{n+\alpha} + [1 - n] [\beta] q^{n+k}) v_{n+k}; \end{aligned}$$

$$\begin{aligned} C(\lambda, \alpha) &:= \bigoplus_{k \in \mathbb{Z}} \mathbb{C} v_k, & J(v_k) &:= \lambda q^k v_k, \\ d_n(v_k) &:= -\lambda^{-1} \left( q^\alpha [n + 1] [\alpha + n + k] + [n] \frac{q^{-n-k-1}}{q - q^{-1}} \right) v_{n+k}; \end{aligned}$$

$$\begin{aligned} D(\lambda, \alpha) &:= \bigoplus_{k \in \mathbb{Z}} \mathbb{C} v_k, & J(v_k) &:= \lambda q^k v_k, \\ d_n(v_k) &:= -\lambda^{-1} q^n \left( q^\alpha [1 - n] [\alpha + k] - [n] \frac{q^{-k-1}}{q - q^{-1}} \right) v_{n+k} \end{aligned}$$

and

$$E(\lambda) := \bigoplus_{k \in \mathbb{Z}} C v_k, \quad J(v_k) := \lambda q^k v_k,$$

$$d_n(v_k) := \lambda^{-1} \frac{q^{n^3 + 2nk - k}}{q - q^{-1}} v_{n+k},$$

where  $n, k \in \mathbb{Z}$ .

Although both  $A(1, \alpha, \beta)$  and  $B(1, \alpha, \beta)$  become the module of tensor fields over the Witt algebra when  $q \rightarrow 1$ , this is not the case for other modules. The third theorem in this paper is the  $q$ -analogue of the Theorem 2 proved by I. Kaplansky in [5].

**THEOREM 3.** *Let  $q$  be not a root of unity and  $V = \bigoplus_{k \in \mathbb{Z}} C v_k$  a  $\mathbb{Z}$ -graded  $U(W_q)$ -module with  $J(v_k) \in C v_k$  for  $k \in \mathbb{Z}$ . If  $d_1$  and  $d_{-1}$  are injective operators on  $V$ , then there exists some  $(\lambda, \alpha, \beta) \in C^* \times C \times C$  such that  $V$  is isomorphic to one of the following  $U(W_q)$ -modules:*

$$A(\lambda, \alpha, \beta), \quad B(\lambda, \alpha, \beta), \quad C(\lambda, \alpha), \quad D(\lambda, \alpha), \quad E(\lambda).$$

By Theorem 3, we have completed the classification of the  $\mathbb{Z}$ -graded  $U_q(W)$ -modules  $V = \bigoplus_{k \in \mathbb{Z}} C v_k$  which satisfy the condition that  $d_1$  and  $d_{-1}$  are injective operators on  $V$ .

Details of the proofs of these results will appear elsewhere.

#### REFERENCES

1. G.M.Benkart and R.V.Moody, *Algebras, Groups and Geometries* 3 (1986), 456-492.
2. T.L.Curtright and C.K.Zachos, *Physics Letters B* Vol 243, number 3 (1990), 237-244.
3. B.L.Feigin and D.B.Fuchs, *Funct. Anal. Appl.*(English Translation) 16 (1982), 114-126.
4. V.G.Kac and A.K.Raina, "Highest weight representations of infinite dimensional Lie algebras," World Scientific Publishing Co Pte Ltd, 1987.
5. I.Kaplansky, *Comm. Math. Phys.* 86 (1982), 49-54.
6. K.-Q.Liu, *C. R. Math. Rep. Acad. Sci. Canada XIII*, No.4, August (1991), 135-140.
7. R.V.Moody and A.Pianzola, "Lie algebras with triangular decomposition," preprint.
8. H.C.Myung, "Malcev-admissible algebras," *Progress in mathematics*, Birkhauser Boston, Inc., 1986.
9. M.L.Tomber, *Hadronic J.* 4 (1981), 183-197.

#### ACKNOWLEDGMENT

I would like to thank Prof. R.V.Moody for his guidance on the work above.

Department of Mathematics, University of Alberta, Edmonton, Alberta, Canada, T6G 2G1<sup>1</sup>  
and  
Department of Mathematics, Wuhan, P. R. China

<sup>1</sup>This is my mailing address.

# ON NOISY PATTERN MATCHING UNDER GEOMETRICAL CONSTRAINTS

Salvatore D. Morgera, *F.I.E.E.E.*

Presented by G.F.D. Duffé, *F.R.S.C.*

**Abstract:** Noisy pattern matching problems arise in many areas, e.g., computational vision, astronomy, and high energy physics. Least-squares pattern matching over the Euclidean space  $\mathbb{E}^n$  for unordered sets of cardinality  $p$  is commonly formulated as a combinatorial optimization problem having complexity  $p \cdot p!$ ,  $p \gg n$ . Since  $p$  may be  $10^3$  or larger in typical applications, less than satisfactory suboptimal methods are usually employed. A powerful hybrid approach is described which casts the pattern matching problem in a differentiable setting using rigid motion constraints which often apply and reduces the complexity to  $l_{21} \cdot n^4 + l_{12} \cdot p^3$ , where  $l_{12}$  and  $l_{21}$  are the number of iterations required by steepest ascent and singular value decomposition (SVD)-based procedures, respectively.

**1. Introduction.** Given two sets of points in  $\mathbb{E}^n$ ,  $\{\mathbf{x}_i : i = 1, 2, \dots, p\}$  and a *template*  $\{\mathbf{y}_i : i = 1, 2, \dots, p\}$ , we wish to match these patterns through a *joint* choice of  $\phi \in \Phi$ , a representation of the orthogonal Lie group, and of  $\pi \in \Pi_p$ , the set of all permutations of the integers  $\mathcal{Z}_p$ . The model which we employ in this work is of the form  $\mathbf{x}_{\pi(i)} = \phi(\mathbf{y}_i) + \mathbf{b} + \mathbf{n}_i$ , where  $\mathbf{b}$  is a translation and the  $\mathbf{n}_i$  are stochastic zero mean vectors which are mutually statistically independent,  $i = 1, 2, \dots, p$ . Without loss of generality, we set  $\mathbf{b} = \mathbf{0}$ . Define the  $(n \times p)$ -dimensional matrices  $X = [\mathbf{x}_1 \mathbf{x}_2 \cdots \mathbf{x}_p]$  and  $Y = [\mathbf{y}_1 \mathbf{y}_2 \cdots \mathbf{y}_p]$ . Let the orthogonal Lie group act on  $\mathbb{E}^n$  according to  $\phi(\mathbf{x}) \mapsto \Theta^T \mathbf{x}$ , where  $\Theta$  belongs to an  $n$ -dimensional matrix Lie group  $\mathcal{O}(n)$  for which  $\Theta^T \Theta = I$ , and let  $\Pi$  be the  $p$ -dimensional permutation matrix which acts on  $\{\mathbf{x}_i : i = 1, 2, \dots, p\}$  according to  $\{\mathbf{x}_{\pi(i)} : i = 1, 2, \dots, p\} \rightarrow X\Pi$ . Using a least-squares criterion, the function to be minimized with respect to  $(\Theta, \Pi)$  is

$$\epsilon(\Theta, \Pi) = \sum_{i=1}^p \|\phi(\mathbf{x}_{\pi(i)}) - \mathbf{y}_i\|^2 = \|\Theta^T X\Pi - Y\|^2. \quad (1)$$

An approach to minimizing (1) when the sets are ordered, i.e., when  $\Pi = I$ , was first discussed by Nadás (1978) and, in a different context,

by Schwartz and Sharir (1985). In this case, the solution is of the form  $\Theta^* = M^{-T}(M^T M)^{1/2}$ , where  $M = XY^T$  and the square root is the symmetric, positive definite square root. Brockett (1989) was the first to treat the unordered pattern matching case described above; he assumes that the optimization problems in  $\Theta$  and  $\Pi$  are separable.

Investigations by Morgera and Lie Chin Cheong (1990) and Lie Chin Cheong and Morgera (1991) have rigorously shown that minimization of (1) can indeed be carried out in a separable manner by solving the following subproblems:

1. Find the orthogonal matrix  $\Theta^* \in \mathcal{O}(n)$  which maximizes the function  $g(\Theta) = \sum_{i=1}^m \text{tr}[\Theta^T Q_i \Theta N_i]$ , where  $Q_i = Q_i(X)$  and  $N_i = N_i(Y)$  are  $m$  appropriately chosen symmetric matrices.
2. Find the permutation matrix  $\Pi^*$  which maximizes the function  $f(\Pi) = \text{tr}[Z^* \Pi Y^T]$ , where  $Z^* = \Theta^{*T} X$ .

We note that if  $\phi(x)$  acts linearly, then the orthogonal Lie group acts on an element  $Q$  of the space of symmetric matrices via  $\phi(Q) \mapsto \Theta^T Q \Theta$ . Step 1 results in a best least-squares match between  $m$  corresponding *symmetric* functions of the patterns, which are independent of  $\Pi$ , while step 2 conducts a search for that permutation which minimizes (1). In step 1, the solution is unaltered if we set  $m = 1$  and simply match the second moments of the patterns by using the symmetric matrices  $Q = XX^T$  and  $N = YY^T$ . We have observed, however, that one advantage to using  $m > 1$  symmetric functions in forming  $g(\Theta)$  is an improved convergence rate of the optimization process when a steepest ascent algorithm is used. At the optimum point, we have

$$\varepsilon(\Theta^*, \Pi^*) = \text{tr}(Q + N) - 2\text{tr}[\Theta^{*T} X \Pi^* Y^T]. \quad (2)$$

The idea of formulating combinatorial optimization problems in a differentiable setting has far-reaching implications and is due to Karmarkar (1984).

**2. Gradient Flow Fields.** The gradient of  $g(\Theta)$ , evaluated at  $\Theta$ , can be found by considering a small, *admissible* perturbation of the form  $\Theta \rightarrow \Theta(I + t\Delta)$ , where  $t > 0$  and  $\Delta$  is in the Lie algebra of  $\mathcal{O}(n)$ . The gradient is given by

$$\lim_{t \rightarrow 0} \frac{g(\Theta(I + t\Delta)) - g(\Theta)}{t} = 2 \sum_{i=1}^m \text{tr}[N_i \Theta^T Q_i \Theta \Delta]. \quad (3)$$

The gradient (3) is proportional to the trace inner product of  $\sum_{i=1}^m (\Theta^T Q_i \Theta N_i)$  and an element  $\Delta$  of the Lie algebra. As is well known, all elements of the Lie algebra associated with  $O(n)$  are skew-symmetric. Projecting the gradient onto the tangent space at  $\Theta$ , equating the result to  $\Theta^T \dot{\Theta}$ , and simplifying, results in the expression for the gradient flow field obtained by Brockett (1989),

$$\dot{\Theta} = \sum_{i=1}^m (Q_i \Theta N_i - \Theta N_i \Theta^T Q_i \Theta). \quad (4)$$

From (4), the optimum solution is seen to satisfy  $\sum_{i=1}^m \Theta^*{}^T Q_i \Theta^* N_i = \sum_{i=1}^m N_i \Theta^*{}^T Q_i \Theta^*$ , which cannot be explicitly solved for  $\Theta^*$ . The nature of the search space can, however, be appreciated by letting  $m = 1$  in  $g(\Theta)$ , defining  $Q$  and  $N$  as in Section 1, and further assuming that  $Q$  and  $N$  each have distinct eigenvalues. In this case, as  $\Theta$  varies over  $O(n)$ , a specialization by Brockett (1989) of the results of von Neumann (1937) for the unitary Lie group shows that  $g(\Theta)$  has  $2^n \cdot n!$  stationary points, of which  $2^n$  are *equivalued* local minima and  $2^n$  are *equivalued* local maxima.

We now describe the manner in which the optimum permutation matrix  $\Pi^*$  may be found. Form the  $p$ -dimensional diagonal matrices  $\mathcal{D}_{z_i}$  and  $\mathcal{D}_{y_i}$  from the  $i$ th rows of  $Z^*$  and  $Y$ , respectively,  $i = 1, 2, \dots, n$ . We may then write  $f(\Pi)$  as

$$f(\Pi) = \sum_{i=1}^n \text{tr}[\Pi^T \mathcal{D}_{z_i} \Pi \mathcal{D}_{y_i}]. \quad (5)$$

It is seen that (5) is just a special case of the functional  $g(\Theta)$ . In fact, since the permutation matrices are a subset of the orthogonal matrices, and, as pointed out by von Neumann (1937), a sufficient condition for a stationary point of  $g(\Theta)$  to be a permutation is that  $Q_i$  and  $N_i$  be of diagonal form, the same procedure described above may be applied to  $f(\Theta) = \sum_{i=1}^n \text{tr}[\Theta^T \mathcal{D}_{z_i} \Theta \mathcal{D}_{y_i}]$ ,  $\Theta \in O(n)$ , to find the gradient flow field. It is not difficult to show that the form of  $\Pi^*$  is that of a permuted diagonal square root of the identity matrix.

**3. Matching Procedures.** The steepest ascent method is, perhaps, the simplest procedure for finding an optimum point for those functionals described in Section 1. For example, the optimization procedure for  $g(\Theta)$  may be carried out using (4) in the following manner:

*Procedure 1: Steepest Ascent Procedure*

1. Choose an initial point,  $\Theta(0)$ , such that  $0 < |\Theta(0)| < \xi/3$ ,  $0 < \xi < 1$ .

2. Choose  $0 < \mu < (1 - \xi) / \sum_{i=1}^m |Q_i| |N_i|$ .
3. For  $k = 0, 1, 2, \dots$ , compute

$$A(k) = \sum_{i=1}^m Q_i \Theta(k) N_i \quad (6)$$

$$\Theta(k+1) = \Theta(k) + \mu [A(k) - \Theta(k) A^T(k) \Theta(k)]. \quad (7)$$

Convergence details are discussed in Morgera and Lie Chin Cheong (1990) and Morgera (1991). In practice, the initial point,  $\Theta(0) \in \mathcal{O}(n)$ , is chosen to be of diagonal form. The iterative process is halted when  $|\Theta(k) - \Theta(k-1)| \leq \delta$ , where  $\delta > 0$  is a suitably chosen tolerance. Essentially the same procedure may be employed to carry out the maximization of (5) with one important difference. The initial point must *not* be chosen to be of diagonal form to avoid convergence to a saddle point.

A more direct approach to the optimization problems may be taken using an SVD-based procedure as described in Lie Chin Cheong (1991). A description of the role that the SVD plays in a variety of statistical signal processing problems may be found in Scharf (1991). In the pattern matching problem, the SVD may be employed at each recursive step to compute the inverse positive definite square root matrix of a product of the form  $A^T A$ , where  $A = A(\Theta) = \sum_{i=1}^m Q_i \Theta N_i$ . Setting the gradient flow field (4) equal to zero and solving for  $\Theta^*$  in terms of  $A$  results in

$$\Theta^* = A(\Theta^*) [A^T(\Theta^*) A(\Theta^*)]^{-1/2}. \quad (8)$$

Using (8), the optimization procedure for  $g(\Theta)$  may be carried out in the following manner, where (10) employs the SVD:

*Procedure 2: SVD-Based Procedure*

1. Choose an initial point,  $\Theta(0)$ .
2. For  $k = 0, 1, 2, \dots$ , compute

$$A(k) = \sum_{i=1}^m Q_i \Theta(k) N_i \quad (9)$$

$$B(k) = [A^T(k) A(k)]^{-1/2} \quad (10)$$

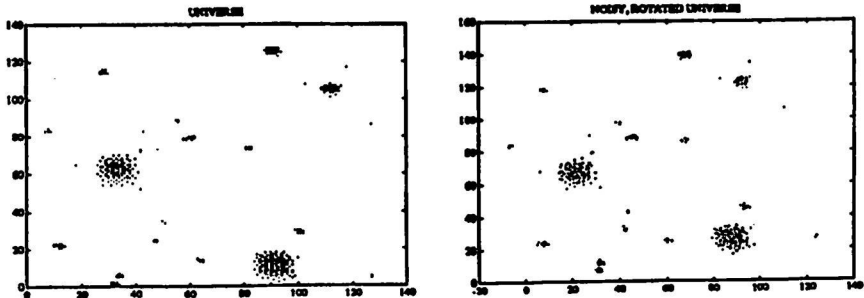
$$\Theta(k+1) = A(k) B(k). \quad (11)$$

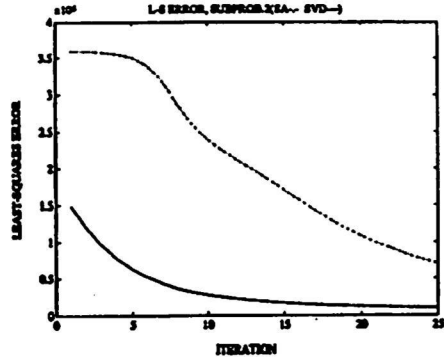
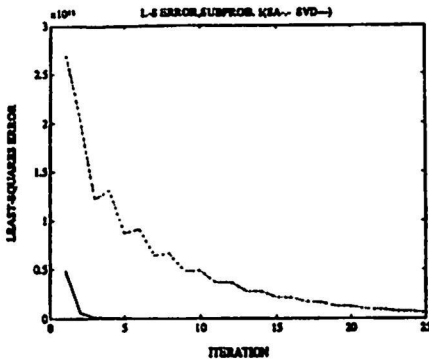
The update approach used in the above procedure is identical to one found in Armour and Morgera (1991). The procedure does not, in fact, require

that  $A^T(k)A(k)$  be nonsingular, as the SVD computation of (10) can still be carried out in pseudoinverse form.

**4. Computational Complexity.** Let  $l_{ij}$  denote the number of iterations required by Procedure  $i$  of Section 3 to obtain a solution to a specified tolerance,  $\delta_j$ , of subproblem  $j$  of Section 1,  $i, j = 1, 2$ . Assume that a mult/div operation "costs" as much as an add/sub operation, and define the combined computational complexity for solution of the overall pattern matching problem using Procedure  $i$  as  $C_i, i = 1, 2$ . We find that  $C_1 \sim l_{11} \cdot mn^3 + l_{12} \cdot p^3$  and that  $C_2 \sim l_{21} \cdot (n^4 + mn^3) + l_{22} \cdot p^4$ , where  $\sim$  denotes *is the order of*. Typically,  $n \leq 4$  and  $m \sim n$ ; in addition, we have observed from experiments that  $l_{21} \ll l_{11}$  and that  $l_{22} \cong l_{12}$ . This leads to the interesting hybrid approach in which the SVD-based procedure is used to maximize  $g(\Theta)$  and the steepest ascent procedure is used to maximize  $f(\Pi)$ . The computational complexity of this hybrid approach is then  $C \sim l_{21} \cdot n^4 + l_{12} \cdot p^3$ .

**5. Experimental Results.** The *universe* is a  $128 \times 128$  pixel binary subimage of a NASA image of the nebulae North America (NGC 7000) in the constellation Cygnus. A noisy ( $\sigma_n = 0.5$ ), rotated ( $10^\circ$ ), and permuted view of the same subimage is also shown. The subimages contain  $p = 386$  (star) points in  $\mathbb{E}^n$  which must be matched. The first l-s error plot shows that the solution of subproblem 1,  $\Theta^*$ , was obtained after only 3 iterations of the SVD-based procedure. The second l-s error plot shows that convergence to the solution of subproblem 2,  $\Pi^*$ , is approximately the same using either procedure; however, special precautions were necessary to assure numerical precision in the SVD-based procedure due to the extremely large matrices in this example. The rapid convergence rate observed in the solution of subproblem 1 also appears to be *independent* of  $p$ . The overall result of the matching process results in a subimage virtually identical to the *universe* with a rotation accuracy well within the nebulae angular diameter of  $1^\circ$ .





## REFERENCES

- [1] Armour, B. and Morgera, S.D. (1991). An exact forward-backward maximum likelihood autoregressive parameter estimation method. *IEEE Trans. Sig. Proc.* 39 1985-1993.
  - [2] Brockett, R.W. (1989). Least squares matching problems. *Lin. Alg. Appl.* 122/123/124 761-777.
  - [3] Karmarkar, N. (1984). A new polynomial-time algorithm for linear programming. *Combinatorica* 4 373-395.
  - [4] Lie Chin Cheong, P. (1991). *Geometrically Constrained Matching Schemes*. Ph.D. Thesis, Dept. of Elect. Engin., McGill University.
  - [5] Lie Chin Cheong, P. and Morgera, S.D. (1991). Matching schemes using steepest-ascent method. In: *Proc. IEEE Intern. Conf. on Acoust., Speech, and Sig. Proc.* Toronto 2349-2352.
  - [6] Morgera, S.D. (1991). Geometrically constrained pattern matching techniques. Presented at: *Fourth SIAM Conf. on Appl. Lin. Alg.* Minneapolis.
  - [7] Morgera, S.D. and Lie Chin Cheong, P. (1990). Geometrically Constrained Matching Schemes Using the Steepest-Ascent/Descent Methods. Submitted to: *IEEE Trans. Patt. Anal. and Mach. Intell.* See also: *ibid.* (1991). — *Research Report 91-R7*. Dept. of Elect. Engin., McGill University.
  - [8] Nádas, A. (1978). Least squares and maximum likelihood estimation of rigid motion. *IBM Research Report RC6945-Mathematics*. Sci. Comp. Dept., Sys. Prod. Div., NY.
  - [9] von Neumann, J. (1937). Some matrix-inequalities and metrization of matric-spaces. *Tomsk Univ. Rev.* 1 286-300. See also: A.H. Taub, ed. (1962). *John von Neumann: Collected Works*. Pergamon, New York 205-218.
  - [10] Scharf, L.L. (1991). The SVD and reduced-rank signal processing. In: R. Vaccaro, ed., *SVD and Signal Processing II*. Elsevier, Amsterdam 3-31.
  - [11] Schwartz, J.T. and Sharir, M. (1985). Identification of partially obscured objects in two dimensions by matching of noisy characteristic curves. *Courant Inst. of Math. Sci. Tech. Report 165*. New York University.
- Information Networks and Systems Laboratory  
 Department of Electrical Engineering  
 McGill University  
 Montréal H3A 2A7 Québec  
 CANADA

Received January 7, 1992

## TAME PRIMITIVITY FOR FREE NILPOTENT ALGEBRAS

Vesselin Drensky<sup>\*)</sup>

Presented by C.K. Gupta, F.R.S.C.

**Abstract.** Let  $F_n(\mathfrak{N})$  be the free algebra of rank  $n$  with free generators  $x_1, \dots, x_n$  in a nilpotent variety  $\mathfrak{N}$  of linear algebras over a field of characteristic 0. Let as a  $GL_n$ -module  $F_n(\mathfrak{N}) \cong \sum m(\lambda)N_n(\lambda)$ , where  $m(\lambda)$  are non-negative integers and  $N_n(\lambda)$  is the irreducible  $GL_n$ -module corresponding to the partition  $\lambda$ . Let  $c$  be a positive integer such that  $\lambda = (\lambda_1, \dots, \lambda_p)$  is a partition in  $\leq c$  parts for all  $m(\lambda) \neq 0$ . Then for every primitive system  $\{w_1, \dots, w_m\}$  of  $m \leq n - c$  elements in  $F_n(\mathfrak{N})$  there exists a tame automorphism  $\phi \in \text{Aut}F_n(\mathfrak{N})$  such that  $\phi(x_j) = w_j$ ,  $j = 1, \dots, m$ . Similarly, when the base field is arbitrary and the variety  $\mathfrak{N}$  is nilpotent of class  $c + 1$ , every primitive system of  $m \leq n - c$  elements in  $F_n(\mathfrak{N})$  can be obtained by a tame automorphism.

## Introduction

The starting point of this paper are two group theoretical results. Let  $F_n$  be the free group of rank  $n$ . C.K.Gupta, N.D.Gupta and V.A.Roman'kov [5] proved that any primitive system of  $m$  elements of the free metabelian nilpotent of class  $c$  group  $M_{n,c} = F_n/\gamma_{c+1}(F_n)F_n''$ ,  $m \leq n - 2$ , can be lifted to a primitive system of  $F_n$ . C.K. Gupta and N.D. Gupta [4] established that any primitive system of  $m$  elements in the free nilpotent of class  $c$  group  $F_{n,c} = F_n/\gamma_{c+1}(F_n)$ ,  $m \leq n - c$ , can also be lifted.

Our purpose is to obtain similar results for arbitrary varieties of nilpotent algebras. Let  $\mathfrak{N}$  be a nilpotent variety of linear algebras over a field  $K$  of characteristic 0, i.e. there exists an integer  $s$  such that all products of length  $s$  vanish in the algebras from  $\mathfrak{N}$ . Let  $F_n(\mathfrak{N})$  be the free algebra of rank  $n$  in  $\mathfrak{N}$  freely generated by  $x_1, \dots, x_n$ . We denote by  $T$  the group of the tame automorphisms of  $F_n(\mathfrak{N})$ ;  $T$  is generated by the invertible linear transformations on  $\text{span}\{x_1, \dots, x_n\}$  and the Jonquières automorphisms defined by  $x_1 \rightarrow x_1 + f(x_2, \dots, x_n)$ ,  $x_j \rightarrow x_j$ ,  $j > 1$ , where  $f(x_2, \dots, x_n) \in F_n(\mathfrak{N})$  does not depend on  $x_1$ . By a result of P.M. Cohn [1], for Lie algebras  $T$  consists of all automorphisms which can

<sup>\*)</sup> This research was carried out when the author was an Alexander von Humboldt fellow in the University of Bielefeld, Germany.

be lifted to automorphisms of the free Lie algebra. A system  $w = \{w_1, \dots, w_m\} \subset F_n(\mathfrak{N})$  is *primitive* if it can be included in a set of free generators of  $F_n(\mathfrak{N})$ , i.e. there exists an automorphism  $\phi \in \text{Aut}F_n(\mathfrak{N})$  such that  $\phi(x_j) = w_j$ ,  $j = 1, \dots, m$ . When there exists a tame automorphism with this property, we call the system  $w$  *tamely primitive*. An endomorphism  $\phi$  of a nilpotent algebra  $R$  is an automorphism if and only if  $\phi$  induces an automorphism on  $R/R^2$ . Hence the primitivity of  $w \subset F_n(\mathfrak{N})$  is equivalent to its linear independence modulo  $F_n^2(\mathfrak{N})$ . In the case of Lie algebras the system  $w$  is tamely primitive if and only if  $w$  can be lifted to a primitive system of the free Lie algebra.

The general linear group  $GL_n$  acts on  $\text{span}\{x_1, \dots, x_n\}$  and  $F_n(\mathfrak{N})$  has a natural  $GL_n$ -module structure. The irreducible polynomial representations of  $GL_n$  are described by partitions. For a partition  $\lambda = (\lambda_1, \dots, \lambda_p)$ ,  $\lambda_1 \geq \dots \geq \lambda_p \geq 0$ , we denote by  $N_n(\lambda)$  the corresponding irreducible  $GL_n$ -module;  $N_n(\lambda)$  is isomorphic to a submodule of the tensor algebra of  $\text{span}\{x_1, \dots, x_n\}$ .

We prove the following result. Let  $F_n(\mathfrak{N}) \cong \sum m(\lambda)N_n(\lambda)$ , where the multiplicities  $m(\lambda)$  are non-negative integers. If the integer  $c$  is such that  $m \leq n - c$  and for all  $m(\lambda) \neq 0$  the partitions  $\lambda = (\lambda_1, \dots, \lambda_p)$  satisfy  $p \leq c$ , then every primitive system of  $m$  elements in  $F_n(\mathfrak{N})$  is tame. The  $GL_n$ -module decomposition of  $F_n(\mathfrak{N})$  is known for many varieties of algebras. If  $\mathfrak{N} = \mathfrak{Q}^2 \cap \mathfrak{N}_s$  is the variety of metabelian nilpotent of class  $s$  Lie algebras, then  $c = 2$  and we obtain a new proof of the result by the author [2] that the primitive systems of  $n - 2$  elements in  $F_n(\mathfrak{Q}^2 \cap \mathfrak{N}_s)$  are tame. A weaker version of the main result holds without restriction on the base field. If  $\mathfrak{N}$  is a nilpotent of class  $c + 1$  variety of algebras over a field of arbitrary characteristic, then every primitive system of  $m$  elements in  $F_n(\mathfrak{N})$  is tame for  $m \leq n - c$ .

### The main results

We fix a field  $K$  of characteristic 0, a nilpotent variety  $\mathfrak{N}$  of linear  $K$ -algebras and a positive integer  $n$ . Let  $F = F_n(\mathfrak{N})$ . We denote by  $F^{[k]}$  the vector subspace of  $F$  of all homogeneous elements of degree  $k$  and by  $F^k$  the ideal  $\sum_{l \geq k} F^{[l]}$ . Let

$$IA_k = \{\phi \in \text{Aut}F \mid \phi(x_j) \equiv x_j \pmod{F^k}, j = 1, \dots, n\}, k \geq 2, IA = IA_2.$$

The group  $\text{Aut}F$  is a split extension of  $IA$  by  $GL_n$ . The action by conjugation of  $GL_n$  on  $IA_k/IA_{k+1}$  has the following useful description (see [3] for details). We consider the elements  $(f_1, \dots, f_n)$  of the vector space  $(F^{[k]})^{\otimes n}$  as  $1 \times n$  matrices with entries from  $F^{[k]}$

and  $g \in GL_n$  as an  $n \times n$  matrix. Let

$$g * (f_1, \dots, f_n) = (f_1(g(x_1), \dots, g(x_n)), \dots, f_n(g(x_1), \dots, g(x_n)))g^{-1}.$$

With this  $GL_n$ -action  $(F^{[k]})^{\otimes n}$  is isomorphic to the tensor product of  $GL_n$ -modules  $F^{[k]} \otimes_K \text{span}\{x_1^*, \dots, x_n^*\}$ , where  $\{x_1^*, \dots, x_n^*\}$  is the dual basis of  $\{x_1, \dots, x_n\}$ . Let  $\phi \in IA_k$ ,  $\phi(x_j) \equiv x_j + f_j \pmod{F^{k+1}}$ , where  $f_j \in F^{[k]}$ ,  $j = 1, \dots, n$ . We define a map

$$\nu_k : IA_k/IA_{k+1} \longrightarrow (F^{[k]})^{\otimes n} \text{ by } \nu_k(\bar{\phi}) = \nu_k(\phi.IA_{k+1}) = (f_1, \dots, f_n).$$

It turns out that  $\nu_k$  is a group isomorphism and  $\nu_k(g\bar{\phi}g^{-1}) = g * \nu_k(\bar{\phi})$ ,  $\bar{\phi} \in IA_k/IA_{k+1}$ ,  $g \in GL_n$ . Hence we may consider  $IA_k/IA_{k+1}$  as a  $GL_n$ -module. For the group  $T$  of the tame automorphisms we denote  $IT_k = T \cap IA_k$ . Then  $IT_k/IT_{k+1}$  is a submodule of  $IA_k/IA_{k+1}$ .

**Lemma.** Let  $\lambda = (\lambda_1, \dots, \lambda_c)$ ,  $\lambda_1 + \dots + \lambda_c = k$ , and let  $N_n(\lambda)$  be an irreducible  $GL_n$ -submodule of  $F^{[k]}$ . For every element  $f \in N_n(\lambda)$  there exists a tame automorphism  $\phi \in IT_k$  such that  $\phi(x_{c+1}) \equiv x_{c+1} + f$ ,  $\phi(x_j) \equiv x_j \pmod{F^{k+1}}$ ,  $j = c+2, \dots, n$ .

**Proof.** Let  $w_\lambda = w_\lambda(x_1, \dots, x_c)$  be a non-zero element of  $N_n(\lambda)$  which is homogeneous of degree  $\lambda_j$  in  $x_j$ ,  $j = 1, \dots, c$ ;  $w_\lambda$  is unique up to a multiplicative constant and is the highest weight vector of  $N_n(\lambda)$ . Let  $\psi$  be the automorphism of  $F$  defined by  $\psi(x_{c+1}) = x_{c+1} + w_\lambda(x_1, \dots, x_c)$ ,  $\psi(x_j) = x_j$ ,  $j \neq c+1$ . Clearly  $\psi \in IT_k$ . It is a well known fact from the representation theory of  $GL_n$  that any irreducible polynomial  $GL_n$ -module  $N$  is spanned by  $\{u(w)|u \in U_n^-\}$ , where  $w$  is the highest weight vector and  $U_n^-$  is the group of the lower unitriangular matrices. There exist  $u_1, \dots, u_r \in U_n^-$  and  $a_1, \dots, a_r \in K$  such that  $\sum_{i=1}^r a_i \cdot u_i(w_\lambda(x_1, \dots, x_c)) = f$ . Clearly

$$\nu_k(\psi) = (\underbrace{0, \dots, 0}_c, w_\lambda, \underbrace{0, \dots, 0}_{n-c-1})$$

and since  $u_i$  is a lower unitriangular matrix,

$$u_i * \nu_k(\psi) = (\underbrace{0, \dots, 0}_c, u_i(w_\lambda), \underbrace{0, \dots, 0}_{n-c-1})u_i^{-1} = (v_{i1}, \dots, v_{ic}, u_i(w_\lambda), \underbrace{0, \dots, 0}_{n-c-1})$$

for some  $v_{ii} \in F^{[k]}$ . Hence there exists  $\phi \in IT_k$  such that  $\nu_k(\phi) = \sum_{i=1}^r a_i(u_i * \nu_k(\psi))$  and  $\phi(x_{c+1}) \equiv x_{c+1} + f$ ,  $\phi(x_j) \equiv x_j \pmod{F^{k+1}}$ ,  $j = c+2, \dots, n$ .

**Theorem 1.** Let  $\mathfrak{N}$  be a nilpotent variety of linear algebras over a field of characteristic 0. If  $c$  is a positive integer such that  $m \leq n - c$  and

$$F_n(\mathfrak{N}) \cong \sum m(\lambda_1, \dots, \lambda_c)N_n(\lambda_1, \dots, \lambda_c),$$

then every primitive system of  $m$  elements in  $F_n(\mathfrak{N})$  is tame.

**Proof.** Let  $w$  be a primitive system of  $m$  elements in  $F = F_n(\mathfrak{N})$ . Without loss of generality we may assume that  $m = n - c$  and  $w = \{w_{c+1}, \dots, w_n\}$ . Since  $F^{k+1} = 0$  for  $k$  large enough, it is sufficient to show by induction that for every integer  $k \geq 1$  there exists a tame automorphism  $\tau_k$  such that  $\tau_k(x_j) \equiv w_j \pmod{F^{k+1}}$ ,  $j = c+1, \dots, n$ . The base of the induction is obvious. Since  $w_{c+1}, \dots, w_n$  are linearly independent modulo  $F^2$ , there exists  $\tau_1 = g \in GL_n \subset T$  such that  $\tau_1(x_j) \equiv w_j \pmod{F^2}$ ,  $j = c+1, \dots, n$ . Assume that we have found  $\tau_{k-1} \in T$  such that  $\tau_{k-1}(x_j) \equiv w_j \pmod{F^k}$ ,  $j = c+1, \dots, n$ . If  $v_j = \tau_{k-1}^{-1}(w_j)$ , then  $v_j \equiv x_j \pmod{F^k}$ , i.e.  $v_j \equiv x_j + f_j \pmod{F^{k+1}}$ ,  $f_j \in F^{[k]}$ ,  $j = c+1, \dots, n$ . Let  $F^{[k]} = N_1 \oplus \dots \oplus N_q$ , where  $N_i$  are irreducible  $GL_n$ -submodules of  $F^{[k]}$ . Then  $f_j = \sum_{i=1}^q f_{ij}$ , where  $f_{ij} \in N_i$ ,  $i = 1, \dots, q$ ,  $j = c+1, \dots, n$ . By the lemma, there exist tame automorphisms  $\phi_{ij} \in IT_k$  such that  $\phi_{ij}(x_j) \equiv x_j + f_{ij}$ ,  $\phi_{ij}(x_l) \equiv x_l \pmod{F^{k+1}}$ ,  $l = c+1, \dots, n$ ,  $l \neq j$ . Hence the product  $\phi = \prod_{j=c+1}^n \prod_{i=1}^q \phi_{ij}$  satisfies

$$\phi(x_j) \equiv x_j + \sum_{i=1}^q f_{ij} \equiv x_j + f_j \pmod{F^{k+1}}, \quad j = c+1, \dots, n.$$

Defining  $\tau_k = \tau_{k-1} \phi$  we obtain  $\tau_k \in T$  and

$$\tau_k(x_j) = \tau_{k-1}(\phi(x_j)) \equiv \tau_{k-1}(v_j) = w_j \pmod{F^{k+1}}, \quad j = c+1, \dots, n.$$

**Examples.** We give some examples for varieties of linear algebras when the integer  $c$  for the decomposition  $F_n(\mathfrak{N}) \cong \sum m(\lambda)N_n(\lambda)$ ,  $\lambda = (\lambda_1, \dots, \lambda_c)$ , is known. Since there exist close relations between varieties of groups and Lie algebras, it seems to be interesting to find a group theoretical analogue of Theorem 1 in the cases (v) and (vi).

(i) *The variety  $\mathfrak{N}_s$  of associative nilpotent of class  $s$  algebras.* In the class of all associative algebras  $\mathfrak{N}_s$  is defined by the polynomial identity  $x_1 \dots x_s = 0$ . It is known that  $F_n(\mathfrak{N}_s) \cong \sum \deg(\lambda)N_n(\lambda)$ , where  $\lambda$  is a partition of  $k = 1, \dots, s-1$  and  $\deg(\lambda)$  is the degree of the corresponding  $S_k$ -character  $\chi_k(\lambda)$ . Therefore  $c = s - 1$  and every primitive system of  $n - s + 1$  elements in  $F_n(\mathfrak{N}_s)$  is tame.

(ii) *The variety  $\mathfrak{N}_s \cap \mathfrak{A}$  of associative commutative nilpotent of class  $s$  algebras.* As a subvariety of  $\mathfrak{N}_s$  from (i),  $\mathfrak{N}_s \cap \mathfrak{A}$  is defined by  $x_1 x_2 - x_2 x_1 = 0$ . Since  $F_n(\mathfrak{N}_s \cap \mathfrak{A})$  is a homomorphic image of the non-unitary polynomial algebra  $K[x_1, \dots, x_n]$ ,  $F_n(\mathfrak{N}_s \cap \mathfrak{A}) \cong \sum_{k=1}^{s-1} N_n(k)$  and  $c = 1$ . Every primitive system of  $n - 1$  elements is tame.

(iii) *The variety  $\mathfrak{N}_s$  of all nilpotent of class  $s$  Lie algebras.* In the class of Lie algebras  $\mathfrak{N}_s$  is defined by  $[x_1, \dots, x_{s+1}] = 0$ . The free Lie algebra  $L_n$  is canonically embedded

into the free associative algebra  $K(x_1, \dots, x_n)$  and  $F_n(\mathfrak{N}_s)$  is isomorphic to a submodule of  $\sum \deg(\lambda)N_n(\lambda)$ ,  $\lambda$  being a partition of  $k \leq s$ . The only  $GL_n$ -submodule  $N_n(1^k)$  of  $K(x_1, \dots, x_n)$  is generated by the standard polynomial which is not a Lie element. Therefore  $N_n(1^k)$  does not participate in the decomposition of  $F_n(\mathfrak{N}_s)$  and  $c = s - 1$ .

(iv) The variety  $\mathfrak{A}^2 \cap \mathfrak{N}_s$  of metabelian nilpotent of class  $s$  Lie algebras. It is well known that  $F_n(\mathfrak{A}^2 \cap \mathfrak{N}_s) \cong N_n(1) + \sum_{k=2}^s N_n(k-1, 1)$ , i.e.  $c = 2$ . As a consequence we obtain a new proof of the result of [2]. The example from [2] shows that for  $s > 1$  there exist primitive systems of  $n - 1$  elements which are not tame.

(v) Nilpotent subvarieties of the centre-by-metabelian variety of Lie algebras  $[\mathfrak{A}^2, \mathfrak{E}]$ . It can be obtained from a result of S.P. Mishchenko [6] that  $F_n([\mathfrak{A}^2, \mathfrak{E}]) \cong N_n(1) + \sum_{k>1} (N_n(k-1, 1) + N_n(2k-2, 1^2) + N_n(2n-1, 2))$ . Hence  $c = 3$ .

(vi) Nilpotent subvarieties of the nilpotent of class  $s$ -by-abelian variety of Lie algebras  $\mathfrak{N}_s \mathfrak{A}$ . As a module  $F_n(\mathfrak{N}_s \mathfrak{A})$  is isomorphic to  $N_n(1) + \sum_{k=1}^s (L'_n)^k / (L'_n)^{k+1}$  and  $(L'_n)^k / (L'_n)^{k+1}$  is spanned on commutators  $[u_1, \dots, u_k]$ , where  $u_i = [x_{i_1}, \dots, x_{i_t}]$ ,  $i_1 > i_2 \leq \dots \leq i_t$ ,  $t \geq 2$ . It is easy to see that the  $GL_n$ -module  $(L'_n)^k / (L'_n)^{k+1}$  is a homomorphic image of the tensor power  $(L'_n / L''_n)^{\otimes k} \cong (\sum_{p>1} N_n(p-1, 1))^{\otimes k}$ . The Littlewood-Richardson rule gives that  $N_n(p_1 - 1, 1) \otimes \dots \otimes N_n(p_k - 1, 1)$  is a sum of irreducible submodules  $N_n(\lambda_1, \dots, \lambda_{2k})$ . Hence  $c = 2s$  for any nilpotent subvariety of  $\mathfrak{N}_s \mathfrak{A}$ .

(vii) Let  $R$  be a finite dimensional algebra of dimension  $c$ . It is well known that  $F_n(\text{var} R) \cong \sum m(\lambda)N_n(\lambda)$  with  $\lambda = (\lambda_1, \dots, \lambda_c)$ .

Till the end of the paper we consider algebras over an arbitrary field. For an algebra  $R$  we define  $R^1 = R$  and inductively  $R^k = \sum_{p+q=k} R^p R^q$ . The algebra  $R$  is nilpotent of class  $s$  if  $R^s = 0$ . The following theorem is an exact ring theoretical analogue of [4].

**Theorem 2.** *Let  $\mathfrak{N}$  be a nilpotent of class  $c + 1$  variety of algebras over an arbitrary field. Every primitive system on  $n - c$  elements in  $F_n(\mathfrak{N})$  is tame.*

**Proof.** Let  $F = F_n(\mathfrak{N})$  be freely generated by  $x_1, \dots, x_n$ . As in the proof of Theorem 1, it is sufficient to show that for any monomial  $u = u(x_1, \dots, x_n)$  of length  $k$ ,  $2 \leq k \leq c$ , and any  $a \in K$  there exists a tame automorphism  $\phi$  such that  $\phi(x_{c+1}) \equiv x_{c+1} + au$ ,  $\phi(x_j) \equiv x_j \pmod{F^{k+1}}$ ,  $j > c + 1$ . Let the degree of  $u$  in  $x_{c+1}$  is equal to  $p$ . We use induction on  $p$ . For  $p = 0$  we define  $\phi$  by  $\phi(x_{c+1}) = x_{c+1} + au$ ,  $\phi(x_j) = x_j$ ,  $j \neq c + 1$ . Since  $u$  does not depend on  $x_{c+1}$ ,  $\phi$  is tame. Let  $p > 0$ . Hence  $k - p < c$  and  $u$  does not depend on a variable  $x_1, \dots, x_c$ . Let for example  $u = u(x_2, \dots, x_n)$ . We define  $\psi \in T$ ,  $\psi(x_1) = x_1 - au(x_2, \dots, x_n)$ ,  $\psi(x_j) = x_j$ ,  $j \neq 1$ , and  $g \in GL_n$ ,  $g(x_{c+1}) = x_1 + x_{c+1}$ ,

$g(x_j) = x_j, j \neq c + 1$ . Easy calculations show that

$$g\psi g^{-1}(x_{c+1}) = x_{c+1} + au(x_2, \dots, x_c, x_1 + x_{c+1}, x_{c+2}, \dots, x_n), g\psi g^{-1}(x_j) = x_j, j > c + 1.$$

Clearly  $u(x_2, \dots, x_c, x_1 + x_{c+1}, x_{c+2}, \dots, x_n) = u(x_2, \dots, x_{c+1}, \dots, x_n) + v(x_1, \dots, x_n)$ , where  $v(x_1, \dots, x_n)$  is a linear combination of monomials with degree in  $x_{c+1}$  less than  $p$ . By induction, there exists a tame automorphism  $\theta$  such that  $\theta(x_{c+1}) \equiv x_{c+1} + av(x_1, \dots, x_n)$ ,  $\theta(x_j) \equiv x_j \pmod{F^{k+1}}, j > c + 1$ , and  $\phi = g\psi g^{-1}\theta$  is the desired tame automorphism.

### Acknowledgements

The author is very grateful to C.K. Gupta for the stimulating discussions.

### References

1. P.M. Cohn, *Subalgebras of free associative algebras*, Proc. London Math. Soc. (3) 14 (1968), 618–632.
2. V. Drensky, *Lifting of primitivity for free metabelian nilpotent Lie algebras*, preprint.
3. V. Drensky, C.K. Gupta, *Automorphisms of free nilpotent Lie algebras*, Canad. J. Math. 17 (1990), 259–279.
4. C.K. Gupta, N.D. Gupta, *Lifting primitivity of free nilpotent groups*, Proc. Amer. Math. Soc. (to appear).
5. C.K. Gupta, N.D. Gupta, V.A. Roman'kov, *Primitivity in free groups and free metabelian groups*, preprint.
6. S.P. Mishchenko, *Varieties of center-by-metabelian Lie algebras over a field of characteristic zero*, Mat. Zametki 30 (1981), 649–657 (Russian). Translation: Math. Notes 30 (1981), 822–826.

Institute of Mathematics, Bulgarian Academy of Sciences  
Acad. G. Bonchev Str., block 8, 1113 Sofia, Bulgaria

---

Received January 31, 1992

# Nielsen type numbers for:- fibre preserving maps, coincidences of fibre preserving maps, and for periodic points of fibre preserving maps

Philip R. Heath

*Presented by S. Halperin, F.R.S.C.*

## Abstract

In this note we announce results concerning three kinds of Nielsen type numbers. Firstly a Nielsen type number for fibre preserving maps ([H2]), secondly a Nielsen type number for fibre preserving coincidences of fibre preserving maps, and thirdly Nielsen type numbers for periodic points of fibre preserving maps. Each number is an appropriate lower bound for the number of fixed points, respectively coincidences, respectively periodic points, within the fibre homotopy classes of the maps involved. Moreover each number shares with the usual Nielsen number such properties as homotopy invariance and commutativity.

**Key words:** Nielsen Numbers, coincidences, periodic points, fibre spaces.

**AMS subject classifications:** primary 55M20, secondary 55R05.

## 1 A Nielsen type number for fibre preserving maps <sup>1</sup>

Much has been written about the ordinary Nielsen number of a fibre map (see for example [B], [BF], [Y], [J], [H] etc.). The idea has been to produce a product formula writing the Nielsen number of the map between the total space in terms of (appropriate) Nielsen numbers of the maps between the base

---

<sup>1</sup>The full text of the ideas from this section will appear in [H2].

and fibre. Of necessity the maps and homotopies involved in such calculation have to be fibre preserving. What has been calculated then is the Nielsen number of a fibre preserving map, but without exploiting any advantage that the restriction 'fibre preserving' might bring. The following example, which is an adaption of one given in [BF], illustrates this advantage, showing that the ordinary Nielsen number can be a poor lower bound for the least number of fixed points of a fibre preserving map.

**Example 1.1** Let  $S^1 \rightarrow S^3 \xrightarrow{p} S^2$  be the Hopf fibration, where  $S^1$  is the fibre over a point  $b \in S^2$ . Let  $f : S^1 \rightarrow S^1$  be a fibre preserving map of  $p$  with induced map  $\bar{f} : S^2 \rightarrow S^2$ . Suppose that  $\bar{f}$  has degree  $d \neq \pm 1$ , with  $b$  as a fixed point. Then  $f$  restricts to a self map  $f_b : S^1 \rightarrow S^1$  of the fibre. It is easy to see that the degree of  $f_b$  is  $d$ . Now  $f$  can be regarded as a self map of the pair  $(S^3, S^1)$ , and so according to relative Nielsen theory ([Sc]) has at least  $N(f; S^3, S^1) = N(f_b) = |1 - d|$  fixed points (use [Sc; Theorem 2.6]). Here  $N(f; S^3, S^1)$  is the relative Nielsen number of the map  $f$  of the pair  $(S^3, S^1)$ . On the other hand since  $L(f) = 1 - d^2 \neq 0$ , then  $N(f) = 1$ .

Let  $p : E \rightarrow B$  be a fibration in which  $E$ , and  $B$  are compact connected ANR's, then every fibre of  $p$  is also a compact connected ANR. If  $(f, \bar{f})$  is a fibre preserving map of  $p$ , (that is  $pf = \bar{f}p$ ), and  $b \in \Phi(\bar{f})$  (the fixed point set of  $\bar{f}$ ), then the restriction of  $f_b$  of  $f$  to  $p^{-1}(b) := F_b$  is a self map, and hence has its own Nielsen number  $N(f_b)$ .

**Definition 1.2** Let  $(f, \bar{f})$  be a fibre preserving map from  $p$  to itself, we say that  $(f, \bar{f})$  is essentially fibre uniform, if  $N(f_b)$  is independent of  $b$  in any essential class of  $\bar{f}$ .

$N(f_b)$  is independent of  $b$  within a single Nielsen class of  $\bar{f}$ , so  $(f, \bar{f})$  is essentially fibre uniform if  $N(\bar{f}) = 0$ , or 1. In addition  $(f, \bar{f})$  is essentially fibre uniform for all  $(f, \bar{f})$  if  $p$  is orientable (in the homotopy sense) (see [Se], [J], [Y] or [H]). In particular  $(f, \bar{f})$  is essentially fibre uniform if  $p$  is a trivial fibration or if  $B$  is simply connected, or if  $p$  is a principal bundle with path connected structure group [Se;p.445].

**Definition 1.3** Let  $(f, \bar{f})$  be a fibre preserving map of  $p$  which is essentially fibre uniform then the Nielsen type number  $N_{\mathcal{F}}(f, p)$  of  $(f, \bar{f})$  is the number  $N(\bar{f})N(f_b)$ , where  $N(f_b)$  is the Nielsen number of the restriction of  $f$  to the fibre over any point  $b$  in any essential class of  $\bar{f}$ .

The reader will notice the similarity of the definition of  $N_{\mathcal{F}}(f, p)$  with the so called naïve product formula for fibre maps due originally to Brown in [B], (see comments of Jiang [J;p. 87]). It seems to this author that our Nielsen type number expresses, but now accurately, the intuition Brown had when he was trying to write down his product formula [B]. In particular we have:-

**Theorem 1.4** *Let  $p$  be a fibration, and  $(f, \bar{f})$  a self fibre preserving map that is essentially fibre uniform, then the number  $N_{\mathcal{F}}(f, p)$  is a lower bound for  $M_{\mathcal{F}}(f, p)$ , where  $M_{\mathcal{F}}(f, p) = \min\{\#\Phi(g) : \text{where } g \text{ is fibre homotopic to } f\}$ .*

The proofs of 1.4, and the following proposition exhibit and exploit a connection between  $N_{\mathcal{F}}(f, p)$  and relative Nielsen theory. In fact  $N_{\mathcal{F}}(f, p)$  can be shown to be the relative Nielsen number of  $f$  taken with respect to a certain subspace of  $E$ .

**Proposition 1.5** *Let  $(f, \bar{f})$  be a self fibre preserving map of a fibration  $p$  that is essentially fibre uniform, then*

- (i)  $N_{\mathcal{F}}(f, p) \geq N(f)$ , and if  $N(f) = 0$  or  $N(\bar{f}) = 0$ , or  $N(f_b) = 0$  then equality holds.
- (ii) If  $N(\bar{f}) \neq 0$ , then for any  $b$  in any essential class of  $\bar{f}$  we have that

$$N_{\mathcal{F}}(f, p) \geq N(f; E, F_b) \geq N(f_b).$$

Using a theorem of You [Y], we are able to give necessary and sufficient conditions for  $N_{\mathcal{F}}(f, p)$  and  $N(f)$  to coincide.

Following [F], we call a space  $X$ , for which an index is defined, a Wecken space if, for every map  $f : X \rightarrow X$ , there exists a map  $g : X \rightarrow X$  homotopic to  $f$ , and with the property that  $g$  has exactly  $N(f)$  fixed points. It is known that every compact connected polyhedron without local cut points is a Wecken space, if it is not a surface of negative Euler characteristic (see [J2;p.760] and [Sc2;Lemma 2,p 525]).

We use the theory of cofibrations to prove:-

**Theorem 1.6** *Let  $(f, \bar{f})$  be a fibre preserving map from  $p$  to itself that is essentially fibre uniform, and with the property that  $B$  and every fibre of  $p$  is a Wecken space, then there is a fibre preserving map  $(g, \bar{g})$  that is fibre homotopic to  $(f, \bar{f})$  and with the property that  $g$  has exactly  $N_{\mathcal{F}}(f, p)$  fixed points.*

## 2 A fibred Nielsen coincidence number

If  $f, g : X \rightarrow Y$  are maps a coincidence of  $f$  and  $g$  is a point  $x \in X$  with  $f(x) = g(x)$ . Coincidence theory is a generalization of fixed point theory where  $g$  is taken to be the identity. If  $X$  and  $Y$  are compact connected orientable  $n$ -dimensional manifolds without boundary, then there is a Nielsen coincidence number  $N(f, g)$  that is a lower bound for the number  $M(f, g) = \min\{\#C(f', g')\}$ , where  $C$  denotes 'the set of coincidences of', and where  $f'$  and  $g'$  are homotopic to  $f$  and  $g$  respectively. We will be considering locally trivial bundles  $(E, p, B)$  and  $(E', p', B')$  in which the total spaces, bases and all fibres are compact connected closed orientable topological manifolds with respectively equal dimensions.

In a recent publication [Je], Jezierski gives a product formula for the Nielsen coincidence number of a fibre preserving map. His work generalizes (among other things) the work of Chengye You [Y]. Jezierski shows for a pair of fibre preserving maps  $(f, \bar{f})$  and  $(g, \bar{g})$  from  $p$  to  $p'$  as above, that one can write the Nielsen coincidence number  $N(f, g)$  in terms of (appropriate) Nielsen coincidence numbers of the base and fibre. Inspired by Jezierski's work we generalize our theory of Nielsen type numbers of fibre preserving maps and define a Nielsen coincidence number of fibre preserving maps as follows:-

**Definition 2.1** *Let  $(f, \bar{f}), (g, \bar{g}) : p \rightarrow p'$  be a pair of fibre preserving maps of  $p$  and  $p'$  as above, and suppose that  $(E', p', B')$  is orientable, then the fibred Nielsen coincidence number  $N_{\mathcal{F}}(f, g)$  of  $(f, \bar{f})$  and  $(g, \bar{g})$  is the number  $N(\bar{f}, \bar{g})N(f_b, g_b)$ , where  $N(f_b, g_b)$  is the Nielsen coincidence number of the restrictions of  $f$  respectively  $g$  to the fibre over any point  $b$  in any essential coincidence class of  $\bar{f}$  and  $\bar{g}$ .*

Let  $M_{\mathcal{F}}(f, g)$  be the fibred version of  $M(f, g)$  (compare 1.4) then we have:-

**Theorem 2.2** *Under the hypotheses of definition 2.1 we have that  $N_{\mathcal{F}}(f, g)$  is a lower bound of  $M_{\mathcal{F}}(f, g)$ .*

Comparison theorems of the type outlined in section 1 are forthcoming, together with theorems of the usual type concerning homotopy invariance and commutativity. Proofs here are variants of those indicated in section 1, since relative Nielsen theory for coincidences does not precisely generalize ordinary Nielsen theory.

### 3 Nielsen type numbers for periodic points of fibre preserving maps

Let  $f : X \rightarrow X$  be a given self map of a compact connected ANR  $X$ , and  $n$  a positive integer. The fixed points of the iterates of  $f$  are called periodic points. Let  $P_n(f)$  denote the set of periodic points of  $f$  of least period  $n$ . In [J], [HPY] and [HY] Nielsen type numbers  $NP_n(f)$ , and  $NF_n(f)$  have been given that are lower bounds for the numbers  $MP_n(f) = \min\{\#P_n(g) : g \simeq f\}$ , and  $M\Phi_n(f) = \min\{\#\Phi(g^n) : g \simeq f\}$ .

**Definition 3.1** Let  $(f, \bar{f})$  be a fibre preserving map of a fibration  $p$  which is orientable then the Nielsen type number  $NP_{\bar{f}}^n(f, p)$  of  $(f, \bar{f})$  is the number  $NP_n(\bar{f})N((f^n)_b)$ , where  $N((f^n)_b)$  is the Nielsen number of the restriction of  $f^n$ , the  $n$ th iterate of  $f$ , to the fibre over any point  $b$  in any essential class of  $\bar{f}^n$ .

As usual we have a lower bound theorem.

**Theorem 3.2** Under the hypotheses of definition 3.1 we have that the number  $NP_{\bar{f}}^n(f, p)$  is a lower bound for  $MP_{\bar{f}}^n(f, p)$ , where  $MP_{\bar{f}}^n(f, p)$  is the fibred version of  $MP_n(f)$ .

The definition and properties of the number  $N\Phi_{\bar{f}}^n(f, p)$  are more difficult to summarize, we do have however that

**Theorem 3.3**  $M\Phi_{\bar{f}}^n(f, p) \geq \sum_{m|n} NP_{\bar{f}}^m(f, p)$ .

#### References

- [B] Brown R. F., The Nielsen number of a fibre map, Ann. Math. 85 (1967), 483-493.
- [BF] —, and Fadell E., Corrections to "The Nielsen number of a fibre map", Annals of Math. 95 (1972), 365-367.
- [F] Fadell, E., Recent results in the fixed point theory of continuous maps, Bull. Amer. Math. Soc. 76 (1970) 10-29.
- [H] Heath P.R., Product Formulae for Nielsen Numbers of Fibre Maps, Pacific J. Math vol 117, No. 2 (1985) 267-289.
- [H2] —, A Nielsen type number for fibre preserving maps, to appear.

[HPY] ———, Piccinini R.A., You C., Nielsen type numbers of Periodic Points I. International Conference on Topological Fixed point Theory and Applications. Tianjin China. Springer Verlag Lecture Notes 1411.

[HY] ———, You C., Nielsen type numbers of Periodic Points II. To appear: Topology and Its applications.

[J] Jiang B., Lectures on Nielsen Fixed point Theory, Contemporary Mathematics v.14, Amer. Math. Society, Providence, Rhode Island, 1983.

[J2] ———, On the least number of fixed points, Amer. J. Math., 102 (1980), 749-763.

[Je] Jezierski J., The Nielsen number product formula for coincidences, Fund. Math. 134, (1989), p183-212.

[Sc] Schirmer H., A Relative Nielsen Number, Pacific J. Math., vol 122, No. 2, (1986), 459-473.

[Sc2] ———, Mappings of polyhedra with prescribed fixed points and fixed point indices, Pacific J. Math., 63 (1976), 521-530.

[Se] Serre J.P., Homologie singulière des espaces fibrés, Annals of Math. 54 (1951), 425-505.

[Y] You C.Y., Fixed point classes of a fibre map, Pacific J. Math. 100 (1982), 217-241.

Philip R. Heath  
Department of Mathematics,  
Memorial University of Newfoundland,  
Newfoundland, A1C 5S7  
Canada

---

Received February 19, 1992

$L_p$ -FUNCTIONALS OF WEIGHTED PARTIAL SUM PROCESSES

Barbara Szyszkowicz

*Presented by Miklós Csörgő, F.R.S.C.*

**Abstract:** We present optimal asymptotic results for the weighted  $L_p$ -distance,  $0 < p < \infty$ , of the partial sum process of independent identically distributed random variables. We consider also  $L_p$ -functionals of other processes which are of interest in changepoint analysis.

**1. Introduction.** Let  $X_1, X_2, \dots$  be independent, identically distributed random variables with  $EX_1 = 0$  and  $EX_1^2 = 1$ . The weak convergence of the weighted partial sum process  $n^{-1/2}S_{[nt]}/q(t)$ ,  $0 < t \leq 1$ , where  $S_{[nt]} = X_1 + \dots + X_{[nt]}$ , in supremum norm was proved by O'Reilly (1974) for continuous weight functions under the assumption of  $E|X_1|^3 < \infty$ . He asserted also, without proof, that the third moment condition could be dropped to two, and his theorem would remain true.

For an extension of the Komlós, Major and Tusnády (1975, 1976) approximation of partial sums to weighted supremum norm approximations which improve also the just mentioned result of O'Reilly (1974) in terms of the optimal class of weight functions as in Csörgő, Csörgő, Horváth and Mason (1986), we refer to Csörgő and Horváth (1988b) and the references given there.

Assuming the existence of two moments only, Szyszkowicz (1991b) obtained the following weighted supremum norm approximation (cf. Theorem 2.1 and Remark 2.2 there).

Let  $Q$  be the class of positive functions on  $(0, 1]$ , i.e.,  $\inf_{\delta \leq t \leq 1} q(t) > 0$  for all  $0 < \delta < 1$ , which are nondecreasing in a neighbourhood of zero. Let also

$$I(q, c) = \int_0^1 t^{-1} \exp(-ct^{-1}q^2(t)) dt, \quad c > 0.$$

**THEOREM A.** *Let  $X_1, X_2, \dots$  be i.i.d.r.v.'s such that  $EX_1 = 0$  and  $EX_1^2 = 1$ . Then a standard Wiener process  $\{W(t), 0 \leq t < \infty\}$  can be constructed in such a way that the following hold true.*

(a) *Let  $q \in Q$ . Then*

$$\sup_{0 < t \leq 1} |n^{-1/2}(S_{[nt]} - W(nt))/q(t)| = o_p(1)$$

*if and only if  $I(q, c) < \infty$  for all  $c > 0$ .*

(b) Let  $q \in Q$ . Then

$$\sup_{0 < t \leq 1} |n^{-1/2}(S_{[nt]} - W(nt))|/q(t) = O_P(1)$$

if and only if  $I(q, c) < \infty$  for some  $c > 0$ .

We note that part (a) of Theorem A gives the limiting distribution of any continuous in sup-norm functional of  $n^{-1/2}S_{[nt]}/q(t)$  whenever  $I(q, c) < \infty$  for all  $c > 0$ .

Here we consider the asymptotic behaviour of  $L_p$ -approximations and functionals of weighted partial sum processes and obtain the following result.

**THEOREM 1.1** Let  $X_1, X_2, \dots$  be i.i.d.r.v.'s such that  $EX_1 = 0$  and  $EX_1^2 = 1$ . We assume that  $0 < p < \infty$  and  $q \in Q$ .

(a) A standard Wiener process  $\{W(t), 0 \leq t < \infty\}$  can be constructed in such a way that

$$\int_0^1 |n^{-1/2}(S_{[nt]} - W(nt))|^p/q(t)dt = o_P(1)$$

if and only if

$$(1.1) \quad \int_0^1 t^{p/2}/q(t)dt < \infty.$$

(b) Let  $\{W(t), 0 \leq t < \infty\}$  be a standard Wiener process. Then

$$\int_0^1 |n^{-1/2}S_{[nt]}|^p/q(t)dt \xrightarrow{D} \int_0^1 |W(t)|^p/q(t)dt$$

if and only if (1.1) holds.

Weighted  $L_p$ -approximations of the empirical and quantile processes were studied by Csörgő, Horváth and Shao (1991).

We call attention to the fact that (a) and (b) of Theorem 1.1 are equivalent. The above theorem enables us to obtain also other results of interest.

**2. U-statistic type processes.** In the context of the so-called changepoint problem, Csörgő and Horváth (1988a) study tests based on processes of  $U$ -statistics which are generalizations of Wilcoxon-Mann-Whitney type statistics. They consider

$$(2.1) \quad Z_k = \sum_{1 \leq i \leq k} \sum_{k+1 \leq j \leq n} h(X_i, X_j), \quad 1 \leq k < n,$$

where  $h(x, y)$  is a symmetric or an antisymmetric function and study its limiting behaviour in unweighted and weighted sup-norm metrics. Typical choices of  $h$  are  $xy$ ,  $(x - y)^2/2$  (sample variance),  $|x - y|$  (Gini's mean difference),  $\text{sign}(x + y)$  (Wilcoxon's one-sample statistic) (cf. Serfling, 1980). Taking  $h(x, y) = x - y$  gives  $Z_k$  equal to  $n(S_k - \frac{k}{n}S_n)$ , or, equivalently, to

$$k(n - k) \left( \frac{1}{k} \sum_{i=1}^k X_i - \frac{1}{n - k} \sum_{i=k+1}^n X_i \right).$$

The case of  $h(x, y) = \text{sign}(x - y)$  (cumulative rank tests) has gained special attention in the literature. For reviews of results concerning changepoint problems, we refer to Csörgő and Horváth (1988b) and the references given there.

Let  $X_1, X_2, \dots$  be independent identically distributed random variables. We assume that  $h$  is a symmetric function, i.e.,  $h(x, y) = h(y, x)$ , and that

$$(2.2) \quad Eh^2(X_1, X_2) < \infty.$$

We let  $Eh(X_1, X_2) = \theta$ ,  $\bar{h}(t) = E\{h(X_1, t) - \theta\}$ . Condition (2.2) implies that  $E\bar{h}^2(X_1) < \infty$  and we assume

$$(2.3) \quad 0 < \sigma^2 = E\bar{h}^2(X_1).$$

We consider the following process

$$U_k = Z_k - k(n - k)\theta, \quad 1 \leq k < n,$$

introduced by Csörgő and Horváth (1988a), which they expressed as a linear combination of  $U$ -statistics. Hence the study of  $U_k$  can be based on the projection of a  $U$ -statistic on the basic observations (cf. Chap. 5 of Serfling, 1980).

Let  $Q^*$  be the class of positive functions on  $(0, 1)$  i.e.  $\inf_{\delta \leq t \leq 1 - \delta} q(t) > 0$  for all  $\delta \in (0, 1/2)$ , which are nondecreasing near zero and nonincreasing near one. We define the Gaussian process  $\Gamma$  by

$$\Gamma(t) = (1 - t)W(t) + t\{W(1) - W(t)\}, \quad 0 \leq t \leq 1,$$

where  $\{W(t), 0 \leq t < \infty\}$  is a Wiener process.

**THEOREM 2.1** *We assume that (2.2) and (2.3) hold and  $0 < p < \infty$ .*

(a) *We can define a sequence of Gaussian processes  $\{\Gamma_n(t), 0 \leq t \leq 1\}$  such that for each  $n \geq 1$   $\{\Gamma_n(t), 0 \leq t \leq 1\} \stackrel{D}{=} \{\Gamma(t), 0 \leq t \leq 1\}$  and with  $q \in Q^*$  we have*

$$\int_0^1 \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{\lfloor (n+1)t \rfloor} - \Gamma_n(t) \right|^p / q(t) dt = o_P(1)$$

if and only if

$$(2.4) \quad \int_0^1 (t(1-t))^{p/2} / q(t) dt < \infty.$$

(b) If  $q \in Q^*$ , then

$$\int_0^1 \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{\lfloor (n+1)t \rfloor} \right|^p / q(t) dt \xrightarrow{\mathcal{D}} \int_0^1 |\Gamma(t)|^p / q(t) dt$$

if and only if (2.4) holds.

When considering an antisymmetric kernel, i.e.  $h(x, y) = -h(y, x)$ , similar type of results hold. However the limiting process is different. In this case  $Eh(X_1, X_2) = 0$  and, as in the symmetric case, we let  $\tilde{h}(t) = Eh(t, X_1)$ . We assume

$$(2.5) \quad Eh^2(X_1, X_2) < \infty \quad \text{and} \quad 0 < \sigma^2 = E\tilde{h}^2(X_1).$$

Now we have  $\mathcal{U}_k = Z_k$ , where  $Z_k$  is defined by (2.1).

**THEOREM 2.2.** We assume that (2.5) holds and  $0 < p < \infty$ .

(a) We can define a sequence of Brownian bridges  $\{B_n(t), 0 \leq t \leq 1\}$  such that with  $q \in Q^*$  we have

$$\int_0^1 \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{\lfloor (n+1)t \rfloor} - B_n(t) \right|^p / q(t) dt = o_P(1)$$

if and only if (2.4) holds.

(b) If  $q \in Q^*$  then

$$\int_0^1 \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{\lfloor (n+1)t \rfloor} \right|^p / q(t) dt \xrightarrow{\mathcal{D}} \int_0^1 |B(t)|^p / q(t) dt$$

if and only if (2.4) holds.

The weighted approximation results for  $\mathcal{U}_k$  process in supremum norm were proved by Csörgő and Horváth (1988a), assuming  $E|h(X_1, X_2)|^\nu < \infty$  for some  $\nu > 2$ , and were extended to contiguous alternatives by Szyszkowicz (1991a). These results were reproved by Szyszkowicz (1991b) under the initial assumption of two moments only for an unweighted approximation (cf. Theorems 2.1 and 4.1 of Csörgő and Horváth, 1988a).

**3. Sequential ranks.** Let  $X_1, X_2, \dots$  be independent continuously distributed random variables and let  $\xi_1, \dots, \xi_n$  denote the normalized sequential ranks

$$\xi_k = k^{-1} \sum_{i=1}^k 1\{X_i < X_k\}, \quad k = 1, \dots, n,$$

of the first  $n$  of the random variables  $X_1, X_2, \dots$ .

Statistical tests based on sequential ranks are frequently used because of their convenient properties. Bhattacharya and Frierson (1981) propose a nonparametric control chart based on partial sums of sequential ranks for detecting small changes in the distribution of a random sample. Khmaladze and Parjanadze (1986) study changepoint problems under contiguous alternatives using linear statistics of sequential ranks. They consider the process

$$n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_k), \quad 0 \leq t \leq 1$$

for  $a \in L^2(0, 1)$  and such that  $\int_0^1 a(u)du = 0$ , where  $\xi_1, \xi_2, \dots$  are normalized sequential ranks. Szyszkowicz (1991b) extends their results in supremum norm to using weight functions. For further results along these lines we refer to Lombard (1983). Here we present the following description of the asymptotic behaviour of  $L_p$ -functionals of the process  $\{n^{-1/2} \sum_{i \leq nt} a(\xi_k)/q(t), 0 < t < 1\}$ .

Let  $Q$  be the class of functions  $q$  as defined in Section 1.

**THEOREM 3.1** *Let  $0 < p < \infty$ . We assume that  $a \in L^2[0, 1]$ , where  $\int_0^1 a(u)du = 0$ ,  $\int_0^1 a^2(u)du = 1$  and  $a(u)$  has a uniformly bounded derivative on  $[0, 1]$ .*

(a) *Let  $q \in Q$ . Then there exists a sequence of Wiener processes  $\{W_n(t), 0 \leq t \leq 1\}$  such that*

$$\int_0^1 \left| n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i) - W_n(t) \right|^p / q(t) dt = o_p(1)$$

*if and only if (1.1) holds.*

(b) *If  $q \in Q$ , then*

$$\int_0^1 \left| n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i) \right|^p / q(t) dt \xrightarrow{D} \int_0^1 |W(t)|^p / q(t) dt$$

*if and only if (1.1) holds.*

**Acknowledgements** This research was financially supported by an NSERC Canada Postgraduate Scholarship and a scholarship from the Faculty of Graduate Studies and Research of Carleton University, Ottawa. These results constitute a part of the author's Ph.D. dissertation in preparation. The author wishes to acknowledge her gratitude for the guidance and continued advice of Professor Miklós Csörgő.

## REFERENCES

- [1] BHATTACHARYA, P.K. and FRIERSON, D. JR. (1981). A nonparametric control chart for detecting small disorders, *Ann. Statist.* 9 544-554.
- [2] CSÖRGÖ, M., CSÖRGÖ, S., HORVÁTH, L. and MASON, D. (1986). Weighted empirical and quantile processes. *Ann. Probab.* 14 31-85.
- [3] CSÖRGÖ, M. and HORVÁTH, L. (1988a). Invariance principles for changepoint problems. *J. Multivariate Anal.* 27 151-168.
- [4] CSÖRGÖ, M. and HORVÁTH, L. (1988b). Nonparametric methods for changepoint problems. *Handbook of Statistics*, Vol. 7, Elsevier Science Publishers B.V. (North-Holland).
- [5] CSÖRGÖ, M., HORVÁTH, L. and SHAO, Q.M. (1991). Convergence of integrals of uniform empirical and quantile processes, in Tech. Rep. Ser. Lab. Statist. Probab., Carleton University, No. 168. To appear in *Stoch. Proc. Appl.*
- [6] KHMALADZE, E.V. and PARJANADZE, A.M. (1986). Functional limit theorems for linear statistics of sequential ranks. *Probab. Theory Related Fields* 73 322-334.
- [7] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975). An approximation of partial sums of independent R.V.'s and the sample D.F.I., *Z. Wahrsch. Verw. Gebiete* 32 111-131.
- [8] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1976). An approximation of partial sums of independent R.V.'s and the sample D.F.II., *Z. Wahrsch. Verw. Gebiete* 34 33-58.
- [9] LOMBARD, F. (1983). Asymptotic distributions of rank statistics in the change-point problem. *South African Statist. J.* 17 83-105.
- [10] O'REILLY, N. (1974). On the weak convergence of empirical processes in sup-norm metrics. *Ann. Probab.* 2 642-651.
- [11] SERFLING, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.
- [12] SZYSZKOWICZ, B. (1991a). Changepoint problem and contiguous alternatives. *Statist. and Probab. Letters* 11 299-309.
- [13] SZYSZKOWICZ, B. (1991b). Weighted stochastic processes under contiguous alternatives. *C.R. Math. Rep. Acad. Sci. Canada*, Vol. XIII, No. 5 161-166.

DEPARTMENT OF MATHEMATICS & STATISTICS  
 CARLETON UNIVERSITY  
 OTTAWA, ONTARIO K1S 5B6  
 CANADA

---

Received February 20, 1992

# Large Deviations For Markov Processes With Mean Field Interaction And Unbounded Jumps

Shui Feng

*Presented by D.A.Dawson, F.R.S.C.*

**Abstract:** An  $N$ -particle system with mean field interaction is considered. Large deviation estimates for the empirical distributions as  $N$  goes to infinity are obtained under conditions which are satisfied by many interesting models including the first and the second Schlögl models.

1. Introduction Let  $E = \{0, 1, \dots\}$ ,  $E^{\otimes N}$  be the  $N$ -fold product of  $E$ .  $x^{(N)}(t) = (x_1^{(N)}(t), \dots, x_N^{(N)}(t))$  on  $E^{\otimes N}$  is a Markov process generated by

$$(1.1) \quad \Omega^{(N)}\psi(x^{(N)}) = \sum_{k=1}^N Q_{\varepsilon_{x^{(N)}}}^{(k)}\psi(x^{(N)})$$

where  $\varepsilon_{x^{(N)}} = \frac{1}{N} \sum_{k=1}^N \delta_{x_k^{(N)}}$ ,  $x^{(N)} = (x_1^{(N)}, \dots, x_N^{(N)}) \in E^{\otimes N}$ .  $Q_\nu$  is defined by

$$(1.2) \quad Q_\nu f(x) = \sum_{y \in E} q_{x,y} (f(y) - f(x)) + \|\nu\| (f(x+1) - f(x))$$

$Q = (q_{x,y})$  is a  $Q$ -matrix.  $\|\nu\|$  is the first moment of  $\nu$ .  $Q_\nu^{(k)}$  is used instead of  $Q_\nu$  when it acts on the  $k$ -th variable of  $\psi \in C(E^{\otimes N})$ .

This is an  $N$ -particle system with mean field interaction. The jump rates  $q_{i,j}$  can be unbounded. Consider the empirical distribution  $\eta_{x^{(N)}(\cdot)}$  of the  $N$ -particle system

$$(1.3) \quad \eta_{x^{(N)}(\cdot)} = \frac{1}{N} \sum_{k=1}^N \delta_{x_k^{(N)}(\cdot)}$$

This is a random measure on  $D([0, T], E)$ , the space of all right continuous functions  $\omega : [0, T] \rightarrow E$  which have left limits at each  $t \in (0, T]$  and are left continuous at  $T$  with the

Skorohod metric  $d$  on it.  $\delta_x$  denotes the Dirac distribution on path  $x(\cdot)$ . Assume,  $N \geq 1$ , that  $(x_1^{(N)}(0), \dots, x_N^{(N)}(0))$  are independent and identically distributed with common distribution  $\nu \in M_1(E)$  (the set of all probability measures on  $E$ ). We prove that  $\eta_{x^{(N)}(\cdot)}$  converges in law to a certain measure on  $D([0, T], E)$  whose marginal distribution is a solution of the following nonlinear master equation

$$(1.4) \quad \frac{du(t)(\cdot)}{dt} = \sum_{y \neq \cdot} \{u(t)(y)q_{y,\cdot} - u(t)(\cdot)q_{\cdot,y}\} + \|u(t)\| \{u(t)(\cdot - 1) - u(t)(\cdot)\}.$$

The main topic of the present paper is the large deviation problem for the distributions of  $\eta_{x^{(N)}(\cdot)}$  as  $N$  goes to infinity.

Several recent papers have discussed similar kind of large deviation problem for related models. Among them see for instance [1], [2], [3] and [4]. However this paper is the first to obtain the complete large deviation principle for Markov process with mean field interaction and unbounded jumps.

Proofs, as well as further details, will appear in a forthcoming paper by the author.

**2. Notation and Preliminary.** Let  $E$  be equipped with the discrete topology and  $\rho$  be the discrete metric. For fixed  $T > 0$ , let  $D = D([0, T], E)$ . For each  $t \geq 0$ , let  $\mathcal{F}_t = \sigma(x(s) : 0 \leq s \leq t)$ .  $\mathcal{M} = M_1(D)$  denotes the space of all probability measures on  $D$  with weak topology. For each  $R \geq 1$ , define  $\mathcal{M}_R = \{P \in \mathcal{M} : \sup_{0 \leq t \leq T} \int \varphi(\omega(t)) P(d\omega) \leq R\}$  equipped with the subspace topology of  $\mathcal{M}$ , where  $\varphi$  is defined as  $\varphi : E \rightarrow R$ ,  $x \rightarrow 1 + x \log \log(x + 2)$ . Let  $\mathcal{M}_\infty = \bigcup_{R \geq 1} \mathcal{M}_R$  be equipped with the "inductive topology". By definition, a set  $V$  is open in  $\mathcal{M}_\infty$  if and only if  $V \cap \mathcal{M}_R$  is open in  $\mathcal{M}_R$  for each  $R \geq 1$ . The "inductive topology" is stronger than the weak topology and  $\mathcal{M}_\infty$  is not a Polish space. Let  $Q = (q_{x,y})_{x,y \in E}$  be a totally stable conservative  $Q$ -matrix satisfying:

$$(2.1) \quad \inf_{x \in E} \{q_{x,x+1}\} > 0$$

$$(2.2) \quad \exists \Lambda > 0, \text{ such that } q_{x,y} = 0 \text{ for } |x - y| \geq \Lambda.$$

$$(2.3) \quad \forall x \in E, \exists \lambda > 0 \text{ such that } \sum_{y \in E} q_{x,y}(y - x) \leq \lambda x, \quad \sum_{x,y} q_{x,y}(\varphi(y) - \varphi(x)) \leq \lambda \varphi(x)$$

$$(2.4) \quad \exists 0 < c < \infty, \text{ such that } \forall x, y \in E, y > x,$$

$$\begin{aligned}
& \sum_{z \neq 0} (q_{y, y+z} - q_{z, z+z})z \\
& + 2 \sum_{z=1}^{\infty} [(q_{y, z-z} - q_{z, 2z-y+z}) \vee 0 + (q_{z, y+z} + q_{y, 2y-z+z}) \vee 0] z \leq c(y-x) \\
(2.5) \quad & \forall l \geq 0, \exists \lambda(l) > 0, \exists
\end{aligned}$$

$$\begin{aligned}
& \sup_{z \in E} \max \left\{ \sum_{y \in E} q_{z, y} (e^{y-z} - 1) + (e - 1)z, \right. \\
& \left. \sum_{y \in E} q_{z, y} (e^{\varphi(y) - \varphi(z)} - 1) + (e^{\varphi(x+1) - \varphi(x)} - 1)l \right\} \leq \lambda(l).
\end{aligned}$$

**Remark.** Conditions (2.3) and (2.4) guarantee that the corresponding martingale problems are well-posed. Conditions (2.1), (2.2) and (2.5) are needed only for the large deviation result.

These conditions are satisfied by the following two Schlögl models. (cf. [5])

The first Schlögl model:

$$(2.6) \quad q_{z, y} = \begin{cases} \beta_0 + \beta_1 z & \text{for } y = z + 1, z \geq 0 \\ \alpha_1 z + \alpha_2 z(z-1) & \text{for } y = z - 1, z \geq 1 \\ -\sum_{y \neq z} q_{z, y} & \text{for } y = z \\ 0 & \text{otherwise} \end{cases}$$

The second Schlögl model:

$$(2.7) \quad q_{z, y} = \begin{cases} \beta_0 + \beta_2 z(z-1) & \text{for } y = z + 1, z \geq 0 \\ \alpha_1 z + \alpha_3 z(z-1)(z-2) & \text{for } y = z - 1, z \geq 1 \\ -\sum_{y \neq z} q_{z, y} & \text{for } y = z \\ 0 & \text{otherwise} \end{cases}$$

where all coefficients are positive constants.

$\forall u \in M_1(E)$ ,  $u(\cdot) \in D([0, T], M_1(E))$  let  $\langle \varphi, u \rangle = \int_E \varphi u(dx) < \infty$ ,  $\sup_{0 \leq t \leq T} \langle \varphi, u(t) \rangle < \infty$  (all  $u$  and  $u(\cdot)$  below have these properties), then we have

**Theorem 2.1** Under assumptions (2.1)-(2.5) we have

(a). The time-inhomogeneous martingale problem for  $Q_{u(t)}$  with initial distribution  $u$  is well-posed. The solution is denoted by  $P_{u(\cdot), u}$ .

(b). The nonlinear martingale problem for  $Q$  with initial distribution  $u$  is well-posed.

(c). For each  $N \geq 1$ , the martingale problem for  $\Omega^{(N)}$  on  $D([0, T], E)^{\otimes N}$  with initial distribution  $u^{\otimes N}$  is well-posed. This solution is denoted by  $P_u^{(N)}$ .

For  $x, y \in E; u, \nu \in M_1(E)$ , define

$$(2.8) \quad Q_u(x, y) = \begin{cases} q_{x, x+1} + \|u\| & \text{for } y = x + 1 \\ q_{x, y} & \text{for } y \neq x, y \neq x + 1 \\ 0 & \text{otherwise} \end{cases}$$

$$(2.9) \quad q_u^\nu(x, y) = \frac{Q_\nu(x, y)}{Q_u(x, y)}.$$

For each  $\omega \in D$  and  $A \subset E$ , define

$$(2.10) \quad N(t, A; \omega) = \#\{s : \omega(s) \in A, \omega(s) \neq \omega(s-), s \leq t\}$$

where  $Q_{u(s)}(\omega(s), A) = \sum_{y \in A} Q_{u(s)}(\omega(s), y)$  and  $u(\cdot) \in D([0, T], M_1(E))$ .

Let  $P_u^{(N)}$  be the unique solution to the martingale problem for  $\Omega^{(N)}$  with initial distribution  $u^{\otimes N}$ ,  $P_{u(\cdot), u}^{(N)}$  denote the  $N$ -fold independent product of  $P_{u(\cdot), u}$ .  $\forall P \in \mathcal{M}$  let  $\pi(P)(\cdot) = P \circ x_{(\cdot)}^{-1}$ . Then we have

**Theorem 2.2**  $P_{u(\cdot), u}^{(N)}$  and  $P_v^{(N)}$  are mutually absolutely continuous and

$$(2.11) \quad \frac{dP_{u(\cdot), u}^{(N)}}{dP_v^{(N)}} = \exp(H_T^{(N)}(\omega)), \text{ where}$$

$$(2.12) \quad H_T^{(N)}(\omega) = \sum_{i=1}^N \int_{[0, T] \times E} \{\log q_{u(s-)}^{\pi(\eta_{x^{(N)}})(s-)}(\omega_i(s-), y)\} N(ds, dy; \omega_i) \\ - \sum_{i=1}^N \int_0^T \{ \|\pi(\eta_{x^{(N)}})(s)\| - \|u(s)\| \} ds.$$

**3. The Main Result.** Let  $X$  be a Hausdorff topological space.  $\{P_N\} \subset M_1(X)$ .  $\{a_N\}$  is a sequence of positive numbers tending to  $\infty$ .  $I$  is a function from  $X$  to  $[0, \infty)$ .

**Definition 3.1**  $(X, P_N, a_N)$  is said to be a large deviation system with action functional  $I$  if

i). for every open subset  $G$  of  $X$

$$(3.1) \quad \liminf_{N \rightarrow \infty} a_N^{-1} \log P_N(G) \geq - \inf_{x \in G} I(x).$$

ii). for every closed subset  $F$  of  $X$

$$(3.2) \quad \limsup_{N \rightarrow \infty} a_N^{-1} \log P_N(F) \leq - \inf_{x \in F} I(x).$$

iii). the level sets  $\{x \in X : I(x) \leq s\}$  are compact for all  $s \geq 0$ .

Note: The infimum is replaced by  $+\infty$  if the set is empty.

Let  $I_u^{u(\cdot)} : \mathcal{M}_\infty \rightarrow [0, \infty]$  be defined as

$$(3.3) \quad I_u^{u(\cdot)}(P) = \sup_{\Phi \in C_b(D)} \left\{ \int \Phi dP - \log E^{P_{u(\cdot), u}}(e^\Phi) \right\}$$

where  $P \in \mathcal{M}_\infty$  and  $C_b(D)$  is the set of all bounded continuous functions on  $D$ . Let  $I_u = I_u^{u(P)}(P)$ ,  $\mathcal{P}_u^{(N)}$  be the image of  $\mathcal{P}_u^{(N)}$  under the map  $\eta_{x^{(N)}, (\cdot)}$ . We have the following main result

**Theorem 3.3** Let  $u \in M_1(E)$  satisfy  $\int_E e^{\psi(x)} u(dx) < \infty$ , under conditions (2.1)–(2.5)  $(\mathcal{M}_\infty, \mathcal{P}_u^{(N)}, N)$  is a large deviation system with action functional  $I_u$ .

The main idea of the proof is the following. First by Sanov's theorem we have

**Lemma 3.4**  $(\mathcal{M}, \mathcal{P}_{u(\cdot), u}^{(N)}, N)$  is a large deviation system with action functional  $I_u^{u(\cdot)}(\cdot)$ .

This combined with theorem 2.2 yields the following estimates:

**Lemma 3.5** For any  $r > 0$ , there exists an  $R_0 > 0$  such that for all  $R \geq R_0$  and  $N \geq 1$ , we have

$$(3.4) \quad \mathcal{P}_u^{(N)}\{\mathcal{M}_\infty \setminus \mathcal{M}_R\} \leq \exp(-Nr).$$

**Lemma 3.6** For any  $R > 0, C > 0$ , there exists a compact set  $K_R \subset \mathcal{M}_\infty$  such that

$$(3.5) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(\mathcal{M}_R \setminus K_R) < -C.$$

**Lemma 3.7** For every  $\bar{P} \in \mathcal{M}_\infty$  and  $\gamma > 0$ , there exists an open neighborhood  $V$  of  $\bar{P}$  in  $\mathcal{M}$  such that

$$(3.6) \quad \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(V) \leq -I_u(\bar{P}) + \gamma.$$

**Lemma 3.8** For every  $\bar{P} \in \mathcal{M}_\infty$  and any open neighborhood  $V$  of  $\bar{P}$ , we have

$$(3.7) \quad \liminf_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(V) \geq -I_u(\bar{P}).$$

**Lemma 3.9** Let  $u \in M_1(E)$  satisfy  $\int_E e^{v(x)} u(dx) < \infty$ . Under assumptions (2.1)-(2.5), we have that for every  $r > 0$ ,  $\Phi_u(r) = \{P \in \mathcal{M}_\infty : I_u(P) \leq r\}$  is compact.

Theorem 3.3 is a direct result of these lemmas.

**Acknowledgment.** This research was financially supported by a scholarship from the Faculty of Graduate Studies and Research of Carleton University and the NSERC operating grant of D.A.Dawson. These results are part of my Ph.D. dissertation. I wish to thank my supervisor Professor D.A.Dawson for his guidance, patience and very helpful suggestions.

## References

- [1] Dawson, D.A. and Gärtner, J.(1987). Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, 20:247-308.
- [2] Gärtner, J.(1977). On large deviations from the invariant measure. *Theory Probab. Appl.*, 22:24-39.
- [3] Léonard, C.(1990). Large deviations in the dual of a normed vector space. *Prépublication de l'Université d'Orsay*.
- [4] Léonard, C.(1990). On large deviations for particle systems associated with spatially homogeneous Boltzmann type equations. preprint.
- [5] Schlögl, F.(1972). Chemical reaction models for non-equilibrium phase transitions. *Z.Physik*, 253:147-161.

Department of Mathematics and Statistics

Carleton University

Ottawa, Ontario K1S 5B6

CANADA.

---

Received February 27, 1992

# Extinction of the Multilevel $M(M(R^d))$ -Valued Branching Diffusion Process

Yadong Wu

*Presented by Donald A Dawson, F.R.S.C.*

## Abstract

In this paper we consider a multilevel  $M(M(R^d))$ -valued branching diffusion process. The long time behavior of the process is studied. The main result is that if  $d \leq 4$  then the two level  $M(M(R^d))$ -valued process suffers local extinction.

## 1 Introduction

In this paper, we study the multilevel branching diffusion process which was formulated by Dawson and Hochberg in 1991. We describe the continuous limiting process as a  $M(M(R^d))$ -valued process  $Y(t)$ . We investigate the long time behaviour of the process  $Y(t)$  and obtain an analogue of the result in the one level case (see, Dawson (1977) [3]) that if the initial value is given by Lebesgue measure on  $R^d$  then existence or non-existence of a non-trivial limiting distribution of the one level classical critical branching Brownian motion is equivalent to transience or recurrence of the spatial motion. The main result of this paper is to prove that:

*The  $M(M(R^d))$ -valued process  $Y(t)$  suffers local extinction if  $d \leq 4$ .*

The method of proof for local extinction which we use here was suggested by Dawson.

## 2 Description of the Model

We define  $\hat{R}^d = R^d \cup \{\tau\}$ ,  $\tau$  being an isolated adjoined point. Let  $M_\rho(\hat{R}^d)$  be the space of  $\rho$ -tempered measures introduced by Iscoe (1983) with the  $\rho$ -vague topology, that is, the smallest topology making the maps  $\mu \rightarrow \langle \phi, \mu \rangle$  continuous for  $\phi \in C_c(R^d) \cup \{\phi_\rho\}$ , where  $d < \rho \leq d + 2$  and for  $x \in R^d$

$$(2.1) \quad \phi_\rho(x) = \frac{1}{1 + |x|^\rho}, \quad \text{and} \quad \phi_\rho(\tau) = 1.$$

Let

$$M_\rho^2(\dot{R}^d) \equiv \{ \nu \in M(M_\rho(\dot{R}^d)) : \int \int \phi_\rho(x) \mu(dx) \nu(d\mu) < \infty \}$$

and endow  $M_\rho^2(\dot{R}^d)$  with the smallest topology which make continuous the maps

$$\nu \longrightarrow \int \int \phi(x) \mu(dx) \nu(d\mu)$$

for all  $\phi \in C_c(\mathbb{R}^d) \cup \{ \phi_\rho \}$ .

**Remark 2.1** *The measure  $\delta_{\delta_x} dx \in M(M(\mathbb{R}^d))$  is very important for the next section. We know that  $\delta_{\delta_x}$  is an atomic measure on  $M(\mathbb{R}^d)$  with the mass at the atomic measure  $\delta_x$ . So the measure  $\delta_{\delta_x} dx$  can be considered as  $\delta_{\delta_x}$  in which the location  $x$  of the atomic measure  $\delta_{\delta_x}$  is given by  $d$ -dimensional Lebesgue measure in  $\mathbb{R}^d$ .*

The basic process in this paper is a  $M_\rho^2(\mathbb{R}^d)$ -valued process  $Y(t)$ . The reason that we consider  $M_\rho^2(\mathbb{R}^d)$  as the state space of the process is that we wish to allow measure  $\delta_{\delta_x} dx$  to be initial value of process  $Y(t)$  and it is easy to show that  $\delta_{\delta_x} dx \in M_\rho^2(\mathbb{R}^d)$ .

We will not discuss in detail the construction of the process in this paper but explain briefly how it can be characterized as the continuous limit of a two level branching diffusion particle system.

We consider a two level branching random field. A "two level" system consists of two different level particles. Each level two particle (i.e., superparticle) is a collection of level one particles (i.e., particle) in  $\mathbb{R}^d$ . We say that a superparticle is of size  $i$  if it consists of exactly  $i$  particles. We suppose that after an exponentially distributed time, each particle undergoes a level one binary branching process with the branching rate  $\gamma_1$  in which case we note that the total number of superparticles is unchanged. We also assume that each superparticle performs level two binary branching with the branching rate  $\gamma_2$ , that is, either dies or produces a copy, each with probability one-half, after an exponentially distributed holding time with the parameter  $\gamma_2$ . During the holding times of both level one and level two, each particle moves in  $\mathbb{R}^d$  according to the  $d$ -dimensional Brownian motion. The entire system can be represented as a random atomic measure on  $N(\mathbb{R}^d)$ :

$$(2.2) \quad Y_\alpha(t) \equiv \sum_{i=1}^{\infty} \sum_{k=1}^{n_i} \delta_{\sum_{r=1}^i \delta_{x_{i,k,r}(t)}} + n_0(t) \delta_{\delta_\phi}$$

where  $x_{i,k,r}(t)$  denotes the location in  $\mathbb{R}^d$  of the  $r$ th particle in the  $k$ th superparticle  $X_{i,k}(t)$  of size  $i$  at time  $t$ ,  $n_i$  is the number of the superparticles of size  $i$  at time  $t$  and  $n_0(t)$  denotes the number of the null superparticles at time  $t$  where the term "null superparticle" means that it does not contain any particles.

Consider the test function  $F(\nu)$  on  $M_\rho^2$  of the following form:

$$(2.3) \quad F(\nu) = f(\langle\langle h_1(\langle h_2, \cdot \rangle), \nu \rangle\rangle)$$

where  $\nu \in M_\rho^2$ ,  $h_1, f \in C_b^2(R)$  and  $h_2 \in C_c^2(R^d)$  and  $\langle\langle g(\mu), \nu \rangle\rangle \equiv \int g(\mu)\nu(d\mu)$ .

To study the continuous limit of the two level branching diffusion particle systems, we rescale the system as follows: If  $B(M_\rho(\hat{R}^d))$  denotes the Borel  $\sigma$ -algebra of  $M_\rho(\hat{R}^d)$ , (ref. [8], p90), then for  $A \in B(M_\rho(\hat{R}^d))$  and  $t > 0$ , we define

$$Y_n(t, A) = \frac{1}{n} Y_n(nt, A_n) \quad \text{where } A_n = \{\mu : \frac{1}{n}\mu \in A\}$$

We can show that when  $n \rightarrow \infty$ ,  $Y_n(t)$  converges weakly in  $D_{[0, \infty)}(M_\rho^2(R^d))$  to a  $M_\rho^2$ -valued process  $Y(t)$  which can be characterized as the unique solution of the martingale problem for the limiting generator  $G_c^{(2)}$  which is given by

$$(2.4) \quad G_c^{(2)} F(\nu) = \langle\langle \mathcal{L}F'(\nu, \cdot), \nu \rangle\rangle + \gamma_2 \langle\langle F''(\nu), \delta_{\mu_1}(d\mu_2)\nu(d\mu_1) \rangle\rangle.$$

In (2.4)  $\mathcal{L}$  denotes the generator of the  $M_\rho(R^d)$ -valued branching process, i.e.

$$(2.5) \quad \begin{aligned} \mathcal{L}F'(\nu, \mu) &= \mathcal{L}_1 F'(\nu, \mu) + \mathcal{L}_2 F'(\nu, \mu) \\ &= f'(\langle\langle h_1(\langle h_2, \cdot \rangle), \nu \rangle\rangle) h_1'(\langle h_2, \mu \rangle) \langle \Delta h_2, \mu \rangle \\ &\quad + \gamma_1 f'(\langle\langle h_1(\langle h_2, \cdot \rangle), \nu \rangle\rangle) h_1''(\langle h_2, \mu \rangle) \langle h_2^2, \mu \rangle \end{aligned}$$

where  $\Delta$  is the  $d$ -dimensional Laplacian,

$$(2.6) \quad \mathcal{L}_1 F'(\nu, \mu) = \int \Delta \frac{\delta F'(\nu, \mu)}{\delta \mu(x)} \mu(dx)$$

and

$$(2.7) \quad \mathcal{L}_2 F'(\nu, \mu) = \gamma_1 \int \int \frac{\delta^2 F'(\nu, \mu)}{\delta \mu(x) \delta \mu(y)} \delta_x(dy) \mu(dx)$$

where

$$(2.8) \quad \begin{aligned} F'(\nu, \mu) &= \frac{\partial F(\nu)}{\partial \nu(\mu)} = \frac{d}{d\epsilon} [F(\nu + \epsilon \delta_\mu)]_{\epsilon=0} \\ &= f'(\langle\langle h_1(\langle h_2, \cdot \rangle), \nu \rangle\rangle) h_1(\langle h_2, \mu \rangle) \end{aligned}$$

The Laplace transition functional of the  $M_\rho^2$ -valued process  $Y(t)$  is given by

$$(2.9) \quad \begin{aligned} L_{t,\nu}(H) &= E[\exp(-\int_{M_\rho(R^d)} H(\mu) Y(t, d\mu) \mid Y(0) = \nu)] \\ &= \exp\{-\int u(t, \mu)\nu(d\mu)\} \end{aligned}$$

where  $u(t, \mu)$  is a solution of the weak form of the following integral equations :

$$(2.10) \quad \begin{aligned} u(t, \mu) &= T_t u(0, \mu) - \gamma_2 \int_0^t [T_{t-s} u^2(s, \cdot)](\mu) ds \\ u(0, \mu) &= H(\mu) \end{aligned}$$

and  $H(\mu) = f(\langle \phi, \mu \rangle)$ ,  $\phi \in C_c(R^d)$  and  $f \in C_b(R)$ .  $T_t$  denotes the semigroup of operators of the one level  $M_\rho(R^d)$ -valued branching process with the generator  $\mathcal{L}$ .

### 3 Extinction

**Definition 3.1** We shall say that the two level  $M_\rho^2(\mathbb{R}^d)$ -valued process  $Y(t)$  suffers local extinction if for every compact set  $C$  in  $\mathbb{R}^d$ ,  $\epsilon_1 > 0$  and  $\epsilon_2 > 0$ ,

$$(3.11) \quad \lim_{t \rightarrow \infty} P(\langle\langle 1_{K_{C,\epsilon_2}}, Y(t) \rangle\rangle > \epsilon_1) = 0$$

where  $K_{C,\epsilon_2}$  is given by

$$K_{C,\epsilon_2} := \{\mu \in M_\rho(\mathbb{R}^d) : \mu(C) > \epsilon_2\}$$

and  $1_{K_{C,\epsilon_2}}$  is the indicator function of the set of  $K_{C,\epsilon_2}$ .

Let  $u(t, \mu)$  denote the solution of the equation (2.10) with initial values  $u(0, \mu) = \langle \phi, \mu \rangle$ ,  $\nu_0 = \delta_{\delta_x} dx$  and define  $T_t^*$  by

$$(3.12) \quad \int [1 - \exp\{-\langle \phi, \cdot \rangle\}](\mu) T_t^* \nu_0(d\mu) \equiv \int T_t [1 - \exp\{-\langle \phi, \cdot \rangle\}](\mu) \nu_0(d\mu)$$

where  $T_t$  is the semigroup of operators of the one level  $M_\rho(\mathbb{R}^d)$ -valued branching process with generator  $\mathcal{L}$ . Then we can show that for  $d \geq 3$

$$(3.13) \quad T_t^* \nu_0 \implies R_\infty$$

where  $R_\infty$  is the canonical equilibrium measure for the one level measure branching process. Our main theorem requires the following lemma which was given by Dawson and Perkins (1991).

**Lemma 3.1** Let  $\mathfrak{R}_d$  denotes the Borel  $\sigma$ -algebra in  $\mathbb{R}^d$ . Then for each  $B \in \mathfrak{R}_d$  and  $d \geq 3$ , we have

$$(3.14) \quad \begin{aligned} R_\infty(\{\mu : \frac{1}{t} \mu(\sqrt{t}B) > a\}) \\ = t^{\frac{d-2}{2}} R_\infty(\{\mu : \mu(B) > a\}) \end{aligned}$$

where  $R_\infty$  is the canonical equilibrium measure as above.

**Proof.** See D. Dawson and E. Perkins (1991) [2].

□

**Theorem 3.1** If  $d \leq 4$  and  $Y_0 = \nu_0 \equiv \delta_{\delta_x} dx$ , then the two level process  $Y(t)$  suffers local extinction.

**Outline of the Proof :** Since the one level  $M_\rho(R^d)$ -valued process with the generator  $\mathcal{L}$  suffers local extinction for  $d \leq 2$ , it is clear that the theorem is true in this case. So we only need to show the theorem for  $d = 3$  and  $d = 4$ . In the remainder of this section, we always assume that  $d \geq 3$ . Noting that for each  $\mu \in M(R^d)$ , we have

$$(3.15) \quad \gamma_2 \int \int_0^t T_{t-s}[u(s)]^2(\mu) ds R_\infty(d\mu) \leq \int T_t u(0)(\mu) R_\infty(d\mu) < K < \infty$$

for all  $t$ , therefore by the invariance of  $R_\infty$  we obtain

$$(3.16) \quad \lim_{t \rightarrow \infty} \int \int_0^t u^2(s, \mu) ds R_\infty(d\mu) < \infty.$$

On the other hand, if

$$(3.17) \quad \lim_{t \rightarrow \infty} \int u(t, \mu) R_\infty(d\mu) = c_1 > 0$$

then we obtain

$$(3.18) \quad \lim_{t \rightarrow \infty} \int \int_0^t u^2(s, \mu) ds R_\infty(d\mu) = \infty.$$

Thus we obtain

$$(3.19) \quad \lim_{t \rightarrow \infty} \int u(t, \mu) R_\infty(d\mu) = 0$$

by contradiction. Therefore, together with the fact that for each  $t > 0$

$$(3.20) \quad \begin{aligned} \limsup_{t_1 \rightarrow \infty} \int u(t_1, \mu) \nu_0(d\mu) &= \limsup_{t_1 \rightarrow \infty} \int u(t_1 + t, \mu) \nu_0(d\mu) \\ &\leq \int T_t u(0, \mu) R_\infty(d\mu) - \gamma_2 \int \int_0^t T_{t-s}[u(s)]^2(\mu) ds R_\infty(d\mu) \\ &= \int u(t, \mu) R_\infty(d\mu) \end{aligned}$$

we have

$$(3.21) \quad \limsup_{t_1 \rightarrow \infty} \int u(t_1, \mu) \nu_0(d\mu) = 0.$$

as claimed. □

**Acknowledgement.** This paper is a part of the author's doctoral thesis which was written under the supervision of Professor D.Dawson. The author would like to thank Professor Dawson for his continued advice and guidance in this project.

## References

- [1] Dawson, D. and Hochberg, K. (1991). A multilevel branching model. *Adv. Appl. Prob.*, 23:701-715.

- [2] Dawson, D. A. and Perkins. E. A. (1991). Historical processes. *Memoir. Amer. Math. Soc.*, 454.
- [3] Dawson, D. A. (1977). The critical measure diffusion. *Z. Wahrsch. Verw. Gebiete*, 40:125–145.
- [4] Dawson, D. A. (1978). Limit theorem for interaction free geostochastic systems. *Colloquia Math. Soc. J. Bolyai*, 24:27–47.
- [5] Dawson, D. A. (1986). Measure-valued stochastic processes: construction, qualitative behavior and stochastic geometry. pages 69–93, Springer-Verlag. *Lecture Notes in Mathematics*.
- [6] Etheridge, A. (1990). Asymptotic behaviour of measure-valued critical branching processes. *To appear*.
- [7] Ethier, S.N. and Kurtz, T.G.(1985). *Markov processes: characterization and convergence*. Wiley, New York.
- [8] Iscoe, I. (1986). A weighted occupation time for a class of measure-valued critical branching Brownian motion. *Probab. Th. Rel. Fields*, 71:85–116.
- [9] Wu, Y. (1992) . Dynamic particle systems and multilevel measure branching processes, Ph.D. thesis. *Carleton University*.
- [10] Zähle, U. (1988). Self-similar random measures I. Notion, carrying Hausdorff dimension and hyperbolic distribution. *Probab. Th. Rel. Fields*, 80:79–100.

Department of Mathematics and Statistics  
Carleton University  
Ottawa, Ontario K1S 5B6  
Canada

---

Received February 27, 1992

**CONSTRUCTION D'UN EXEMPLE D'ANNEAU DE JAFFARD  
LOCAL FACTORIEL NON NOETHERIEN DE DIMENSION 2  
ET DE CARACTERISTIQUE 0.**

**SOUAD AMEZIANE HASSANI**

*Presented by P. Ribenboim, F.R.S.C.*

**S 1 INTRODUCTION.**

Dans [2, Proposition 2.12 (b)] A. Bouvier et S. Kabbaj ont énoncé que pour tout  $n \geq 2$ , il existe un anneau de Jaffard, factoriel, local, non noethérien, de dimension  $n$  et de caractéristique 0. Pour ce faire, ils ont utilisé une construction de J.W. Brewer, D.L. Costa et E. Lady [3, Theorem D], où l'anneau en question était de dimension  $n \geq 3$ , en omettant ainsi le cas où  $n = 2$ . La question restait donc ouverte pour  $n = 2$ .

Dans ce papier, en s'inspirant d'un exemple dans [3, p. 306-307], nous allons construire un anneau de Jaffard, factoriel, local, non noethérien, de dimension 2 et de caractéristique 0.

Soient  $A$  un anneau et  $G$  un groupe abélien noté additivement. L'anneau de

groupe de  $G$  sur  $A$ , noté  $A[G]$  est l'ensemble des fonctions définies de  $G$  vers  $A$  nulles presque partout.

En fait,  $A[G]$  est isomorphe au  $A$ -module libre  $\bigoplus_{g \in G} AX^g$ , où  $X^g X^{g'} = X^{g+g'}$  pour tout  $g, g' \in G$ .

Notons que  $A[G]$  est intègre si et seulement si  $A$  est intègre et  $G$  est sans torsion [6, Theorem 15.8].

## S 2 EXEMPLE.

Ce paragraphe est consacré à l'exemple d'anneau de Jaffard, local, factoriel, non noethérien, de dimension 2 et de caractéristique 0.

On reprend entièrement la construction faite dans [3, p. 306-307]. Soient  $p$  un entier naturel premier,  $\mathbb{Z}[1/p] = \{ a/p^n; a \in \mathbb{Z}, n \in \mathbb{N} \}$ ,  $G = \mathbb{Z}[1/p] \oplus \mathbb{Z}[1/p] \oplus \mathbb{Z}[1/p]$ ,  $H = \mathbb{Z}[1/p] \oplus \mathbb{Z}[1/p]$  un sous groupe de  $G$  et  $G_n = H \oplus p^{-n}\mathbb{Z}$ .

Il est clair que  $G = \bigcup_{n \geq 0} G_n$  et que  $\mathbb{Z}[G] = \bigcup_{n \geq 0} \mathbb{Z}[G_n] = \bigcup_{n \geq 0} \mathbb{Z}[H][p^{-n}\mathbb{Z}]$  ( $\mathbb{Z}[G_n] = \mathbb{Z}[H][p^{-n}\mathbb{Z}]$ , d'après [8, Theorem 7.1]).

Soit  $\delta_1, \delta_2, \dots$  une suite d'éléments de  $\mathbb{Z}[H] - \{0\}$  telle que  $cl(\delta_n)$  n'est pas nulle dans  $\mathbb{Z}/p\mathbb{Z}[H]$  où  $cl(\delta_n)$  désigne la classe de  $\delta_n$  modulo  $p\mathbb{Z}[H]$  et  $cl(\delta_n)^p = cl(\delta_{n-1})$  pour tout  $n \in \mathbb{N}$ .

(la suite de terme général  $\delta_n = \chi(p^{-n}, 0) + \chi(0, p^{-n})$  convient).

$P_n = (p, \delta_n + \gamma p^{-n})$  est un idéal premier de  $\mathbb{Z}[H][\gamma p^{-n}, \gamma^{-p^{-n}}]$  qui est identifié à  $\mathbb{Z}[H][p^{-n}\mathbb{Z}]$ , et  $P = \bigcup_{n \geq 0} P_n$  est un idéal premier de  $\mathbb{Z}[G]$ .

Soit  $A = \mathbb{Z}[G]_P$ ; il est établi dans [3, p. 306-307] que :

$A$  est factoriel.

- . A est non noethérien.
- . A est local, de dimension 2 et de caractéristique 0.

Nous affirmons alors que :

- . A est un anneau de Jaffard.

Comme  $\dim \mathbb{Z}[G]_P = 2$ , on va prouver que  $\dim \mathbb{Z}[G]_P[X_1, X_2] = 2.2 = 4$ .

Il suffit d'établir que  $\dim \mathbb{Z}[G]_P[X_1, X_2] \leq 4$ .

$\mathbb{Z}[G]_P[X_1, X_2] = S^{-1}(\mathbb{Z}[G][X_1, X_2])$ , où  $S = \mathbb{Z}[G] - P$ . Donc

$\text{spec}(\mathbb{Z}[G]_P)$  est homéomorphe à  $\{Q \in \text{spec}(\mathbb{Z}[G][X_1, X_2]) / Q \cap S = \emptyset\}$ .

Soit  $Q \in \text{spec}(\mathbb{Z}[G]_P[X_1, X_2])$  tel que  $Q \cap S = \emptyset$ . Posons  $q = Q \cap \mathbb{Z}[G]$ , et alors  $Q \cap S = \emptyset$  entraîne que  $q \subseteq P$ .

On sait que  $\text{ht } Q \leq 2 + \text{ht } q[X_1, X_2]$  ([4, Theorem 1]),

$$\leq 2 + \text{ht } P[X_1, X_2].$$

Il suffit donc de prouver que  $\text{ht } P[X_1, X_2] \leq 2$ . On a :  $P = \cup_{n \geq 0} P_n$  et par suite,

$$P[X_1, X_2] = \cup_{n \geq 0} P_n[X_1, X_2].$$

Soit l'idéal premier  $Q_0 = (p, \delta_0 + Y)$  dans  $\mathbb{Z}[H][Y]$ .

$\text{ht } P_0[X_1, X_2] = \text{ht } Q_0[X_1, X_2]$  car  $P_0[X_1, X_2] = T^{-1}Q_0[X_1, X_2]$  où  $T = \{Y^n / n \in \mathbb{N}\}$ .

Nous allons établir que  $\text{ht } Q_0[X_1, X_2] = 2$ . Sinon, supposons que

$$\text{ht } Q_0[X_1, X_2] > 2 \quad (\text{ht } Q_0 = 2).$$

Dans  $\mathbb{Z}[H][Y, X_1, X_2]$ ,  $Q_0[X_1, X_2] \cap \mathbb{Z}[H] = p\mathbb{Z}[H]$  et

$$\text{ht } p\mathbb{Z}[H][Y, X_1, X_2] + \dim \mathbb{Z}/p\mathbb{Z}[H][Y, X_1, X_2] \leq \dim \mathbb{Z}[H][Y, X_1, X_2].$$

$\mathbb{Z}/p\mathbb{Z}[H][Y, X_1, X_2]$  est isomorphe à  $\mathbb{Z}/p\mathbb{Z}[Y, X_1, X_2][H]$  et  $\mathbb{Z}[H][Y, X_1, X_2]$

est isomorphe à  $\mathbb{Z}[Y, X_1, X_2][H]$  ( Il s'agit des isomorphismes canoniques).

Comme  $\text{rang } H = 2$ , alors  $\dim \mathbb{Z}/p\mathbb{Z}[Y, X_1, X_2][H] = 5$  et

$$\dim \mathbb{Z}[Y, X_1, X_2][H] = 6 \quad [7, \text{Corollary 2}].$$

Par suite,  $\text{ht } p\mathbb{Z}[H][Y, X_1, X_2] + 5 \leq 6$ . Il en résulte que  $\text{ht } p\mathbb{Z}[H][Y, X_1, X_2] \leq 1$ .  
 $\mathbb{Z}[H][Y, X_1, X_2]$  est intègre et  $p\mathbb{Z}[H][Y, X_1, X_2] \neq 0$  donc  $\text{ht } p\mathbb{Z}[H][Y, X_1, X_2] = 1$ .

D'après [4, Corollary 3] la hauteur de  $Q_0[X_1, X_2]$  peut être réalisée comme la longueur d'une chaîne spéciale, et comme  $\text{ht } p\mathbb{Z}[H][Y, X_1, X_2] = 1$ , il existe alors  $Q' \in \text{spect } \mathbb{Z}[H][Y, X_1, X_2]$  tel que :  $p\mathbb{Z}[H][Y, X_1, X_2] \subset Q' \subset Q_0[X_1, X_2]$ . Absurde, puisque  $Q_0[X_1, X_2]$  est un supérieur de  $p\mathbb{Z}[H][X_1, X_2]$  dans  $\mathbb{Z}[H][X_1, X_2][Y]$  (qu'on identifie avec  $\mathbb{Z}[H][Y, X_1, X_2]$ ). Donc,  $\text{ht } P_0[X_1, X_2] = 2$ .

Pour tout  $n \in \mathbb{N}$ , la construction de  $P_n$  dans  $\mathbb{Z}[H][\gamma p^{-n}, \gamma^{-p^{-n}}]$  est similaire à celle de  $P_0$  dans  $\mathbb{Z}[H][Y, Y^{-1}]$ . Par conséquent pour tout  $n \in \mathbb{N}$ ,  $\text{ht } P_n[X_1, X_2] = 2$ .

On a  $\text{ht } P = 2$  [3, p. 306-307], donc  $\text{ht } P[X_1, X_2] \geq 2$ .

Supposons que  $\text{ht } P[X_1, X_2] > 2$ , il existe alors  $Q_1, Q_2, Q_3 \in \text{spect}(\mathbb{Z}[G][X_1, X_2])$  tels que :  $(0) \subset Q_1 \subset Q_2 \subset Q_3 = P[X_1, X_2]$  (\*).

Solent  $x_1, x_2, x_3 \in P[X_1, X_2]$  tels que  $x_1 \in Q_1 - (0)$ ,  $x_2 \in Q_2 - Q_1$  et  $x_3 \in Q_3 - Q_2$ .

$\mathbb{Z}[G] = \bigcup_{n \geq 0} L_n$ , où  $L_n = \mathbb{Z}[H][\gamma p^{-n}, \gamma^{-p^{-n}}]$ . Donc,  $\mathbb{Z}[G][X_1, X_2] = \bigcup_{n \geq 0} L_n[X_1, X_2]$ , il existe alors  $n_0 \in \mathbb{N}$  tel que :  $x_1, x_2, x_3 \in L_{n_0}[X_1, X_2]$ .

$P[X_1, X_2] \cap L_{n_0}[X_1, X_2] = (P \cap L_{n_0})[X_1, X_2]$ . Or,  $P \cap L_{n_0} = P_{n_0}$  :

En effet,  $P_{n_0} \subseteq P \cap L_{n_0}$ . Inversement, soit  $z \in P \cap L_{n_0}$ , comme  $P = \bigcup_{n \geq 0} P_n$ , il existe  $m > n_0$  tel que  $z \in P_m$ ; et donc  $cl(z) \in P_m/p\mathbb{Z}[H][\gamma p^{-m}, \gamma^{-p^{-m}}]$ , par suite,  $cl(z) = cl(f)(cl(\delta_m) + \gamma p^{-m})$ , où  $f \in L_m = \mathbb{Z}[H][\gamma p^{-m}, \gamma^{-p^{-m}}]$ .

$cl(z)^p = cl(f)^p (cl(\delta_{m-1}) + \gamma p^{-(m-1)})$ ,  $cl(f)^p \in \mathbb{Z}/p\mathbb{Z}[H][\gamma p^{-(m-1)}, \gamma^{-p^{-(m-1)}}]$ .

Donc,  $cl(z)^p \in P_{m-1}/p\mathbb{Z}[H][\gamma p^{-(m-1)}, \gamma^{-p^{-(m-1)}}]$ . En répétant le même

processus  $(m-n_0)$  fois, on obtient :

$$cI(z)^{p^{m-n_0}} = cI(f)^{p^{m-n_0}} (cI(\delta_{n_0}) + \gamma p^{-n_0}) \in P_{n_0} / p \mathbb{Z}[H][\gamma p^{-n_0}, \gamma^{-p^{-n_0}}].$$

Et comme  $z^{p^{m-n_0}} \in L_{n_0}$  ( $z \in L_{n_0}$ ), alors  $z \in P_{n_0}$ . D'où,  $P \cap L_{n_0} = P_{n_0}$ .

Par suite,  $P[X_1, X_2] \cap L_{n_0}[X_1, X_2] = P_{n_0}[X_1, X_2]$ .

Il en résulte alors par contraction de la suite (\*) sur  $L_{n_0}[X_1, X_2]$  que

ht  $P_{n_0}[X_1, X_2] \geq 3$  (absurde).

D'où  $A = \mathbb{Z}[G]_p$  est l'exemple désiré. ■

Remarquons que pour pouvoir appliquer [5, Lemma 2.1], on doit vérifier que  $\mathbb{Z}[G][X_1, X_2]_p[X_1, X_2] = \cup_{n \geq 0} \mathbb{Z}[G_n][X_1, X_2]_{p_n}[X_1, X_2]$  pour être dans les conditions du lemme. Une telle vérification est loin d'être immédiate.

## REFERENCES

- [1] D.F. ANDERSON – A. BOUVIER – D.E. DOBBS – M. FONTANA – S. KABBAJ.

On Jaffard domains.

Expositiones Mathematicae 6 (1988) 145-175.

- [2] A. BOUVIER.– S. KABBAJ.

Examples of Jaffard domains.

Journal of pure and applied algebra 54 (1988) 155- 165.

- [3] **J.W. BREWER – D.L. COSTA – E.L. LADY.**  
Prime ideals and localization in commutative group rings.  
Journal of algebra 34 (2) (1975) 300-308.
- [4] **J.W. BREWER – P.R. MONTGOMERY –E.A. RUTTER – W.J. HEINZER.**  
Krull dimension of polynomial rings.  
Springer Lecture Notes Math. 311 (1972) 26-45.
- [5] **D.E. DOBBS – M. FONTANA.– S. KABBAJ.**  
Direct limits of Jaffard domains and S-domains.  
Commentarii Mathematici Universitatis Sancti-Pauli 39 (2) (1990)  
143-155.
- [6] **R. GILMER.**  
Multiplicative ideal theory.  
Dekker, New York 1972.
- [7] **R. GILMER.**  
A two-dimensional non-noetherian factorial ring.  
Proc. Amerc. Math. Soc. 44 (1) (1974) 25-30.
- [8] **R. GILMER.**  
Commutative semigroup rings.  
Chicago Lectures Math. univ. Chicago. Press, Chicago 1984.

Département de Mathématiques et informatique.

Faculté des Sciences, Université S.M. Ben Abdellah

B.P. 1796 ( Atlas ), Fès - Maroc.

Received March 6, 1992

# The Diophantine equation $x^4 + dy^4 = z^p$

Nobuhiro Terai and Hiroyuki Osada

*Presented by P. Ribenboim, F.R.S.C.*

## 1 Introduction

Let  $p$  be an odd prime, and let  $d$  be a square-free positive integer. In [2], Powell proved that the Diophantine equation

$$x^4 \pm y^4 = z^p$$

has no integral solutions  $x, y, z$ , where  $(x, y) = 1$ ,  $p \nmid xyz$  and  $p \not\equiv \pm 1 \pmod{8}$  for the case  $x^4 + y^4 = z^p$ . In the present paper we shall prove that the Diophantine equation

$$x^4 + dy^4 = z^p$$

has no integral solutions  $x, y, z$  under some conditions, using the method of factoring  $z^p$  as  $(x^2 + \sqrt{-d}y^2)(x^2 - \sqrt{-d}y^2)$  over the imaginary quadratic field  $\mathbb{Q}(\sqrt{-d})$  as in [2].

This type of equation is treated in the book of Shorey and Tijdeman [3], especially chapter 8, 10. The Diophantine equation

$$x^4 + dy^4 = z^2$$

has been considered by Fermat, Euler, Nagell and others for a long time, using the method of infinite descent (see Mordell's book [1], chapter 4).

## 2 Lemma

In this section, we prove the following lemma which we use in the next section.

**Lemma.** *Let  $x, y$  be relatively prime integers such that  $x$  is odd,  $y$  is even,  $(x, d) = 1$  and  $p \nmid xy$ . Suppose that for some integers  $a, b$*

$$x^2 + \sqrt{-d}y^2 = (a + b\sqrt{-d})^p. \tag{1}$$

*If  $p \not\equiv \pm 1 \pmod{8}$ , then the equation (1) has no integral solutions  $x, y$ .*

*Proof.* It follows from (1) that

$$x^2 = a \sum_{j=0}^{(p-1)/2} \binom{p}{2j} a^{p-(2j+1)} b^{2j} (-d)^j = aA, \tag{2}$$

$$y^2 = b \sum_{j=0}^{(p-1)/2} \binom{p}{2j+1} a^{p-(2j+1)} b^{2j} (-d)^j = bB. \tag{3}$$

We first show that  $(a, A) = 1$  and  $(b, B) = 1$ . Since  $(x, y) = 1$  and  $x$  is odd,  $(a, b) = 1$  and  $a$  is odd. We also have  $(a, d) = 1$  and  $p \nmid ab$  from  $(x, d) = 1$  and  $p \nmid xy$ . Hence, since  $A \equiv pb^{p-1}(-d)^{p-1/2} \pmod{a^2}$  and

$$B \equiv pa^{p-1} \pmod{b^2}, \quad (4)$$

$(a, A) = 1$  and  $(b, B) = 1$ . Therefore from (2),(3) we obtain

$$a = \pm u^2, A = \pm U^2, b = \pm v^2 \text{ and } B = \pm V^2 \quad (5)$$

for some non-zero integers  $u, U, v$  and  $V$ .

We next show that  $b$  is even and  $B$  is odd. Let  $d$  be even. Then it is easily seen that  $B$  is odd, so  $b$  is even since  $y$  is even. Let  $d$  be odd and suppose that  $b$  is odd. Then we have

$$\sum_{j=0}^{(p-1)/2} \binom{p}{2j} \equiv A \equiv 1 \pmod{2}$$

and

$$\sum_{j=0}^{(p-1)/2} \binom{p}{2j+1} \equiv B \pmod{2},$$

so

$$2^p = \sum_{k=0}^p \binom{p}{k} = \sum_{j=0}^{(p-1)/2} \binom{p}{2j} + \sum_{j=0}^{(p-1)/2} \binom{p}{2j+1} \equiv 1 + B \pmod{2}.$$

Thus  $B$  is odd. But this contradicts that  $y^2 = bB$  is even. Hence  $b$  is even, so  $B$  is odd since  $(b, B) = 1$ . Therefore it follows from (4),(5) that

$$\pm V^2 \equiv pa^{p-1} \pmod{8},$$

so  $\pm 1 \equiv p \pmod{8}$  since  $V$  and  $a$  are odd. This contradicts our assumption  $p \not\equiv \pm 1 \pmod{8}$ , which completes the proof of Lemma.  $\square$

### 3 Theorems

We use Lemma in section 2 to prove the following theorems:

**Theorem 1** *Let  $p$  be an odd prime, let  $d \not\equiv 3 \pmod{4}$  be a square-free positive integer, and let  $h(-d)$  be the class number of the imaginary quadratic field  $\mathbb{Q}(\sqrt{-d})$ . If  $p \not\equiv \pm 1 \pmod{8}$  and  $p \nmid h(-d)$ , then the Diophantine equation*

$$x^4 + dy^4 = z^p \quad (6)$$

*has no integral solutions  $x, y, z$  with  $(x, y) = 1$ ,  $p \nmid xy$  and  $y$  even.*

**Theorem 2** *Let  $p$  be an odd prime and let  $m \equiv 1 \pmod{4}$  be a square-free positive integer. If  $p \equiv 3 \pmod{8}$  or  $p \equiv 5 \pmod{16}$  and  $p \nmid h(-2m)$ , then the Diophantine equation*

$$x^4 + 2my^4 = z^p \quad (7)$$

*has no integral solutions  $x, y, z$  with  $(x, y) = 1$ ,  $p \nmid xy$ .*

*Proof of Theorem 1.* Since  $(x, y) = 1$  and  $y$  is even, we see that  $x$  and  $z$  are odd,  $(x, z) = 1$  and  $(x, d) = 1$ . Now we show that  $x^2 + \sqrt{-d}y^2$  and  $x^2 - \sqrt{-d}y^2$  are relatively prime in  $\mathbb{Z}[\sqrt{-d}]$ , the ring of integers in  $\mathbb{Q}(\sqrt{-d})$ . Suppose that there is a prime ideal  $\mathfrak{p}$  in  $\mathbb{Z}[\sqrt{-d}]$  which divides both  $x^2 + \sqrt{-d}y^2$  and  $x^2 - \sqrt{-d}y^2$ . Then  $\mathfrak{p} \nmid 2$  since  $x^4 + dy^4$  is odd. Hence  $\mathfrak{p} \mid x$  and  $\mathfrak{p} \mid y\sqrt{-d}$ . If  $\mathfrak{p} \nmid \sqrt{-d}$ , then  $\mathfrak{p}$  divides both  $x$  and  $y$ , which is impossible since  $(x, y) = 1$ . Therefore  $\mathfrak{p} \mid \sqrt{-d}$ , so  $\mathfrak{p} \mid d$ , which contradicts that  $(x, d) = 1$  since  $\mathfrak{p} \mid x$ . Thus  $x^2 + \sqrt{-d}y^2$  and  $x^2 - \sqrt{-d}y^2$  are relatively prime in  $\mathbb{Z}[\sqrt{-d}]$ . Hence it follows from (6) that there is an ideal  $A$  in  $\mathbb{Z}[\sqrt{-d}]$  such that

$$(x^2 + y^2\sqrt{-d}) = A^p.$$

Since  $p \nmid h(-d)$  and  $d \not\equiv 3 \pmod{4}$ , we obtain

$$x^2 + y^2\sqrt{-d} = (a + b\sqrt{-d})^p \tag{8}$$

for some integers  $a, b$ . Therefore (8) has no integral solutions  $x, y$  by Lemma. This completes the proof of Theorem 1.  $\square$

*Proof of Theorem 2.* Suppose that our assumptions are all satisfied. Then by Theorem 1, we have only to prove that  $y$  is even in (7). On the contrary, suppose that  $y$  is odd. We use the notations in the proof of Lemma. Since  $(x, y) = 1$ , we see that  $x$  and  $z$  are odd,  $(x, z) = 1$  and  $(x, 2m) = 1$ . Therefore we obtain

$$x^2 + y^2\sqrt{-2m} = (a + b\sqrt{-2m})^p, \tag{9}$$

as in the proof of Theorem 1. Hence by the proof of Lemma, we have

$$A = \sum_{j=0}^{(p-1)/2} \binom{p}{2j} a^{p-(2j+1)} b^{2j} (-2m)^j, \quad x^2 = aA, \tag{10}$$

$$B = \sum_{j=0}^{(p-1)/2} \binom{p}{2j+1} a^{p-(2j+1)} b^{2j} (-2m)^j, \quad y^2 = bB, \tag{11}$$

$$A = \pm U^2, \quad B = \pm V^2. \tag{12}$$

Since  $x$  and  $y$  are odd, we see that  $a, A, b$  and  $B$  are all odd.

Suppose  $p \equiv 3 \pmod{8}$ . Then it follows from (10),(12) that

$$A = \pm U^2 \equiv a^{p-1} + \binom{p}{2} a^{p-3} b^2 (-2m) \pmod{8},$$

so  $\pm 1 \equiv 1 - 6m \pmod{8}$ . Thus  $m \equiv 0, 3 \pmod{4}$ , which is a contradiction since  $m \equiv 1 \pmod{4}$ . Hence  $y$  is even.

Suppose  $p \equiv 5 \pmod{16}$ . Then by (11) and (12), we have

$$B = \pm V^2 \equiv \binom{p}{1} a^{p-1} + \binom{p}{3} a^{p-3} b^2 (-2m) + \binom{p}{5} a^{p-5} b^4 (-2m)^2 \pmod{8},$$

so  $\pm 1 \equiv 5 - 20m + 4m^2 \pmod{8}$ . Thus  $4m \equiv 0, -2 \pmod{8}$ , which contradicts that  $m$  is odd. Therefore  $y$  is even. This completes the proof of Theorem 2.  $\square$

**Remark 1** If  $d = 1$  in (6), it follows unconditionally that  $p \nmid h(-1)$  and  $y$  is even. Therefore Theorem 1 gives a generalization of Powell's result for the case  $x^4 + y^4 = z^p$ .

**Remark 2** If some conditions of Theorem 1 are not satisfied, then the equation (6) sometimes has integral solutions. Such examples are given in the table below. We note that if  $(x, y) > 1$ , then there are many  $d$ 's such that the equation (6) has integral solutions. The condition  $(x, y) = 1$  is essential one for the equation (6) to have no integral solutions (see [2], p222).

Table. integral solutions of (6) which do not satisfy some conditions of Theorem 1

$x$	$y$	$z$	$p$	$d$	$h(-d)$
7	4	33	3	131	5
7	2	17	3	157	6
55	20	225	3	14	4
3	5	11	3	2	1
6	1	5	5	1829	40
96	96	24	7	53	6

**Acknowledgement.** The authors would like to thank the referee for his valuable suggestions.

## References

- [1] L.J. Mordell, *Diophantine Equations*, Academic Press, 1969.
- [2] B. Powell, Sur l'équation Diophantine  $x^4 \pm y^4 = z^p$ , *Bull. Sc. Math.*, 107(1983), 219-223.
- [3] T.Ñ. Shorey and R. Tijdeman, *Exponential diophantine equations*, Cambridge Tracts in Mathematics 87, Cambridge University Press, 1986.

Nobuhiro Terai  
 Department of Mathematics  
 School of Science and Engineering  
 Waseda University  
 Okubo, Shinjuku, Tokyo 169, Japan

Hiroyuki Osada  
 Department of Mathematics  
 Rikkyo University  
 Nishi-Ikebukuro, Tokyo 171, Japan

---

Received February 27, 1992

Mailing Addresses

1. V. Drensky  
Institute of Mathematics  
Bulgarian Academy of Sciences  
Acad. G. Bonchev Str., Block 8  
1113 Sofia, Bulgaria
2. S. Feng  
Department of Mathematics and Statistics  
Carleton University  
Ottawa, Ontario, Canada, K1A 5B6
3. S.A. Hassani  
Département de Mathématiques et Informatique  
Faculté des Sciences  
Université S.M. Ben Abdellah  
B.P. 1796 (Atlas), Fès-Maroc
4. P.R. Heath  
Department of Mathematics  
Memorial University of Newfoundland  
St. John's, Newfoundland, Canada, A1C 5S7
5. T.-Y. Lee  
Department of Mathematics  
University of Maryland  
College Park, MD 20742, U.S.A.
6. K.-Q. Liu  
Department of Mathematics  
University of Alberta  
Edmonton, Alberta, Canada, T6G 2G1
7. S.D. Morgera  
Department of Electrical Engineering  
McGill University  
Montréal, Québec, Canada, H3A 2A7
8. H. Osada  
Department of Mathematics  
Rikkyo University  
Nishi-Ikebukuro, Tokyo 171, Japan
9. B. Remillard  
Département de Mathématiques  
Université du Québec - Trois-Rivières  
Trois-Rivières, Québec, Canada, G1K 7P4
10. B. Szyszkowicz  
Department of Mathematics and Statistics  
Carleton University  
Ottawa, Ontario, Canada, K1S 5B6
11. N. Terai  
Department of Mathematics  
School of Science and Engineering  
Waseda University  
Okubo, Shinjuku, Tokyo 169, Japan
12. Y. Wu  
Department of Mathematics and Statistics  
Carleton University  
Ottawa, Ontario, Canada, K1S 5B6