SEPTEMBER/ SEPTEMBRE 2003

IN THIS ISSUE / DANS CE NUMÉRO

Published by / Publié par
The Academy of Science of the Royal Society of Canada
L'Académie des sciences de la Société royale du Canada
c/o Department of Mathematics and Statistics
Carleton University
Ottawa, Ontario, K1S 5B6
Tel/fax: 613-520-2166

C. R. Math. Rep. Acad. Sci. Canada Vol. **25**, (3), 2003 pp. 65–70

ON A DEGENERATE QUASILINEAR STOCHASTIC PARABOLIC EQUATION

M. SANGO

Presented by Vlastimil Dlab, FRSC

ABSTRACT. Using the splitting-up method, we establish a new existence result for an initial boundary value problem for the model degenerate stochastic quasilinear parabolic equation

$$d(|y|^{\alpha-2}y) - [\Delta y - c_1|y|^{2\mu-2}y]dt = c_2\sum_{l=1}^{d} |y|^{\sigma_l-1}ydW_t^l,$$

where W_t^l are one-dimensional Wiener processes defined on a complete probability space, α, μ and σ_l are some non negative numbers satisfying appropriate restrictions.

RÉSUMÉ. A l'aide d'une méthode de décomposition (éclatement d'opérateurs), nous établissons un nouveau résultat d'existence pour un problème aux limites initial pour l'équation stochastique parabolique non linéaire dégenérée

$$d(|y|^{\alpha-2}y) - [\Delta y - c_1|y|^{2\mu-2}y]dt = c_2\sum_{l=1}^{d} |y|^{\sigma_l-1}ydW_t^l,$$

où W_t^l sont des procéssus de Wiener unidimensionnels définis sur un espace de probabilités complet, α, μ et σ_l sont des réels non négatifs satisfaisant à des restrictions appropriées.

1. Introduction. Let D be a bounded domain in the Euclidean space $\mathbf{R}^n$ ($n \geq 2$) with a sufficiently smooth boundary ∂D. For $0 < T < \infty$, we denote by Q_T the cylinder $D \times (0,T)$. Let $(\boldsymbol{\Omega}, \mathbf{F}, \{\mathbf{F}_t\}_{0\leq t\leq T}, \mathbf{P})$ be a complete filtered probability space on which are defined the one-dimensional Wiener processes W_t^l; $l = 1, \ldots, d$, such that $(\{\mathbf{F}_t\}, 0 \leq t \leq T)$ is a natural filtration of $W(t) = (W_t^1, \ldots, W_t^d)$, augmented by all $\mathbf{P}$-negligible sets, we denote by ω the sample points from $\boldsymbol{\Omega}$. In $Q_T \times \boldsymbol{\Omega}$, we investigate the initial boundary value problem for the degenerate quasilinear stochastic parabolic equation

$$d(|y|^{\alpha-2}y) - [\Delta u - c_1|y|^{2\mu-2}y]dt = c_2\sum_{l=1}^{d} |y|^{\sigma_l-1}ydW_t^l, \quad \text{in } Q_T \times \boldsymbol{\Omega}, \tag{1}$$

$$y(x,t,\omega) = 0 \quad \text{on } \partial D \times (0,T) \times \boldsymbol{\Omega}, \tag{2}$$

$$y(x,0,\omega) = y_0(x,\omega) \quad \text{in } D \times \boldsymbol{\Omega}, \tag{3}$$

Received by the editors on April 9, 2002.
AMS subject classification: 60H15, 35D05, 35R60.

where c_1, c_2, α, μ, and σ_l are some nonnegative numbers satisfying the restrictions: $1 < \alpha \leq 2$,

$$\sigma_l \leq \min\{2\mu/\alpha', \alpha - 1\}, \quad 2 \geq 2\mu \geq 1; \tag{4}$$

here and in the sequel we denote by r' the Hölder conjugate of a number $r > 1$, *i.e.*, $r' = r/(r-1)$, Δ denotes the Laplace operator.

We denote

$$g(y) = c_1|y|^{2\mu-2}y, \quad B(y) = c_2(|y|^{\sigma_1-1}y, \dots, |y|^{\sigma_d-1}y).$$

The aim of the present work is to establish an existence result for the problem (1)–(3). We note that the equation (1) is degenerate when $y = 0$. It is a natural stochastic generalization of the famous porous medium equation

$$\frac{\partial(|y|^{\alpha-2}y)}{\partial t} - \Delta y = 0;$$

the study of which has been undertaken by many authors. We refer to [3], [4], [6], [8], just to mention a few. The main technical tool for the investigation of problem (1)–(3) is the splitting-up method, which consists of considering $d(|y|^{\alpha-2}y) + [\Delta y - g(y)]dt - B(y)dW$ as the sum of two operators. For a background on the splitting-up method we refer to the papers [1], [2] and the literature therein.

2. **Preliminaries and formulation of main result.** We consider the well-known function spaces $C_o^\infty(D)$, $L_q(D)$, $W_q^1(D)$, $\overset{o}{W}{}_q^1(D)$ with $q > 1$. By $W_{q'}^{-1}(D)$, we denote the dual of $\overset{o}{W}{}_q^1(D)$ for $q > 1$. Let X be a Banach space. For $r, q \geq 1$, we denote by $\mathcal{L}_{r,q}(0,T,\Omega,X)$ the space of functions $u = u(x,t,\omega)$ with values in X defined on $[0,T] \times \Omega$ and such that

(1) u is measurable with respect to (t,ω) and for each t, u is $\mathcal{F}_t$-measurable.
(2) $u \in X$ for almost all (t,ω) and

$$\|u\|_{\mathcal{L}_{r,q}(0,T,\Omega,X)} = \left(E\left(\int_0^T \|u\|_X^q\,dt\right)^{r/q}\right)^{1/r} < \infty,$$

where E denotes the mathematical expectation.

The space $\mathcal{L}_{r,q}(0,T,\Omega,X)$ so defined is a Banach space. If $r = q$, we shall simply write $\mathcal{L}_r(0,T,\Omega,X)$ for $\mathcal{L}_{r,r}(0,T,\Omega,X)$. When $q = \infty$, the norm in $\mathcal{L}_{r,\infty}(0,T,\Omega,X)$ is given by

$$\|u\|_{\mathcal{L}_{r,\infty}(0,T,\Omega,X)} = (E \sup_{0\leq t\leq T} \|u\|_X^r)^{1/r}.$$

The following compactness results play a central role in the work. In analogy with [4, Chap. 1, Lemma 1.3], we have

LEMMA 1. *Let $(g_\kappa)_{\kappa=1,2\ldots}$ and g be some functions in $\mathcal{L}_q(0,T,\Omega,L_q(D))$ with $q \in (1,\infty)$ such that*

$$\|g_\kappa\|_{\mathcal{L}_q(0,T,\Omega,L_q(D))} \le C, \quad \forall \kappa$$

and as $\kappa \to \infty$

$$g_\kappa \to g \quad \text{for almost all } (x,t,\omega) \in Q_T \times \Omega.$$

Then g_κ weakly converges to g in $\mathcal{L}_q(0,T,\Omega,L_q(D))$.

The next result follows from [7, Section 5, Theorem 5 and Remark 8.2]. It refines an earlier result due to Dubinskii [3].

LEMMA 2. *Let B and B_1 be some Banach spaces such that B is continuously embedded into B_1 and let Y be a subset of B such that $\lambda Y \subset Y$ for all $\lambda \in \mathbf{R}$. Assume that Y is endowed with the semi-norm $Y \ni v \to M(v) \in \mathbf{R}_+$ such that*

$$M(\lambda v) = |\lambda| M(v),$$

and the set $\{v : v \in Y, M(v) \le 1\}$ is relatively compact in B.

Let q, $q_1 \in (1,\infty)$. We define $\mathbf{L}_q(0,T,Y)$ to be the set of functions measurable with respect to t and with values in Y such that

$$\|u\|_{\mathbf{L}_q(0,T,Y)} = \left(\int_0^T M[v(t)]^q \, dt\right)^{1/q} < C_1.$$

Let V be a set bounded in $\mathbf{L}_q(0,T,Y)$ such that

$$\lim_{\theta \to 0} \int_0^{T-\theta} \|v(t+\theta) - v(t)\|_{B_1}^{q_1} \, dt = 0, \quad \text{uniformly for all } v \in V.$$

Then V is relatively compact in $\mathbf{L}_q(0,T,B)$.

We note that $|y|^{\alpha-2}y$ has a continuous modification in $t \in [0,T]$, thus y also has a continuous modification in t. Therefore by identifying y with its continuous modification, we see that the initial condition (3) makes sense.

The main result of the paper is:

THEOREM 3. *Let the conditions (4) be satisfied. Assume that*

$$P\{\omega : y_0 \in \overset{o}{W}{}_2^1(D) \cap L_\alpha(D) \cap L_{2\mu}(D)\} = 1.$$

Then there exists a function

$$y \in \mathcal{L}_2(0,T,\Omega,\overset{o}{W}{}_2^1(D)) \cap \mathcal{L}_{\alpha,\infty}(0,T,\Omega,L_\alpha(D)) \cap \mathcal{L}_{2\mu}(0,T,\Omega,L_{2\mu}(D)),$$

solution of the problem (1)–(3) for almost every $\omega \in \Omega$.

3. Sketch of the proof. For the proof we introduce the splitting-up algorithm. Let $k = T/(N+1)$, where $N = 0, 1, 2, \dots$. We split the interval $[0, T]$ into subintervals $[rk, (r+1)k[$, $r = 0, \dots, N$ and consider the number $\rho \in (0,1)$.

On the probability space $(\mathbf{\Omega}, \mathbf{F}, \{\mathbf{F}_t\}_{0\le t\le T}, \mathbf{P})$ we define the stochastic processes y_k and $\tilde{y}_k$ such that

(5) $$\frac{\partial\big(|y_k(t)|^{\alpha-2}y_k(t)\big)}{\partial t} - \Delta y_k(t) + (1-\rho)g\big(y_k(t)\big) = 0, \quad t \in \big(rk, (r+1)k\big),$$

(6) $$y_k(rk) = y_k^r,$$

(7) $$d\big(|\tilde{y}_k(t)|^{\alpha-2}\tilde{y}_k(t)\big) + \rho g\big(\tilde{y}_k(t)\big)dt = B\big(y_k(t)\big)dW, \quad t \in \big(rk, (r+1)k\big)$$

(8) $$\tilde{y}_k(rk) = y_k\big((r+1)k - 0\big).$$

Next we set

(9) $$y_k^{r+1} = \tilde{y}_k\big((r+1)k - 0\big),$$

and we start with

(10) $$y_k^0 = \tilde{y}_k^0 = y_0;$$

the processes y_k and $\tilde{y}_k$ are subject to the homogenous Dirichlet boundary condition (2) on $\partial D \times (0,T) \times \mathbf{\Omega}$.

The first problem is deterministic and is known to be solvable under the boundary conditions (2) (see *e.g.*, [9]). Equation (7) admits a unique solution under the assumptions (4).

We have:

LEMMA 4. *If*

$$P\{\omega : \tilde{y}_k^0, y_k^0 \in \overset{o}{W}{}_2^1(D) \cap L_\alpha(D) \cap L_{2\mu}(D)\} = 1,$$

then the processes y_k and $\tilde{y}_k$ satisfy the following a priori estimates

(11)
$$E \sup_{0\le t\le T} \|y_k(t)\|_{L_\alpha(D)}^\alpha + E\int_0^T \|y_k(s)\|_{\overset{o}{W}{}_2^1(D)} ds + E\int_0^T \|y_k(s)\|_{L_{2\mu}(D)}^{2\mu} ds \le C_1,$$

(12) $$E \sup_{0\le t\le T} \|\tilde{y}_k(t)\|_{L_\alpha(D)}^\alpha + E\int_0^T \|y_k(s)\|_{L_{2\mu}(D)}^{2\mu} ds \le C_2,$$

(13) $$E\int_0^{T-\theta} \big\| |y_k|^{\alpha-2}y_k(t+\theta) - |y_k|^{\alpha-2}y_k(t)\big\|_{W_2^{-1}(D)}^{p'} dt = O(\theta),$$

(14) $$E\int_0^{T-\theta} \big\| |\tilde{y}_k|^{\alpha-2}\tilde{y}_k(t+\theta) - |\tilde{y}_k|^{\alpha-2}\tilde{y}_k(t)\big\|_{W_2^{-1}(D)}^{p'} dt = O(\theta),$$

(15) $$E\big\| |y_k(t)|^{\alpha-2}y_k(t) - |\tilde{y}_k(t)|^{\alpha-2}\tilde{y}_k(t)\big\|_{W_2^{-1}(D)}^{p'} \le C_3 k,$$

for all $t \in [0,T]$, *where* $O(\theta) \to 0$ *as* $\theta \to 0$, *and* C_i $(i = 1,2,3)$ *are some constants independent of* k.

We give a brief outline of the proof of Theorem 3.

By Lemma 4, we have the following convergences as $k \to 0$,

$$y_{k}, \rightharpoonup y \text{ weakly}^* \text{ in } \mathcal{L}_{\alpha,\infty}(0,T,\Omega,L_\alpha(D)), \tag{16}$$

$$y_k(T) \rightharpoonup \bar{y} \text{ weakly}^* \text{ in } L_\alpha(D) \quad \text{for almost all } \omega, \tag{17}$$

$$y_{k}, \rightharpoonup y \text{ weakly in } \mathcal{L}_{2\mu}(0,T,\Omega,L_{2\mu}(D)), \tag{18}$$

$$\tilde{y}_{k}, \rightharpoonup \tilde{y} \text{ weakly}^* \text{ in } \mathcal{L}_{\alpha,\infty}(0,T,\Omega,L_\alpha(D)), \tag{19}$$

$$\tilde{y}_{k}, \rightharpoonup \tilde{y} \text{ weakly in } \mathcal{L}_{2\mu}(0,T,\Omega,L_{2\mu}(D)). \tag{20}$$

Thus

$$y \in \mathcal{L}_2(0,T,\Omega,\overset{o}{W}{}_2^1(D)) \cap \mathcal{L}_{\alpha,\infty}(0,T,\Omega,L_\alpha(D)) \cap \mathcal{L}_{2\mu}(0,T,\Omega,L_{2\mu}(D)),$$

and

$$\tilde{y} \in \mathcal{L}_{2\mu}(0,T,\Omega,L_{2\mu}(D)) \cap \mathcal{L}_{\alpha,\infty}(0,T,\Omega,L_\alpha(D)).$$

Furthermore

$$|y_k|^{\alpha-2}y_k \to v \text{ weakly}^* \text{ in } \mathcal{L}_{\alpha',\infty}(0,T,\Omega,L_{\alpha'}(D)), \tag{21}$$

$$\Delta y_k \rightharpoonup f \text{ weakly in } \mathcal{L}_2(0,T,\Omega,W_2^{-1}(D)), \tag{22}$$

$$|\tilde{y}_k|^{\alpha-2}\tilde{y}_k \to \tilde{v} \text{ weakly}^* \text{ in } \mathcal{L}_{\alpha',\infty}(0,T,\Omega,L_{\alpha'}(D)), \tag{23}$$

$$|\tilde{y}_k|^{2\mu-2}\tilde{y}_k \rightharpoonup \tilde{g} \text{ weakly in } \mathcal{L}_2(0,T,\Omega,L_2(D)), \tag{24}$$

$$|y_k|^{2\mu-2}y_k \to g \text{ weakly in } \mathcal{L}_2(0,T,\Omega,L_2(D)), \tag{25}$$

$$B(y_k) \to h \text{ weakly}^* \text{ in } \mathcal{L}_{\alpha',\infty}(0,T,\Omega,L_{\alpha'}(D)), \tag{26}$$

$$\int_D B(y_k)dW \to \int_D h\,dW \text{ weakly in } L_{\alpha'}(D), \quad \text{for almost all } \omega \text{ and } t. \tag{27}$$

Using Lemmas 1 and 2, and some compactness results of Prokhorov and Skorokhod as in [1], we show that for almost all $(x,\omega,t) \in Q_T \times \Omega$

$$v = |y|^{\alpha-2}y, \quad \tilde{v} = |\tilde{y}|^{\alpha-2}\tilde{y}, \quad g = |y|^{2\mu-2}y, \quad \tilde{g} = |\tilde{y}|^{2\mu-2}\tilde{y}, \quad h = B(y).$$

By the estimate (15), we have

$$\mathbf{P}\{\omega : y = \tilde{y} \text{ for almost all } x,t\} = 1. \tag{28}$$

Next, arguing as in [5, Theorem 3, Chap. 3, Section 6] and taking in account the relations (16)–(28) we obtain by passage to the limit in appropriate integral expressions that as $k \to 0$, $\bar{y} = y(T)$, and for almost all $(x,t,\omega) \in Q_T \times \Omega$

$$|y(t)|^{\alpha-2}y(t) + \int_0^t f\,ds + \int_0^t |y|^{2\mu-2}y\,ds = \int_0^t B(y)\,dW + |y_0|^{\alpha-2}y_0. \tag{29}$$

Finally, one proves that

$$f(x,t,\omega) = \Delta y, \quad \text{for almost all } (x,t,\omega) \in Q_T \times \Omega. \tag{30}$$

References

1. A. Bensoussan, *Some existence results for stochastic partial differential equations.* In: Stochastic Partial Differential Equations and Applications (Trento, 1990), Pitman Res. Notes Math. Ser. **268**, Longman Sci. Tech., Harlow, UK, 1992, 37–53.
2. A. Bensoussan, R. Glowinski and A. Rascanu, *Approximation of some stochastic differential equations by the splitting-up method.* Appl. Math. Optim. **25**(1992), 81–106.
3. J. Dubinskii, *Weak convergence for non linear elliptic and parabolic equations.* (Russian) Mat. Sbornik **67**(1965), 609–642.
4. J. L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires.* Dunod, Gauthiers-Villars, Paris, 1969 (Russian translation by Mir).
5. E. Pardoux, *Equations aux derivées partielles stochastiques non linéaires monotones.* Thèse, Université Paris XI, 1975.
6. P. A. Raviart, *Sur la résolution de certaines équations paraboliques non linéaires.* J. Funct. Anal. **5**(1970), 299–328.
7. J. Simon, *Compact sets in the space $L^p(0,T;B)$.* Ann. Mat. Pura Appl. (4) **146**(1987), 65–96.
8. M. Tsutsumi, *On solutions of some doubly nonlinear degenerate parabolic equations with absorption.* J. Math. Anal. Appl. **132**(1988), 187–212.

Department of Mathematics
Vista University
Private Bag X1311
Silverton 0127
Pretoria
South Africa
email: sango-m@marlin.vista.ac.za

C. R. Math. Rep. Acad. Sci. Canada Vol. **25**, (3), 2003 pp. 71–75

A NOTE ON THE DIVISIBILITY OF CLASS NUMBERS OF REAL QUADRATIC FIELDS

FLORIAN LUCA

Presented by M. Ram Murty, FRSC

RÉSUMÉ. Dans cet article, on prouve que si $g \geq 2$ est un entier positif pair fixé, alors pour les valeurs positives grandes de x, il y a un nombre de $\gg x^{1/g}/\log x$ des entiers positifs et libres des careés $d < x$, tel que le nombre de classes du corps $\mathbf{Q}[\sqrt{d}]$ est divisible par g.

Let $d > 1$ be a squarefree positive integer. We put $\mathbf{K} := \mathbf{Q}[\sqrt{d}]$, and we write $\mathcal{O}_{\mathbf{K}}$, $N_{\mathbf{K}}$, and $h_{\mathbf{K}}$, for the ring of algebraic integers in $\mathbf{K}$, the norm map over $\mathbf{K}$, and the class number of $\mathbf{K}$, respectively. We fix a positive integer d, and we shall construct "many" such numbers d, such that $h_{\mathbf{K}}$ is a multiple of g. We let x be a large real number, we write $\log x$ for its natural logarithm, and $N_g(x)$ for the number of squarefree numbers $d < x$ so that $g | h_{\mathbf{K}}$. We use the Vinogradov symbols $\gg$ and $\ll$ with their usual meanings.

In [3], M. R. Murty showed that if g is odd, then for every $\varepsilon > 0$, the inequality $N_g(x) \gg x^{1/(2g)-\varepsilon}$ holds for large values of x. This result was recently improved by G. Yu [5], who showed that $N_g(x) \gg x^{1/g-\varepsilon}$ holds for large values of x. In both these estimates, the constant understood in $\gg$ may depend on both ε and g.

Murty's argument from [3] also gives $N_g(x) \gg x^{1/(2g)-\varepsilon}$ when $g \equiv 2 \pmod 4$, but when $4 | g$ his argument only gives $N_g(x) \gg x^{1/(4g)-\varepsilon}$.

Results of the same type for the case of the complex quadratic number fields appear in [3] and [4].

In this note, we obtain the same lower bound on $N_g(x)$ as Yu's for the case in which g is even.

For any integer $g > 1$, write $G := \mathrm{lcm}[g, 2]$. Clearly, $G = g$ when g is even. Our result is the following:

THEOREM. *The inequality $N_G(x) \gg x^{1/G}/\log x$ holds for sufficiently large values of x, where the constant understood in $\gg$ above depends on g.*

PROOF. We assume that $g > 1$, that x is large, we set $y := x^{1/G}$, and we write $\mathcal{P}$ for the set of all odd primes $p < y$ so that the polynomial $X^G - 2 \in \mathbf{Z}[X]$ is irreducible. A standard application of Chebotarev's Density Theorem together with the Prime Number Theorem gives $|\mathcal{P}| \gg \pi(y) \gg y/\log y \gg x^{1/G}/\log x$. For $p \in \mathcal{P}$ we write

Received by the editors on January 13, 2003.
AMS subject classification: 11E41.
Keywords: Pell equations, class numbers of quadratic fields.

$$p^G + 1 = dz^2, \tag{1}$$

with some positive integers z and d, and with d squarefree. Clearly, $d \leq p^G + 1 < x + 1$. Since p is odd and G is even, we have $2||d$. Let $\mathcal{P}_1$ be the subset of those $p \in \mathcal{P}$ with $d = 2$. If $p \in \mathcal{P}_1$, then $(X, Y) := (p^{G/2}, z)$ is a solution of the Pell equation $X^2 - 2Y^2 = -1$. From the theory of Pell equations, it follows that there exists an odd integer $m > 1$ so that the formula

$$p^{G/2} = \frac{\alpha^m + \beta^m}{2} \tag{2}$$

holds, where $(\alpha, \beta) := (1 + \sqrt{2}, 1 - \sqrt{2})$. Since $x^{1/2} > p^{G/2}$, we get that $x^{1/2} \gg (1 + \sqrt{2})^m$, therefore $m \ll \log x$. Thus, $|\mathcal{P}_1| \ll \log x$; hence, $|\mathcal{P} \setminus \mathcal{P}_1| \gg x^{1/G}/\log x$. We remark in passing that $\mathcal{P}_1$ is empty when $G > 2$ by Theorem 6.1 in [1].

We now show that the d's appearing from equation (1) for $p \in \mathcal{P} \setminus \mathcal{P}_1$ are mutually distinct, and that $h_{\mathbf{K}}$ is a multiple of G for every one of these d's, which will finish the proof of the Theorem.

THE FUNDAMENTAL UNIT. The pair of integers $(X, Y) := (p^{G/2}, z)$ is a solution of the Pell equation $X^2 - dY^2 = -1$. For a positive integer m let (X_m, Y_m) stand for the m-th solution in positive integers to the Pell equation $X^2 - dY^2 = \pm 1$. By the well known properties of solutions to Pell equations, it follows that $p^{G/2} = X_m$ must hold with some odd integer m. We prove that $m = 1$. If $m > 1$, then $X_1 | X_m$ (because m is odd), and $X_1 > 1$ (because $d \neq 2$). Thus, X_1 must be a power of p, which shows that every prime divisor of X_m is a prime divisor of X_1 as well. This is impossible by the argument from p. 162 in [2]. Thus, $m = 1$, and therefore $\zeta := p^{G/2} + z\sqrt{d}$ is the fundamental unit in $\mathcal{O}_{\mathbf{K}}$, and is of norm -1.

AN ELEMENT OF THE CLASS GROUP WHOSE ORDER DIVIDES G. Since $2||d$, it follows that $\{1, \sqrt{d}\}$ is a base for $\mathcal{O}_{\mathbf{K}}$. The discriminant of $\mathbf{K}$ is $4d$, p is coprime to z (by equation (1)), and $(4d|p) = (dz^2|p) = (p^G + 1|p) = 1$, where for an integer a we used the notation $(a|p)$ for the Legendre symbol of a with respect to p. Thus, p splits into two prime ideals in $\mathcal{O}_{\mathbf{K}}$, lets call them π_1 and π_2. Rewriting (1) as $p^G = 1 - dz^2 = (1 - \sqrt{d}z)(1 + \sqrt{d}z)$, and passing to ideals, we get $\pi_1^G \cdot \pi_2^G = [1 + \sqrt{d}z][1 - i\sqrt{d}z]$. The two ideals $[1 + \sqrt{d}z]$ and $[1 - \sqrt{d}z]$ are easily seen to be coprime (for the sum of their generators is 2, and the product of their generators is a power of p), and by unique factorization for ideals in $\mathcal{O}_{\mathbf{K}}$, it follows that we may assume that the equation

$$\pi_1^G = [1 + \sqrt{d}z] \tag{3}$$

holds. Equation (1) immediately implies that the order of π_1 in the ideal class group divides G. It remains to show that this order is exactly G. Assuming that it is not, we deduce the existence of a prime number $r|G$, and of an element $\alpha = u + \sqrt{d}v \in \mathcal{O}_{\mathbf{K}}$ so that the equation

$$[1 + \sqrt{d}z] = [\alpha^r] \tag{4}$$

holds. Here, u and v are nonzero integers with $\gcd(u, dv) = 1$. Equation (4) leads to the diophantine equation

$$1 + \sqrt{d}z = \varepsilon\alpha^r\zeta^s, \tag{5}$$

where $\varepsilon \in \{\pm 1\}$, and s is an integer. We distinguish two cases.

CASE 1. $r = 2$.

We may assume that $s \in \{0, -1\}$. Taking norms in (5), we get

$$-p^G = 1 - dz^2 = N_{\mathbf{K}}(1 + \sqrt{d}z) = N_{\mathbf{K}}(\varepsilon)N_{\mathbf{K}}(\alpha^2)N_{\mathbf{K}}(\zeta^s) = \left(N_{\mathbf{K}}(\alpha)\right)^2 \cdot (-1)^s,$$

therefore s is odd. Thus, $s = -1$, and equation (5) becomes $1 + \sqrt{d}z = \varepsilon\alpha^2\zeta^{-1}$, and since $z > 0$, $\zeta < 0$, it follows that $\varepsilon = -1$. Thus, equation (5) can be rewritten as

$$(1 + \sqrt{d}z)\zeta = \alpha^2.$$

Rewriting the above equation as

$$(1 + \sqrt{p^G + 1})(p^{G/2} + \sqrt{p^G + 1}) = \left(u + \frac{v}{z}\sqrt{p^G + 1}\right)^2, \tag{6}$$

and identifying the rational and irrational parts from both sides of equation (6) above, we get

$$u^2 + (p^G + 1)\left(\frac{v}{z}\right)^2 = p^G + p^{G/2} + 1 \quad \text{and} \quad 2u\left(\frac{v}{z}\right) = p^{G/2} + 1. \tag{7}$$

Expressing v/z from the second equation (7) in terms of p and u, and inserting the result into the first equation (7), we get the equation

$$u^4 - (p^G + p^{G/2} + 1)u^2 + \frac{(p^G + 1)(p^{G/2} + 1)^2}{4} = 0,$$

therefore

$$u^2 = \frac{(p^{G/2} + 1)^2}{2} \quad \text{or} \quad u^2 = \frac{p^G + 1}{2}. \tag{8}$$

The first equation (8) is impossible because 2 is not a square, while the second equation (8) is impossible because $d > 2$.

CASE 2. $r > 2$.

We first remark that we may choose α in equation (5) so that $s < 0$ and odd. Indeed, with a fixed equation of the form (5), replace α by $\alpha\zeta^t$, where t is a large positive integer incongruent to s modulo 2. After this replacement, the exponent s in (5) will be replaced by $s - rt$, which is negative for large t and odd, and ε will be replaced by $\varepsilon \cdot (-1)^{rt}$. By replacing now α by $-\varepsilon\alpha$, it follows that we may assume that $\varepsilon = -1$. Thus, we have shown the existence of an equation of the form

$$1 + \sqrt{d}z = -\alpha^r \zeta^{-s}, \tag{9}$$

where $s > 0$ is odd. Computing norms in (9), we get

$$-p^G = 1 - dz^2 = N_{\mathbf{K}}(1 + \sqrt{d}z) = N_{\mathbf{K}}(-1)N_{\mathbf{K}}(\alpha^r)N_{\mathbf{K}}(\zeta^{-s}) = -N_{\mathbf{K}}(\alpha)^r.$$

Thus, $N_{\mathbf{K}}(\alpha) = p^{G/r}$, and therefore we have obtained the equation

$$u^2 - dv^2 = p^{G/r}. \tag{10}$$

We now rewrite (9) as $(1 + \sqrt{d}z)\zeta^s = \alpha^r$, which, with (10), is seen to be of the form

$$(1 + \sqrt{p^G + 1})(p^{G/2} + \sqrt{p^G + 1})^s = (u + \varepsilon_1\sqrt{u^2 - p^{G/r}})^r, \tag{11}$$

where $\varepsilon_1 = \mathrm{sign}(v) \in \{\pm 1\}$. Identifying the rational parts from (11), we get the equation

$$\begin{aligned}(1 + \sqrt{p^G + 1})(p^{G/2} + \sqrt{p^G + 1})^s + (1 - \sqrt{p^G + 1})(p^{G/2} - \sqrt{p^G + 1})^s \\ = (u + \sqrt{u^2 - p^{G/r}})^r + (u - \sqrt{u^2 - p^{G/r}})^r.\end{aligned} \tag{12}$$

The left hand side of (12) modulo p is 2, while the right hand side of (12) modulo p is $(2u)^r$. Complicated proofs of these statements can be achieved by expanding both sides of (12) using the binomial formula. Simple proofs of these statements can be achieved by setting say $X := p$ in the left hand side of (12), treating this left hand side of (12) as a formal polynomial in X, and then computing its value in zero, which is obviously 2, and a similar argument works for the left hand side of (12). Thus, we have arrived at the congruence $(2u)^r \equiv 2 (\mathrm{mod}\, p)$. Setting $x_0 := 2u$, it follows that $X^{G/r} - x_0$ is a factor of $X^G - 2$ modulo p, which is impossible from the way we have defined our set of prime numbers $\mathcal{P}$.

This concludes the proof of the Theorem.

ACKNOWLEDGEMENTS. I thank Professor M. R. Murty for useful suggestions, and for making me aware of Yu's work [5]. This work was supported by Project PAPIIT-IN104602 from the UNAM.

References

1. J. H. E. Cohn, *The diophantine equation $x^n = Dy^2 + 1$*. Acta Arith. (1) **106**(2003), 73–83.
2. F. Luca and P. G. Walsh, *The product of like-indexed terms in binary recurrences*. J. Number Theory **96**(2002), 152–173.
3. M. R. Murty, *Exponents of class groups of quadratic fields*. In: "Topics in Number Theory", Math. Appl. **467**, Kluwer Academic Publishers, Dordrecht, 1999, 229–239.
4. K. Soundararajan, *Divisibility of class numbers of imaginary quadratic fields*. J. London Math. Soc. (2) **61**(2000), 681–690.
5. G. Yu, *A note on the divisibility of class numbers of real quadratic fields*. J. Number Theory **97**(2002), 35–44.

Mathematical Institute, UNAM
Ap. Postal 61-3 (Xangari)
CP 58 089
Morelia, Michoacán
México
email: fluca@matmor.unam.mx

C. R. Math. Rep. Acad. Sci. Canada Vol. 25, (3), 2003 pp. 76–81

ON SOME QUESTIONS OF EYMARD AND BEKKA CONCERNING AMENABILITY OF HOMOGENEOUS SPACES AND INDUCED REPRESENTATIONS

VLADIMIR PESTOV

Presented by G. A. Elliott, FRSC

ABSTRACT. Let $F \subseteq H \subseteq G$ be closed subgroups of a locally compact group. In response to a 1972 question by Eymard, we construct an example where the homogeneous factor space G/F is amenable in the sense of Eymard–Greenleaf, while H/F is not. (In our example, G is discrete.) As a corollary which answers a 1990 question by Bekka, the induced representation $\mathrm{ind}_H^G(\rho)$ can be amenable in the sense of Bekka even if ρ is not amenable. The second example, answering another question by Bekka, shows that $\mathrm{ind}_H^G(\rho)$ need not be amenable even if both the representation ρ and the coset space G/H are amenable.

RÉSUMÉ. Soient $F \subseteq H \subseteq G$ deux sous-groupes fermés d'un groupe localement compact G. En réponse à une question posée en 1972 par Eymard, nous construisons un groupe discret G tel que l'espace homogène G/F est moyennable au sens d'Eymard et Greenleaf, bien que H/F n'est pas moyennable. On obtienne un corollaire qui répond à un problème posé par Bekka en 1990: une représentation induite $\mathrm{ind}_H^G(\rho)$ peut être moyennable au sens de Bekka même si ρ n'est pas moyennable. Le deuxième exemple, qui répond à une autre question de Bekka, montre que $\mathrm{ind}_H^G(\rho)$ n'est pas nécessairement moyennable même si la représentation ρ et l'espace homogène G/H sont l'un et l'autre moyennables.

1. Introduction. Let H be a closed subgroup of a locally compact group G. The homogeneous factor space G/H is *amenable in the sense of Eymard and Greenleaf* [5], [6], if $L^\infty(G/H)$ supports a G-invariant mean. If $H = \{e\}$, one obtains the classical concept of an amenable locally compact group.

A unitary representation ρ of a group G in a Hilbert space $\mathcal{H}$ is *amenable in the sense of Bekka* [1] if there exists a state, ϕ, on the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators that is $\mathrm{Ad}\, G$-invariant: $\phi\big(\pi(g)T\pi(g)^{-1}\big) = \phi(T)$ for every $T \in \mathcal{B}(\mathcal{H})$ and every $g \in G$. For instance, the homogeneous space G/H is Eymard–Greenleaf amenable if and only if the quasi-regular representation $\lambda_{G/H}$ of G in $L^2(G/H)$ is amenable.

Let F and H be closed subgroups of a locally compact group G, such that $F \subseteq H \subseteq G$. In 1972 Eymard had asked [5, p. 55]:

Q1. *Suppose the space G/F is amenable. Is then H/F amenable?*

Received by the editors on January 23, 2003.

Partially supported by the Marsden Fund of the Royal Society of New Zealand and by a University of Ottawa start-up grant.

AMS subject classification: 43A07, 43A65, 43A85, 22D30.

This is of course a classical result in the case $F = \{e\}$.

Let π be a strongly continuous unitary representation of H. In 1990 Bekka has shown [1] that, if the unitarily induced representation $\operatorname{ind}_H^G(\pi)$ is amenable, then G/H is amenable. He asked the following:

Q2. *Does amenability of* $\operatorname{ind}_H^G(\pi)$ *imply that of* π*?*

(This is a more general version of Eymard's question Q1.)

Q3. *Suppose both* G/H *and* π *are amenable. Is then* $\operatorname{ind}_H^G(\pi)$ *amenable?*

Question 3 was partially answered in the positive by Bekka himself [1] under extra assumptions on H. Further discussion can be found in [7].

We show that in general the answer to all three questions above is negative.

REMARK. In a recent independent work Nicolas Monod and Sorin Popa have also solved Eymard's problem (and thus Bekka's Q2 as well). (See their note [8] in the same issue.)

2. Reminders and simple facts.

2.1. A unitary representation π of a locally compact group G in a Hilbert space $\mathcal{H}$ is said to *almost have invariant vectors* if for every compact $K \subseteq G$ and every $\varepsilon > 0$ there is a $\xi \in \mathcal{H}$ satisfying $\|\xi\| = 1$ and $\|\pi_g(\xi) - \xi\| < \varepsilon$ for all $g \in K$.

2.2. *If a unitary representation* π *of a group* G *(viewed as discrete) almost has invariant vectors, then* π *is amenable.*

This follows from Corollary 5.3 of [1], because π weakly contains the trivial one-dimensional representation, $1_G < \pi$. Alternatively, choose for every finite $K \subseteq G$ and every $\varepsilon > 0$ an almost invariant vector $\xi_{K,\varepsilon}$ as above, and set $\phi_{K,\varepsilon}(T) := \langle T\xi_{K,\varepsilon}, \xi_{K,\varepsilon}\rangle$. Every weak* cluster point, ϕ, of the net of states $(\phi_{K,\varepsilon})$ is a G-invariant state.

Note that the converse of 2.2 is not true, as for example every unitary representation of an amenable group is amenable by Theorem 2.2 of [1].

2.3. *Let* H *be a closed subgroup of a locally compact group* G*. The following statements are equivalent.*

- (i) *The homogeneous space* G/H *is amenable.*
- (ii) *The left quasi-regular representation* $\lambda_{G/H}$ *of* G *in* $L^2(G/H)$*, formed with respect to any (equivalently: every) quasi-invariant measure* ν *on* G/H*, almost has invariant vectors.*
- (iii) *The left quasi-regular representation* $\lambda_{G/H}$ *is amenable.*
- (iv) *If* G *acts continuously by affine transformations on a convex compact set* C *in such a way that* C *contains an* H*-fixed point, then* C *contains a* G*-fixed point.*

The equivalences between (i), (ii), and (iv) are due to Eymard [5], while (iii) was added by Bekka [1, Theorem 2.3 (i)]. The condition (ii) is an analogue of Reiter's condition (P_2) for homogeneous spaces. (Interestingly, a natural analogue of Følner's condition for homogeneous spaces fails [6].)

2.4. *Let the homogeneous space G/H be amenable, and let π be a strongly continuous unitary representation of G. If the restriction of π to H is an amenable representation, then π is amenable as well.*

To prove this statement, denote, following Bekka [1, Section 3], by $X(\mathcal{H})$ the C^*-subalgebra of $\mathcal{B}(\mathcal{H})$ formed by all operators T with the property that the orbit map of the adjoint action,

$$G \ni g \mapsto \pi(g)T\pi(g)^{-1} \in \mathcal{B}(\mathcal{H}),$$

is norm-continuous. Then G acts upon $X(\mathcal{H})$ in a strongly continuous way by isometries. It follows that the dual adjoint action of G on the state space of $X(\mathcal{H})$, equipped with the weak* topology, is continuous. Because of the assumed amenability of $\pi|_H$, there is an H-invariant state, ψ, on $\mathcal{B}(\mathcal{H})$ and consequently on $X(\mathcal{H})$. Eymard's conditional fixed point property 2.3 (iv) allows us to conclude that there is a G-invariant state, ϕ, on $X(\mathcal{H})$. This in turn implies amenability of π by Theorem 3.5 of [1].

2.5. Let H be a closed subgroup of a locally compact group G, and let π be a strongly continuous unitary representation of H in a Hilbert space. By $\mathrm{ind}_H^G(\pi)$ we will denote, as usual, the unitarily induced representation of G.

2.6. The following is Mackey's generalization of the Frobenius Reciprocity Theorem [9, Theorem 5.3.3.5].

Let H be a closed subgroup of a locally compact group G, and let π and ρ be finite-dimensional irreducible unitary representations of H and G, respectively. If G/H carries a finite invariant measure, then $\mathrm{ind}_H^G(\pi)$ contains ρ as a discrete direct summand exactly as many times as $\rho|_H$ contains π as a discrete direct summand.

2.7. Let G be a locally compact group, and let H, F be closed subgroups of G such that $F \subseteq H$. Set $\pi = \mathrm{ind}_F^H(1_F)$, where 1_F stands for the trivial one-dimensional representation of F. Then $\pi = \lambda_{H/F}$ is unitarily equivalent to the quasi-regular representation of H in $L^2(H/F)$ [9, Corollary 5.1.3.6]. By the theorem on induction in stages (*ibid.*, Proposition 5.1.3.5), the induced representation $\mathrm{ind}_H^G(\lambda_{H/F})$ is unitarily equivalent to $\mathrm{ind}_F^G(1_F)$, which is just the representation $\lambda_{G/F}$, the quasi-regular representation of G in $L^2(G/F)$.

Now assume that the homogeneous factor space G/F is amenable. By Bekka's result mentioned above (2.3.(iii)), this amounts to the amenability of $\lambda_{G/F}$.

This argument shows that Eymard's Q1 is a particular case of Bekka's Q2.

3. First example.

3.1. Let $H = F_\infty$, where F_∞ denotes the free non-abelian group on infinitely many free generators x_i, $i \in \mathbb{Z}$. For every $n \in \mathbb{Z}$, let Γ_n denote the normal subgroup of F_∞ generated by x_i, $i \leq n$. We will set $F = \Gamma_0$.

As the factor space $H/F = F_\infty/\Gamma_0$ is a free group, it is not amenable.

3.2. Let the group $\mathbb{Z}$ act on F_∞ by group automorphisms via shifting the generators:

$$\tau_n(x_m) := x_{m+n},$$

where $m, n \in \mathbb{Z}$ and τ denotes the action of $\mathbb{Z}$ on F_∞.

Denote by $G = \mathbb{Z} \ltimes_\tau F_\infty$ the semi-direct product formed with respect to the action τ. That is, G is the Cartesian product $\mathbb{Z} \times F_\infty$ equipped with the group operation $(m,x)(n,y) = (m+n, x\tau_m y)$, the neutral element $(0,e)$ and the inverse $(n,x)^{-1} = (-n, \tau_{-n}x^{-1})$.

Let us show that the homogeneous space G/F is Eymard–Greenleaf amenable.

3.3. For every $n \in \mathbb{Z}$,

$$\begin{aligned}(n,e)F &= \{(n,e)(0,y) : y \in F\}\\ &= \{(n,\tau_n y) : y \in \Gamma_0\}\\ &= \{n\} \times \Gamma_n.\end{aligned}$$

Let $S \subset F_\infty$ be an arbitrary finite subset. For $n \in \mathbb{Z}$ sufficiently large, $S \subset \Gamma_n$. Now for each $(0,s) \in S$,

$$\begin{aligned}(0,s)(n,e)F &= \{(0,s)(n,z) : z \in \Gamma_n\}\\ &= \{n\} \times (s\Gamma_n)\\ &= \{n\} \times \Gamma_n\\ &= (n,e)F;\end{aligned}$$

that is, the left F-coset $(n,e)F \in G/F$ is S-invariant.

Consequently, the unit vector $\delta_{(n,e)F} \in \ell^2(G/F)$ is S-invariant. We have proved that the H-module $\ell^2(G/F)$ almost has invariant vectors, and therefore the restriction to H of the left regular representation $\lambda_{G/F}$ is amenable. Since G/H is an amenable homogeneous space (H is normal in G and the factor-group G/H is isomorphic to $\mathbb{Z}$), we conclude by 2.4 that the left regular representation of G in $\ell^2(G/F)$ is amenable, that is, that G/F is an amenable homogeneous space.

3.4. The constructed triple of groups $F \subset H \subset G$ provides a negative answer to the question of Eymard (Q1).

In view of the remarks in 2.7, it provides a negative answer to Bekka's question (Q2) as well. Take as $\pi = \mathrm{ind}_F^H(1_F) = \lambda_{H/F}$ the left quasi-regular representation of H in $\ell^2(H/F)$. While π is not amenable, the induced representation $\mathrm{ind}_H^G(\pi) = \lambda_{G/F}$ is amenable.

3.5. The following way of viewing our example may be instructive.

If considered as a unitary F_∞-module, $\ell^2(G/F)$ decomposes, up to unitary equivalence, into the orthogonal sum

$$\ell^2(G/F) \cong \bigoplus_{n\in\mathbb{Z}} \ell^2(F_\infty/\Gamma_n).$$

None of the unitary F_∞-modules $\ell^2(F_\infty/\Gamma_n)$ is amenable, yet their orthogonal sum is amenable, because the family $\left(\ell^2(F_\infty/\Gamma_n)\right)_{n\in\mathbb{Z}}$ "asymptotically has invariant vectors": every finite $S \subset F_\infty$ acts trivially on the Hilbert space $\ell^2(F_\infty/\Gamma_n)$, provided n is large enough so that $S \subset \Gamma_n$. (Indeed, Γ_n are normal in F_∞.)

4. Second example.

4.1. Let $G = \mathrm{SL}(n, \mathbb{R})$ and $H = \mathrm{SL}(n, \mathbb{Z})$, $n \geq 3$. Since H is a (non-uniform) lattice in G (*cf.* [3, Exercise 7, §2, Chapitre VII]), the homogeneous space G/H is amenable.

4.2. The group $\mathrm{SL}(n, \mathbb{Z})$ is maximally almost periodic (for instance, homomorphisms to the groups $\mathrm{SL}(n, \mathbb{Z}_p)$, where p is a prime number, separate points in $\mathrm{SL}(n, \mathbb{Z})$). Let π be a non-trivial (of dimension > 1) irreducible unitary finite-dimensional representation of H. Being finite-dimensional, π is an amenable representation; *cf.* [1, Theorem 1.3 (i)].

4.3. Applying Mackey's Reciprocity Theorem (quoted above, 2.6) with ρ equal to 1_G, a trivial one-dimensional representation of G, we conclude that the unitarily induced representation $\mathrm{ind}_H^G(\pi)$ has no non-zero invariant vectors.

At the same time, the Lie group $G = \mathrm{SL}(n, \mathbb{R})$, $n \geq 3$, has property (T) (see *e.g.* [4, Chapitre 2.a, Théorème 4]), and the only amenable representations of non-compact simple Lie groups with Kazhdan's property (T), such as G, are those having a non-zero invariant vector ([1, Remark 5.10]; *cf.* also the principal result of [2]).

4.4. We conclude: $\mathrm{ind}_H^G(\pi)$ is non-amenable, even if the unitary representation π and the homogeneous space G/H are both amenable.

This answers in the negative another question posed by Bekka (Q3).

References

1. M. E. B. Bekka, *Amenable unitary representations of locally compact groups*. Invent. Math. **100**(1990), 383–401.
2. M. E. B. Bekka and A. Valette, *Kazhdan's property* (T) *and amenable representations*. Math. Z. **212**(1993), 293–299.
3. N. Bourbaki, *Intégration*. Chapitres 7 et 8, Hermann, Paris, 1963.

4. P. de la Harpe and A. Valette, *La propriété* (T) *de Kazhdan pour les groupes localements compacts.* Astérisque **175**, 1989.
5. P. Eymard, *Moyennes invariantes et représentations unitaires.* Lecture Notes in Math. **300**, Springer-Verlag, Berlin-New York, 1972.
6. F. P. Greenleaf, *Amenable actions of locally compact groups.* J. Funct. Anal. **4**(1969), 295–315.
7. P. Jolissaint, *Invariant states and a conditional fixed point property for affine actions.* Math. Ann. **304**(1996), 561–579.
8. N. Monod and S. Popa, *On co-amenability for groups and von Neumann algebras.* This issue.
9. G. Warner, *Harmonic Analysis on Semi-Simple Lie Groups I.* Grundlehren Math. Wiss. **188**, Springer-Verlag, Berlin-Heidelberg-New York, 1972.

Department of Mathematics and Statistics
University of Ottawa
Ottawa, Ontario
K1N 6N5
email: vpest283@science.uottawa.ca

C. R. Math. Rep. Acad. Sci. Canada Vol. 25, (3), 2003 pp. 82–87

ON CO-AMENABILITY FOR GROUPS AND VON NEUMANN ALGEBRAS

NICOLAS MONOD AND SORIN POPA

Presented by G. A. Elliott, FRSC

RÉSUMÉ. Nous démontrons tout d'abord que la comoyennabilité ne passe pas aux sous-groupes, résolvant ainsi une question posée par P. Eymard en 1972. Nous étudions ensuite la comoyennabilité pour les algèbres de von Neumann; en particulier, nous clarifions le lien avec le cas des groupes.

1. Co-amenable subgroups. A subgroup H of a group G is called *co-amenable* in G if it has the following relative fixed point property:

Every continuous affine G-action on a convex compact subset of a locally convex space with an H-fixed point has a G-fixed point.

This is equivalent to the existence of a G-invariant mean on the space $\ell^\infty(G/H)$ or to the weak containment of the trivial representation in $\ell^2(G/H)$; see [4] (compare also Proposition 3 below). Alternative terminology is *amenable pair* or *co-Følner*, and more generally Greenleaf [5] introduced the notion of an *amenable action* (conflicting with more recent terminology) further generalized by Zimmer [10] to amenable pairs of actions.

In the particular case of a normal subgroup $H \triangleleft G$, co-amenability is equivalent to amenability of the quotient G/H.

One checks that for a triple

$$K < H < G,$$

co-amenability of K in H and of H in G implies that K is co-amenable in G. In 1972, Eymard proposed the following problem [4, N° 3, §8]:

If we assume K co-amenable in G, is it also co-amenable in H?

(It is plain that H is co-amenable in G.)

If K is normal in G, then the answer is positive, since it amounts to the fact that amenability passes to subgroups. We give an elementary family of counter-examples for the general case. Interestingly, we can even get some normality:

THEOREM 1.

(i) *There are triples of groups $K \triangleleft H \triangleleft G$ with K co-amenable in G but not in H.*

Received by the editors on February 5, 2003.
AMS subject classification: 43A07, 43A85, 46L.

(ii) *For any group Q there is a triple $K \lhd H_0 < G$ with K co-amenable in G and $H_0/K \cong Q$.*

Moreover, all groups in (i) *can be assumed ICC (infinite conjugacy classes). In* (ii), *the group G is finitely generated/presented, ICC, torsion-free, etc. as soon as Q has the corresponding property.*

REMARK 1. We shall see that one can also find counter-examples $K < H_0 < G$ with all three groups finitely generated.

REMARK 2. We mention that in a recent independent work V. Pestov also solved Eymard's problem (see V. Pestov's note in this same issue).

We start with the following general setting: Let K be any group and $\vartheta\colon K \to K$ a monomorphism. Denote by $G = K*_\vartheta$ the corresponding HNN-extension, *i.e.*,

$$G \stackrel{\text{def}}{=} \langle K, t \mid tkt^{-1} = \vartheta(k)\ \forall k \in K \rangle.$$

Write $H \stackrel{\text{def}}{=} \bigcup_{k\in\mathbf{Z}} t^{-k}Kt^k$ so that $G = H \rtimes \langle t \rangle$.

PROPOSITION 2. *The group K is co-amenable in G.*

PROOF. Note that H is co-amenable in G, so that by applying the relative fixed point property to the space of means on $\ell^\infty(G/K)$ we deduce that it is enough to find an H-invariant mean

$$\mu\colon \ell^\infty(G/K) \longrightarrow \mathbf{R}.$$

For every $k \in \mathbf{N}$ we define a mean μ_k by $\mu_k(f) = f(t^{-k}K)$ for f in $\ell^\infty(G/K)$. This mean is left invariant by the group $t^{-k}Kt^k$. Since H is the increasing union of $t^{-k}Kt^k$ for $k \to \infty$, any weak-$*$ limit of the sequence $\{\mu_k\}$ is an H-invariant mean. ■

We set $H_0 = t^{-1}Kt$ and consider $K < H_0 < H \lhd G$. Observe that K is not co-amenable in H_0 unless $\vartheta(K)$ is co-amenable in K, and hence also not in H when $K \lhd H$. Thus we will construct our counter-examples by finding appropriate pairs (K, ϑ) with $\vartheta(K)$ not co-amenable in K.

For Theorem 1, let Q be any group, let $K = \bigoplus_{n\geq 0} Q$ be the direct limit of finite Cartesian powers and let ϑ be the shift; then $\vartheta(K)$ is co-amenable in K if and only if Q is amenable. The claims follow; observe that G is the restricted wreath product of Q by $\mathbf{Z}$.

For Remark 1, we may for instance take K to be the free group on two generators and define ϑ by choosing two independent elements in the kernel of a surjection of K to a non-amenable group.

Amenability of groups can be characterized in terms of *bounded cohomology*; see [6]. Here is a generalization, which shows that actually the transitivity and (lack of) hereditary properties of co-amenability were to be expected.

PROPOSITION 3. *The following are equivalent:*

(i) *H is co-amenable in G.*
(ii) *For any dual Banach G-module V and any $n \geq 0$, the restriction $\mathrm{H}^n_b(G, V) \to \mathrm{H}^n_b(H, V)$ is injective.*

PROOF. "(i) ⇒ (ii)" is obtained by constructing a *transfer* map (at the cochain level) using an invariant mean on G/H, along the lines of [7, §8.6].

"(ii) ⇒ (i)" Let V be the G-module $(\ell^\infty(G/H)/\mathbf{C})^*$, where $\mathbf{C}$ is embedded as constants. The extension $0 \to V \to \ell^\infty(G/H)^* \to \mathbf{C} \to 0$ splits over H (by evaluation on the trivial coset) so that in the associated long exact sequence the transgression map $\mathbf{C} \to \mathrm{H}^1_b(H, V)$ vanishes. By (ii) and naturality, the transgression $\mathbf{C} \to \mathrm{H}^1_b(G, V)$ also vanishes, and so the long exact sequence for G yields a G-invariant element μ in $\ell^\infty(G/H)^*$ which does not vanish on the constants. Taking the absolute value of μ (for the canonical Riesz space structure) and normalizing it gives an invariant mean. ■

Co-amenable von Neumann subalgebras. Let N be a finite von Neumann algebra and $B \subset N$ a von Neumann subalgebra. For simplicity, we shall assume N has countably decomposable centre. Equivalently, we assume N has a normal faithful tracial state τ. Let $B \subset N \overset{e_B}{\subset} \langle N, B \rangle$ be the basic construction for $B \subset N, \tau$, *i.e.*, $\langle N, B \rangle = JBJ' \cap \mathcal{B}(L^2(N, \tau))$ with J the canonical conjugation on $L^2(N, \tau)$ and e_B the canonical projection of $L^2(N, \tau)$ onto $L^2(B, \tau)$. Note that, up to isomorphism, the inclusion $N \subset \langle N, B \rangle$ does not depend on τ. We also denote by Tr the unique normal semifinite faithful trace on $\langle N, B \rangle$ given by $\mathrm{Tr}(xe_By) = \tau(xy)$, $x, y \in N$.

DEFINITION 4. ([8], see also [9]) The subalgebra B is *co-amenable* in N if there exists a norm one projection Ψ of $\langle N, B \rangle$ onto N. One also says that N is *amenable relative to B.*

REMARK 3. If $B_0 \subset B \subset N$ and N is amenable relative to B_0 then N is amenable relative to B. Indeed, since $\langle N, B \rangle$ is included in $\langle N, B_0 \rangle$, we may restrict $\Psi \colon \langle N, B_0 \rangle \to N$ to $\langle N, B \rangle$.

The next result from [8], [9] provides several alternative characterizations of relative amenability. The proof, which we have included for the reader's convenience, is an adaptation to the case of inclusions of algebras of Connes' well known results for single von Neumann algebras [3].

PROPOSITION 5. *With $B \subset N, \tau$ as before, the following conditions are equivalent:*

(i) *B is co-amenable in N.*
(ii) *There exists a state ψ on $\langle N, B \rangle$ with $\psi(uXu^*) = \psi(X)$ for all $X \in \langle N, B \rangle$, $u \in \mathcal{U}(N)$.*

(iii) *For all* $\varepsilon > 0$ *and all finite* $F \subset \mathcal{U}(N)$ *there is a projection* f *in* $\langle N, B\rangle$ *with* $\mathrm{Tr}\, f < \infty$ *such that* $\|ufu^* - f\|_{2,\mathrm{Tr}} < \varepsilon \|f\|_{2,\mathrm{Tr}}$ *for all* $u \in F$.

Here (ii) parallels the invariant mean criterion while (iii) is a Følner type condition.

PROOF. "(i) $\Leftrightarrow$ (ii)" If Ψ is a norm one projection onto N, then by Tomiyama's Theorem it is a conditional expectation onto N; thus $\psi = \tau \circ \Psi$ is $\mathrm{Ad}(u)$-invariant $\forall u \in \mathcal{U}(N)$. Conversely, if ψ satisfies (ii) above and $X \in \langle N, B\rangle$ then let $\Psi(X)$ denote the unique element in N satisfying $\tau(\Psi(X)x)$, $\forall x \in N$. It is immediate to see that $X \mapsto \Psi(X)$ is a conditional expectation (thus a norm one projection) onto N.

"(ii) $\Leftrightarrow$ (iii)" Assuming (iii), denote by I the set of finite subsets of $\mathcal{U}(N)$ ordered by inclusion and for each $i \in I$ let $f_i \in \mathcal{P}(\langle N, B\rangle)$ be a non-zero projection with $\mathrm{Tr}(f_i) < \infty$ such that $\|uf_iu^* - f\|_{2,\mathrm{Tr}} < |i|^{-1}\|f\|_{2,\mathrm{Tr}}$, $\forall u \in i$. For each $X \in \langle N, B\rangle$ let $\psi(X) \stackrel{\text{def}}{=} \lim_i \mathrm{Tr}(Xf_i)/\mathrm{Tr}(f_i)$ for some Banach limit over i. Then ψ is easily seen to be an $\mathrm{Ad}(u)$-invariant state on $\langle N, B\rangle$, $\forall u \in \mathcal{U}(N)$.

Conversely, if ψ satisfies (ii) then by Day's convexity trick it follows that for any finite $F \subset \mathcal{U}(N)$ and any $\varepsilon > 0$ there exists $\eta \in L^1(\langle N, B\rangle, \mathrm{Tr})_+$ such that $\mathrm{Tr}(\eta) = 1$ and $\|u\eta u^* - \eta\|_{1,\mathrm{Tr}} \leq \varepsilon$. By the Powers-Størmer inequality, this implies $\|u\eta^{1/2}u^* - \eta^{1/2}\|_{2,\mathrm{Tr}} \leq \varepsilon^{1/2}\|\eta^{1/2}\|_{2,\mathrm{Tr}}$. Thus, Connes' "joint distribution" trick (see [2]) applies to get a spectral projection f of $\eta^{1/2}$ such that $\|ufu^* - f\|_{2,\mathrm{Tr}} < \varepsilon^{1/4}\|f\|_{2,\mathrm{Tr}}$, $\forall u \in F$. ■

The next result relates the co-amenability of subalgebras with the co-amenability of subgroups, generalizing a similar argument for single algebras/groups from [3].

PROPOSITION 6. *Let* G *be a (discrete) group,* (B_0, τ_0) *a finite von Neumann algebra with a normal faithful tracial state and* $\sigma\colon G \to \mathrm{Aut}(B_0, \tau_0)$ *a trace-preserving cocycle action of* G *on* (B_0, τ_0). *Let* $N = B_0 \rtimes_\sigma G$ *be the corresponding crossed product von Neumann algebra with its normal faithful tracial state given by* $\tau(\sum_{g\in G} b_g u_g) = \tau_0(b_e)$. *Let also* $H < G$ *be a subgroup and* $B = B_0 \rtimes_\sigma H$.

Then B *is co-amenable in* N *if and only if the group* H *is co-amenable in* G.

In particular, taking $B_0 = \mathbf{C}$, we deduce:

COROLLARY 7. *Let* $H < G$ *be (discrete) groups. Then the inclusion of group von Neumann algebras* $L(H) \subset L(G)$ *is co-amenable if and only if* H *is co-amenable in* G.

PROOF OF PROPOSITION 6. "$\Leftarrow$" Let $F_i \nearrow G/H$ be a net of finite Følner sets, which we identify with some sets of representatives $F_i \subset G$. For each i and $X \in \langle N, B\rangle$ set

$$\Psi_i(X) = |F_i|^{-1}\Phi\Big(\big(\sum_{g\in F_i} u_g e_B u_g^*\big)X\big(\sum_{h\in F_i} u_h e_B u_h^*\big)\Big),$$

where Φ: $\operatorname{sp} Ne_B N \to N$ is the canonical operator valued weight given by $\Phi(xe_B y) = xy$ for $x, y \in N$. Then clearly Ψ_i are completely positive, normal, unital, B_0-bimodular and satisfy $\Psi_i(\sum_{g\in G} b_g u_g) = \sum_{g\in G} \alpha_g^i b_g u_g$, where $\alpha_g^i = |gF_i \cap F_i|/|F_i|$. Thus, if we put $\Psi(X) \stackrel{\text{def}}{=} \lim_i \Psi_i(X)$ for some Banach limit over i, then $\Psi(X) = X$ for all $X \in N$, $\Psi(\langle N, B\rangle) \subset N$ and Ψ is completely positive and unital. Thus Ψ is a norm one projection.

"$\Rightarrow$" Let $\{g_i\}_i \subset G$ be a set of representatives for the classes G/H and denote by $\mathcal{A} \subset \langle N, B\rangle$ the set of operators of the form $\sum_i \alpha_i u_{g_i} e_B u_{g_i}^*$, where $\alpha_i \in \mathbf{C}$, $\sup_i |\alpha_i| < \infty$. Clearly, $\mathcal{A} \cong \ell^\infty(G/H)$. Note that $\operatorname{Ad}(u_h)$ implements the left translation by $h \in G$ on $\ell^\infty(G/H)$ via this isomorphism. Put $\varphi \stackrel{\text{def}}{=} \tau \circ \Psi|_{\mathcal{A}}$; then φ is clearly a left invariant mean on $\ell^\infty(G/H)$. ■

COROLLARY 8. *There are inclusions of type* II_1 *factors* $N_0 \subset N \subset M$ *such that* N_0 *is co-amenable in* M *but not in* N. *In fact, there are such examples with* N *having property* (T) *relative to* N_0 *in the sense of [1], [8], [9].*

PROOF. Let Q be an infinite group with property (T) and let $K \lhd H_0 < G$ be as in point (ii) of Theorem 1. Set $N_0 = L(K)$, $N = L(H_0)$, $M = L(G)$. Choosing Q to be ICC ensures that these three algebras are factors of type II_1. The claim follows now from Corollary 7. ■

We proceed now to present an alternative example of interesting co-amenable inclusions of type II_1 factors.

Let $\omega \in \beta\mathbf{N}\backslash\mathbf{N}$ be a free ultrafilter on $\mathbf{N}$. Denote as usual by R the hyperfinite II_1 factor and let $R^\omega \stackrel{\text{def}}{=} \ell^\infty(\mathbf{N}, R)/I_\omega$, where $I_\omega \stackrel{\text{def}}{=} \{(x_n)_{n\in\mathbf{N}} \mid \lim_{n\to\omega} \tau(x_n^* x_n) = 0\}$. Then R^ω is well known to be a type II_1 factor with its unique trace given by $\tau_\omega\big((x_n)_{n\in\mathbf{N}}\big) = \lim_{n\to\omega} \tau(x_n)$. Let further $R_\omega = R' \cap R^\omega$, where R is embedded into R^ω as constant sequences. Recall that R_ω, $R \vee R_\omega$ are factors and $(R \vee R_\omega)' \cap R^\omega = \mathbf{C}$.

Also, if ϑ is an automorphism of R then there exists a unitary element $u_\vartheta \in R^\omega$ normalizing both R and $R \vee R_\omega$ such that $\operatorname{Ad}(u)_{|R} = \vartheta$. The unitary element u_ϑ is unique modulo perturbation by a unitary from R_ω and ϑ is properly outer if and only if $\operatorname{Ad}(u)$ is properly outer on $R \vee R_\omega$ (see Connes [2] for all this). Thus, if σ is a properly outer cocycle action of some group G on R then $\{\operatorname{Ad}(u_{\sigma(g)}) \mid g \in G\} \subset \mathcal{N}(R \vee R_\omega)$ implements a cocycle action $\tilde{\sigma}$ of G on $R \vee R_\omega$ (again, see [2]).

THEOREM 9.

(i) R^ω *is amenable both relative to* R_ω *and to* $R \vee R_\omega$.
(ii) *If* N *is the von Neumann algebra generated by the normalizer* $\mathcal{N}(R \vee R_\omega)$ *of* $R \vee R_\omega$ *in* R^ω, *then* $R \vee R_\omega$ *is not co-amenable in* N. *Moreover, if* σ *is a properly outer action of an infinite property* (T) *group* G *on* R *(e.g. by Bernoulli shifts) then* $N_0 \stackrel{\text{def}}{=} (R \vee R_\omega) \cup \{u_{\sigma(g)} \mid g \in G\})'' \simeq (R \vee R_\omega) \rtimes_{\tilde{\sigma}} G$

is not amenable relative to $R \vee R_\omega$. *In fact,* N_0 *has property* (T) *relative to* $R \vee R_\omega$.

PROOF. By Remark 3 it is sufficient to prove that R^ω is amenable relative to R_ω. We show that in fact a much stronger version of condition (iii) in Proposition 5 holds true, namely: $\forall \mathcal{U}_0 = \{u_n\}_n \subset \mathcal{U}(R^\omega)$ countable, $\exists v \in \mathcal{U}(R^\omega)$ such that $f = ve_{R_\omega}v^*$ satisfies $ufu^* = f$, $\forall u \in \mathcal{U}_0$ or, equivalently, $v\mathcal{U}_0v^* \subset R_\omega$.

Let $u_n = (u_n^k)_k$ with $u_n^k \in \mathcal{U}(R)$, $\forall k, n$. Let $R = \overline{\bigcup_p M_{2^p \times 2^p}(\mathbf{C})}$ and for each n let p_n be such that $\|E_{M_{2^{p_n} \times 2^{p_n}}(\mathbf{C})}(u_m^k) - u_m^k\|_2 \leq 1/n$, $\forall k, m \leq n$. Thus, if we put $P \overset{\text{def}}{=} \Pi_{n \to \omega} M_{2^{p_n} \times 2^{p_n}}(\mathbf{C})$ then $\mathcal{U}_0, R \subset P$. In particular, $P' \cap R^\omega \subset R_\omega$. For each n let now $v_n \subset R$ be a unitary element satisfying $v(M_{2^{p_n} \times 2^{p_n}}(\mathbf{C}))v_n^* \subset M_{2^{p_n} \times 2^{p_n}}(\mathbf{C})' \cap R$ (this is trivially possible, since the latter is a type II_1 factor). Thus, if we let $v = (v_n)_n$ then $vPv^* \subset P' \cap R^\omega \subset R_\omega$, showing that $v\mathcal{U}_0v^* \subset R_\omega$ as well.

The rest of the statement is now trivial, because if we denote by $\mathcal{G}$ the quotient group $\mathcal{N}(R \vee R_\omega)/\mathcal{U}(R \vee R_\omega)$ then N itself is a cross product of the form $(R \vee R_\omega) \rtimes \mathcal{G}$. By Proposition 6 (with $G = \mathcal{G}$ and $H = 1$), the co-amenability of $R \vee R_\omega$ in N amounts to the amenability of the group $\mathcal{G}$. But $\mathcal{G}$ contains copies of any countable non-amenable group (*e.g.* property (T) groups). ∎

REFERENCES

1. C. Anantharaman-Delaroche, *On Connes' property T for von Neumann algebras.* Math. Japon. **32**(1987), 337–355.
2. A. Connes, *Classification of injective factors. Cases* II_1, II_∞, III_λ, $\lambda \neq 1$. Ann. of Math. (2) **104**(1976), 73–115.
3. ———, *On the classification of von Neumann algebras and their automorphisms.* In: Symposia Mathematica, Vol. XX (Convegno sulle Algebre C^* e loro Applicazioni in Fisica Teorica, Convegno sulla Teoria degli Operatori Indice e Teoria K, INDAM, Rome, 1975), Academic Press, London, 1976, 435–478.
4. P. Eymard, *Moyennes invariantes et représentations unitaires.* Lecture Notes in Math. **300**, Springer-Verlag, Berlin, 1972.
5. F. P. Greenleaf, *Amenable actions of locally compact groups.* J. Funct. Anal. **4**(1969), 295–315.
6. B. E. Johnson, *Cohomology in Banach algebras.* Mem. Amer. Math. Soc. **127**, 1972.
7. N. Monod, *Continuous bounded cohomology of locally compact groups.* Lecture Notes in Math. **1758**, Springer, Berlin, 2001.
8. S. Popa, *Correspondences.* INCREST preprint, 1986.
9. ———, *Some properties of the symmetric enveloping algebra of a subfactor, with applications to amenability and property T.* Doc. Math. **4**(1999), 665–744 (electronic).
10. R. J. Zimmer, *Amenable pairs of groups and ergodic actions and the associated von Neumann algebras.* Trans. Amer. Math. Soc. **243**(1978), 271–286.

University of Chicago
Chicago, IL 60637
USA
email: monod@math.uchicago.edu

UCLA
Los Angeles, CA 90095
USA
email: popa@math.ucla.edu

C. R. Math. Rep. Acad. Sci. Canada Vol. 25, (3), 2003 pp. 88–91

DIVERGENCE THEOREM FOR SETS OF FINITE PERIMETER

ISIDORE FLEISCHER

Presented by Niky Kamram, FRSC

ABSTRACT. This is a development of the theorem which uses only bare measure theory (and a smidgen of functional analysis), avoiding recourse to the heavy machinery of "mollifiers, (co)area formulae, (relative) isoperimetric inequalities, and the Whitney extension theorem" that the other published treatments are based on. The basic difficulty is to identify the boundary measure, obtained from formal integration by parts from the volume integral over the measurable set of finite perimeter, with Hausdorff measure. This can be achieved with nothing more exotic than an appeal to the compactness in L_1 of bounded BV-norm sets, the Radon-Nikodym and monotone convergence theorems, Fubini and a little multi-dimensional calculus.

RÉSUMÉ. Une dérivation du théorème qui n'utilise que la pure théorie de la mesure (et un brin d'analyse fonctionnelle). Le noyau de la démonstration consiste à reconnaître que la mesure sur la frontière réduite, obtenue par une intégration par parties formelles à partir de l'intégrale de volume sur l'ensemble de périmetre fini, se confonde avec la mesure de Hausdorff. Ceci est achevé par un appel à la compacité en L_1 des ensembles bornés en BV-norme, les théorèmes de Radon-Nikodym et de convergence monotone, Fubini, et un peu de calcul multi-dimensionnel.

In the classical formulation of the divergence theorem,

$$\int_E \nabla \cdot \phi \, dx = \int_{\partial E} \phi \cdot n \, dH,$$

E is an N-dimensional volume bounded by a smooth $(N-1)$ surface ∂E to which n is the exterior unit normal and on which H is $(N-1)$ (hyper)surface measure.

For any compactly supported C' selfmap ϕ of N-space the left side makes sense as soon as E is Lebesgue measurable and furnishes a linear functional on the space of these selfmaps; moreover, the right side bounds the norm of the functional (*i.e.*, its sup on $|\phi| \leq 1$) by the surface measure of ∂E (which is actually for bounded E the sup, since $\phi \cdot n$ is largest when ϕ has the direction of n, when it equals $|\phi|$). Accordingly, declare E to have *finite perimeter* if $\int_E \nabla \cdot \phi \, dx$ is bounded on compactly supported $|\phi| \leq 1$ in C'.

As a function of R^N selfmaps, this form can be construed as a vector (-valued) Radon measure acting componentwise: $\nabla \chi_E = \Sigma e_i (\nabla_i \chi_E)$, $\nabla_i \chi_E \varphi := \nabla \chi_E \varphi_i$

Received by the editors on March 17, 2003.
AMS subject classification: 26B15, 26B20, 28A25, 28A78.

is the restriction to the i-th component of φ, and the bound over $|\phi| \leq 1$ is the value at R^N of a positive (real) Radon, its "variation", measure $\| \nabla \chi_E \|$, whose values on opens are the sup of $\nabla \chi_E$ over $|\phi| \leq 1$ with compact support in the open. (Indeed, these values are continuous on increasing sequences of, and (countably)·additive on disjoint, opens.) Similarly for $\nabla \chi_E$, whose value on ball B is $\Sigma_i \lim \nabla_i \chi_E \phi_n$, $\phi_n \uparrow \chi_B$ pointwise, which determines E a.e. There is continuity of $\| \nabla \chi \|$ for argumentwise convergence of $\nabla \chi_E$, in open U whose boundary has $\| \lim \nabla \chi_E \|$-measure zero. Domination by lim inf commutes with sup: If $F(\phi) \leq \liminf_j F_j(\phi)$ then $\leq \liminf_j \vee_\phi F_j(\phi)$ for every ϕ, hence the latter is $\geq \vee_\phi F(\phi)$. So from $\nabla \chi_E \leq \liminf_j \nabla \chi_{E_j}$ on the ϕ supported in open U follows $(\| \nabla \chi_E \|)U \leq \liminf_j (\| \nabla \chi_{E_j} \|)U$. Dually, from $\nabla \chi_E \geq \limsup \nabla \chi_{E_j}$ on the $\phi = 1$ on closed C follows $\| \nabla \chi_E \| C \geq \limsup \| \nabla \chi_{E_j} \| C$. Hence if also $C \supset U$ with $\| \nabla \chi_E \| (C \backslash U) = 0$, then $\| \nabla \chi_E \| U = \| \nabla \chi_E \| C = \lim \| \nabla \chi_{E_j} \| C = \lim \| \nabla \chi_{E_j} \| U$. This ia a variant of what is called the "Portmanteau theorem" in the probability literature, *cf.* [EG, Section 1.9].

The vector measure $\nabla \chi_E$ is (*i.e.*, its components are) absolutely continuous with respect to its variation $\| \nabla \chi_E \|$ so one can form the Radon-Nikodym derivative, which will be a vector-valued point function $\nu_E(x)$. To see to what it corresponds, recall that for a smoothly bounded volume E the vector measure $\int_E \nabla \cdot \varphi \, dx = \int_{\partial E} \varphi \cdot n \, dH$ has components $\int_{\partial E} \phi_i n_i \, dH$, the variation measure is surface measure H on ∂E, and so the vector derivative is n. This suggests defining the unit normal in general as the derivative. However, the derivative is only determined up to a set of $\| \nabla \chi_E \|$-measure zero on which it can take arbitrary values—to nail it down, we'll restrict to those x for which $\| \nabla \chi_E \| [B(x,r)] > 0$ for arbitrarily small $r > 0$ and such that $\nabla \chi_E [B(x,r)] / \| \nabla x_E \| [B(x,r)]$ approaches a vector of norm 1 as $r \to 0$. This subset of x's is sometimes called the "reduced boundary" $\partial^- E$ and the negative (since χ_E increases when entering E) of the limit, the "generalized exterior normal" $\nu_E(x)$.

Since $\nabla \chi_E$ vanishes on every $\phi \in C'$ supported either in the interior or the exterior of E (for the former use the classical divergence theorem) the reduced boundary $\partial^- E \subset \partial E$, the topological boundary (not conversely, as witnessed by sets of measure zero). Thus $\nabla \chi_E$ and $\| \nabla \chi_E \|$ are effectively "boundary measures"—indeed, they are even carried by the reduced boundary $\partial^- E$ since by standard measure theory the complement of the reduced boundary has $\| \nabla \chi_E \|$-measure zero, hence also $\nabla \chi_E$-measure zero. This yields a kind of Divergence Theorem for E of finite perimeter:

$$\int_E \nabla \cdot \phi \, dx = \nabla \chi_E(\phi) = \int_{\partial^- E} \phi \cdot \nu_E \, d \| \nabla \chi_E \|.$$

It would be more satisfying to be able to replace in general the abstract measure $\| \nabla \chi_E \|$ by the standard Hausdorff surface measure in the right term. This is what the subsequent development will accomplish; the resulting equality acheves the divergence theorem announced in the title.

Not only does the above equality reduce globally to the classical one for smoothly bounded domains, it also melds correctly locally on intersections with such domains. Let the domain D be smoothly bounded with exterior normal n, let f be a real-valued C' function which approximates some continuous function which differs from 1 in D and 0 outside only close to the boundary of D where it is linear on each normal there, and apply the above equality to obtain $\int_E f \nabla \cdot \phi\, dx + \int_E \nabla f \cdot \phi\, dx = \int_E \nabla \cdot (f\phi)\, dx = \int_{\partial^- E} f\phi \cdot \nu\, d\| \nabla \chi_E \|$. As the regions of constancy expand to force the non-constant part of f to approach the negative unit jump across the boundary of D, this becomes

$$\int_{E\cap D} \nabla \cdot \phi\, dx - \int_{E\cap\partial D} \phi \cdot n\, dH = \int_{D\cap\partial^- E} \phi \cdot \nu_E\, d\| \nabla \chi_E \|$$

which is the Divergence Theorem for $E \cap D$ with each portion of the boundary integral evaluated according to its own rule. The passage to the limit in the outer terms, whose integrands have f as a factor, can be justified by the Monotone Convergence Theorem; for the middle term, parametrize the domain of non-constancy of f with rectangular regions and apply Fubini to the product of length along the normal times surface measure.

From now on E will be assumed closed. This results in no loss of generality since $\nabla\chi_E = \nabla\chi_{\bar{E}}$.

Setting ϕ equal to C' vector functions constantly basis vectors on a neighborhood of closed $E \cap D$ yields the vector equality

$$-\int_{E\cap\partial D} n\, dH = \int_{D\cap\partial^- E} \nu_E\, d\| \nabla \chi_E \| = (\nabla\chi_E)D.$$

With D the (closed) centered balls of radius $r \downarrow 0 \in \partial^- E$, $\int_{E\cap\partial D} n\, dH / \|\nabla\chi_E\|(D)$ approaches $\nu(0)$ of unit length so $\liminf_{r\downarrow 0} H(E\cap\partial D)/\|\nabla\chi_E\|(D) \geq 1$; with D balls B of sufficiently small radius r, $\|\nabla\chi_E\|(B) \leq (1+\epsilon)H(E\cap\partial B) \leq (1+\epsilon)H(\partial B) = Mr^{N-1}$ for some constant M.

The change of variable $x = ry$ with constant (scalar) r converts the form $\int_E \nabla_x \cdot \phi(x)\, dx$ to $r^{N-1} \int_{r^{-1}E} \nabla_y \cdot \phi(ry)\, dy$; for B a bounded open ball centered at $0 \in \partial^- E$, $\phi(x)$ has support in rB iff $\phi(ry)$ has support in B, hence $\|\nabla\chi_E\|(rB) = r^{N-1}\| \nabla \chi_{r^{-1}E} \|B$, $\frac{1}{r^{N-1}}\| \nabla \chi_E \| rB = \| \nabla \chi_{r^{-1}E} \| B$ is bounded for small r.

It is known that BV-norm (*i.e.*, variation $+L_1$) bounded sets are precompact in $L_1(B)$; hence there is a sequence $r_j \to 0$ (indeed every zero convergent sequence contains such a subsequence) for which $\chi_{r_j^{-1}E}$ converges in measure on B and so some further subsequence converges a.e.—hence the limit is a.e. in B some characteristic function χ_F.

To identify the set F for which $\chi_F|B$ is the limit of $\chi_{r_j^{-1}E}|B$ for sufficiently small balls centered at 0, take the positive x_1-axis along the generalized exterior normal to E at 0, dot product with which then shows convergence as $r_j \downarrow 0$ of the scalar function $(\nabla_1\chi_{r_jE})B/\| \nabla \chi_{r_jE} \|B$ to -1. Thus $(\nabla_1\chi_F)B = \lim(\nabla_1\chi_{r_j^{-1}E})B = -\lim \| \nabla \chi_{r_j^{-1}E} \| B$ and since sup of a limit is $\leq$ lim inf of the sups (as noticed previously), $\| \nabla \chi_F \| B \leq \lim(\inf) \| \nabla \chi_{r_j^{-1}E} \| B = -(\nabla_1\chi_F)B \leq \| \nabla \chi_F \| B$. Thus χ_F is a.e. a function only of x_1 ($\nabla\chi_F$ annihilates $\partial\varphi/\partial x_i$

for $i > 1$) and since $\nabla_1 \chi_F = -\| \nabla \chi_F \|$ is non-positive, F is a.e. of the form $\{x_1 < \lambda\}$. (If F had two boundary points in whose neighborhoods it had positive measure, one could find an interval on which $\| \nabla \chi_F \| > - \nabla_1 \chi_F$.)

It can be shown (but is not needed for the argument) that $\lambda = 0$ so that $T := \{x_1 = \lambda\}$ is really the "measure-theoretic tangent" hyperplane, *i.e.*, that through 0, rather than through an unspecified λ, orthogonal to the normal. Since F is smoothly bounded by T, $\| \nabla \chi_F \|$ is hyperplane measure, thus vanishes on the boundary of the meet of every ball B with T; by the "Portmanteau theorem" $\| \nabla \chi_{r_j^{-1} E} \| B = 1/r_j^{N-1} \| \nabla \chi_E \| (r_j B) \to \| \nabla \chi_F \| B$—for B a ball centered in T, the latter is the surface measure of its diameter. Thus at every point of $\partial^- E$, the $\| \nabla \chi_E \|$ measure of the centered ball of radius r divided by the surface measure of its diametral section approaches 1 as $r \downarrow 0$. Thus the derivative is identically 1 and surface measure coincides with $\| \nabla \chi_E \|$ on $\partial^- E$.

References

[NB] N. Bourbaki, *Integration*. Chapitre 3.

[EG] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*. CRC Press, 1992.

[HF] H. Federer, *Geometric measure theory*. Springer-Verlag, 1969.

[IF] I. Fleischer, *L_1 compactness of bounded BV sets*. arXiv:math. FA/0212198 v1, 15 Dec. 2002.

[FF] W. Fleming, *Functions of several variables*. Springer-Verlag, 1977.

[GE] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*. Birkhäuser, 1984.

[VM] V. G. Maz'ja, *Sobolev Spaces*. Springer-Verlag, 1985.

[FM] F. Morgan, *Geometric measure theory, A beginners guide*. Academic Press, 1988.

[MM] M. E. Munroe, *Introduction to measure and integration*. Addison-Wesley, 1953.

[SS] S. Saks, *Theory of the Integral*. Stechert, 1937.

[WZ] W. P. Ziemer, *Weakly differentiable functions*. Springer-Verlag, 1989.

Centre de recherches mathématiques
Université de Montréal
C.P. 6128, Succ. Centre-Ville
Montréal, QC
H3C 3J7
email: fleischi@CRM.UMontreal.CA

C. R. Math. Rep. Acad. Sci. Canada Vol. 25, (3), 2003 pp. 92–96

PARAMETER ESTIMATION IN A MULTILEVEL HIERARCHY OF SDES

G. GAUTHIER, D. KAKO AND J. VAILLANCOURT

Presented by Donald A. Dawson, FRSC

ABSTRACT. We obtain the maximum likelihood estimators for the drift coefficients of a hierarchical model of stochastic processes arising as an extension of the mean field model. The closed form of said estimators is explicited and allows for its use in practice. A central limit theorem for these coefficients is derived.

RÉSUMÉ. Nous obtenons explicitement les estimateurs de vraisemblance maximale pour les coefficients de dérive d'un modèle hiérarchique de processus stochastiques généralisant le modèle de champ moyen. La forme fermée des estimateurs permet une utilisation aisée en pratique. Un théorème de la limite centrale pour ces coefficients estimés est aussi obtenu.

1. **The multilevel system.** For the sake of simplicity, the results will be formulated in the case of a model with a fixed number L of levels in the hierarchy, each level having (common) size N describing the structure of the system at that level or strata. While this assumption of a common size is artificial, it will allow many simplifications in the statement of the results. On an appropriate complete probabilized stochastic basis $(\Omega, \mathcal{F}, \{\mathcal{F}_t : t \geq 0\}, \mathcal{P})$, consider the following stochastic system

$$(1)\quad dX_i^N(t) = \sum_{k=1}^{L} C_k^N[\overline{X}_{i,k}^N(t) - X_i^N(t)]dt + C_{L+1}^N[\alpha - X_i^N(t)]dt + \gamma^N dW_i(t)$$

where the multivariate index $i = (i_1, i_2, \ldots, i_L)$ runs over all values of $i_k \in \{1, 2, \ldots, N\}$ for every $k \in \{1, 2, \ldots, L\}$. In these equations, each one of the N^L processes $X_i^N(t)$ is one dimensional and the standard brownian motions $\{W_i(t), t \geq 0\}$ are also one dimensional and mutually independent. The parameters C_k^N with $k \in \{1, \ldots, L\}$, C_{L+1}^N, γ^N and α are real-valued and we define $\overline{X}_{i,k}^N(t) = \frac{1}{N^k} \sum_{i_1, i_2, \ldots, i_k = 1}^{N} X_i^N(t)$ and $\overline{X}_{i,0}^N(t) = X_i^N(t)$.

The system therefore comprises N^L elements with an associated value of $X_i^N(t)$ at time t and classified into categories (say, regions, type of product, sizes, *etc.*). At the first level, the N^{L-1} hierarchical averages $\overline{X}_{i,1}^N(t)$ (which are independent of index i_1 but dependent on all other i_k's) are calculated over only one

Received by the editors on March 14, 2003.

The first author's research was supported by the National Science and Engineering Research Council of Canada and the Institut de Finance Mathématique de Montréal. The third author's research was supported by the National Science and Engineering Research Council of Canada and Valorisation Recherche Québec.

AMS subject classification: Primary: 60K35; secondary: 62M05.

of these categories; at the second level, the N^{L-2} averages $\overline{X}_{i,2}^{N}(t)$ now involve averaging over two of them (i_1 and i_2) and leaving a dependency on i_3, i_4 and so on; more generally, at level k, the N^{L-k} averages $\overline{X}_{i,k}^{N}(t)$ involve averaging over components $\{i_1, i_2, \ldots, i_k\}$ and keeping fixed those amongst $\{i_{k+1}, i_{k+2}, \ldots, i_L\}$. This continues until we reach level L where the overall average value $\overline{X}_{i,L}^{N}(t)$ of all the elements in the system is obtained. From now on index i will be dropped in this last case, since the dependency on any component of vector i has been completely averaged out at the top level.

Realistically, interaction amongst neighbouring elements is more intense than between elements belonging to very different categories, something which one could model by putting $C_k^N = C_k/N^{k-1}$, thus emphasizing a decreasing intensity of the impact of the fluctuations of a single element upon the movement of higher level averages of families of elements it belongs to. Once again the choice of parameters has been made in order to get simpler formulas but there is no difficulty in converting back to the more natural C_k's.

Our system (1) has Lipschitz coefficients and hence possesses a unique strong solution for every starting value (see Lipster and Shiryayev [7]). Another useful feature is that the overall average (but not the intermediate ones!) is actually an Ornstein-Uhlenbeck process, namely $\overline{X}_L^N(t) = \alpha + e^{-C_{L+1}^N t}\left[(\overline{X}_L^N(0) - \alpha) + \gamma^N N^{-L/2} \int_0^t e^{C_{L+1}^N s}\, dW(s)\right]$, where $W(s)$ denotes the appropriate standard brownian motion. This feature makes it possible to come up with a sensible estimator for parameter α.

Throughout the rest of this paper, we assume that diffusion parameter γ^N is known, the problem of estimating it being the subject of ongoing research.

Finally, let us point out that estimation of the parameters in system (1) goes back to Kasonga [5], who achieved it in the mean field case, where there is only one level ($K = 1$), no global market drift $\alpha = 0$ and unitary volatility $\gamma^N = 1$. While the asymptotic behavior of hierarchies of diffusions has been analyzed extensively in mathematical genetics (as in papers [2] and [3], where related models are studied), the elegant closed form of the maximum likelihood estimator (MLE) that characterizes our model is stated here explicitly for the first time. The proofs of all the results stated here will be published elsewhere and are available upon request.

2. The main results.

We first obtain the closed form for the MLE $\widehat{\theta}$ of column vector $\theta = (C_1^N, C_2^N, \ldots, C_d^N)^t$ when the values for α and γ^N are given. We then state a central limit theorem for this estimator. Finally, we give the Euler approximation formulas (corresponding to the more realistic instance when data collection is effected at discrete time points) and explain how to go about estimating α in this particular context.

Since our system is linear in θ, the normal equations that characterize the MLE (for instance, see Barndorff-Nielsen and Sorensen [1]) are given in matrix form by $F(X,t)\widehat{\theta} = V(X,t)$, where $(L+1) \times (L+1)$ matrix $F(X,t) = [F_{qp}(X,t)]$

and vector $V(X,t) = [V_k(X,t)]$ turn out to have a very simple form: for all $k, p, q \in \{1, 2, \dots, L+1\}$ we can write

$$(2) \qquad F_{qp}(X,t) = \sum_{i_1, i_2, \dots, i_L = 1}^{N} \int_0^t \left(\overline{X}^N_{i,\min(p,q)}(s) - X^N_i(s)\right)^2 ds$$

and

$$(3) \qquad V_k(X,t) = \sum_{i_1, i_2, \dots, i_L = 1}^{N} \int_0^t \left(\overline{X}^N_{i,k}(s) - X^N_i(s)\right) ds$$

where we set $\overline{X}^N_{i,L+1}(s) = \alpha$ for all s. Since there holds $F_{pq}(X,t) = F_{qp}(X,t)$, we write this common value as $F_k(X,t)$ from now on with the choice $k = \min(p,q)$. Here and in the sequel, $X = \{X^N(t) : t \geq 0\}$ denotes matrix-valued diffusion process $X^N(t) := [X^N_i(t)]$.

LEMMA 2.1. *For any starting value $F(X,0)$ and any time $t > 0$, matrix $F(X,t)$ defined above is symmetric and invertible. Further, its inverse $F(X,t)^{-1} = [G_{q,p}(X,t)]$ is tridiagonal and given by formulas $G_{k,k-1} = -(F_k - F_{k-1})^{-1}$, $G_{k,k} = (F_k - F_{k-1})^{-1} + (F_{k+1} - F_k)^{-1}$ and $G_{k,k+1} = -(F_{k+1} - F_k)^{-1}$, each of them being well-defined for those values of k such that both indices fall in the given range, as long as t remains positive. All other terms in this inverse are null.*

The MLE is therefore determined by $\widehat{\theta} = F(X,t)^{-1}V(X,t)$ and uniquely so, provided of course one uses all the information at hand and selects the largest value of t available.

THEOREM 2.2. *The MLE $\widehat{\theta} = (\widehat{C^N_1}, \widehat{C^N_2}, \dots, \widehat{C^N_{L+1}})^t$ is given by*

$$(4) \qquad \widehat{C^N_{L+1}} = \left(\frac{V_{L+1} - V_L}{F_{L+1} - F_L}\right)$$

and for each subsequent $k = 1, 2, \dots, L$

$$(5) \quad \widehat{C^N_k} = \left(\frac{V_k - V_{k-1}}{F_k - F_{k-1}}\right) - \left(\frac{V_{k+1} - V_k}{F_{k+1} - F_k}\right) = \left(\frac{V_k - V_{k-1}}{F_k - F_{k-1}}\right) - \sum_{j=k+1}^{L+1} \widehat{C^N_j}$$

by putting $V_0 = F_0 = 0$ in the last case $k = 1$. This MLE (which is a process in t) is also strongly consistent as N goes to infinity, uniformly over compact time sets.

The recursive nature of formula (5) is what allows us to get the Euler approximation in an immediately usable form with no additional conditions. In practice, the full integrals appearing in formulas (2) and (3) above are not known since data is often collected at fixed times and not continuously. The two natural approaches that come to mind to solve this difficulty (discretization of the integrals above and modelling the phenomenon as observed at fixed times and computing the MLE for this discrete model) turn out to yield the same answer here and so we only present the first one, known as the Euler approximation scheme. The equality of these two solutions implies in particular that formulae (4) and (5) remain true when the integrals are replaced by approximating Riemann sums and we make use of this in the sequel.

Assume the process is observed at times $0 < \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_M < T$ for some fixed T. Following Kloeden and Platen [6], it is easy to see that the approximation is given for $i = \{i_1, i_2, \ldots, i_L\}$ and $n \in \{0, \ldots, M-1\}$ by

$$(6) \qquad X_i^N(\tau_{n+1}) - X_i^N(\tau_n) = a_n\big(\tau_n, X_i^N(\tau_n)\big)(\tau_{n+1} - \tau_n) + \gamma^N\big(W_i^N(\tau_{n+1}) - W_i^N(\tau_n)\big)$$

with drift coefficient $a_n\big(\tau_n, X_i^N(\tau_n)\big) = \sum_{k=1}^{L} C_k^*[\overline{X}_{i,k}^N(\tau_n) - X_i^N(\tau_n)] + C_{L+1}^*$ $[\alpha - X_i^N(\tau_n)]$.

The Euler approximation is therefore given by the application of the second half of formula (5) in conjunction with the following, simplified expressions (with d indicating we are now using the discretized integrals in (2) and (3) above). For the reader's convenience, we make the expressions known in final form.

$$(7) \quad \widehat{C_{L+1}^{N,d}} = \left(\frac{V_{L+1}^d - V_L^d}{F_{L+1}^d - F_L^d}\right) = \frac{\sum_{n=0}^{M-1}\big(\alpha - \overline{X}_L^N(\tau_n)\big)\big(\overline{X}_L^N(\tau_{n+1}) - \overline{X}_L^N(\tau_n)\big)}{\sum_{n=0}^{M-1}\big(\alpha - \overline{X}_L^N(\tau_n)\big)^2(\tau_{n+1} - \tau_n)}$$

and, for $k = 1, 2, \ldots, L+1$,

$$(8) \quad \left(\frac{V_{k+1}^d - V_k^d}{F_{k+1}^d - F_k^d}\right) = \frac{\sum_{n=0}^{M-1}\sum_{i_{k+1},\ldots,i_L=1}^{N}\big(\overline{X}_{i,k+1}^N(\tau_n) - \overline{X}_{i,k}^N(\tau_n)\big)\big(\overline{X}_{i,k}^N(\tau_{n+1}) - \overline{X}_{i,k}^N(\tau_n)\big)}{\sum_{n=0}^{M-1}\sum_{i_{k+1},\ldots,i_L=1}^{N}\big(\overline{X}_{i,k}^N(\tau_n) - \overline{X}_{i,k+1}^N(\tau_n)\big)^2(\tau_{n+1} - \tau_n)}.$$

The only remaining difficulty here occurs when α is unknown, as will generally be the case. Since the overall average turns out to be an Ornstein-Uhlenbeck process, we have at once the following representation for its expectation: $E\overline{X}_L^N(T) = \alpha + e^{-C_{L+1}^N T}\big(E\overline{X}_L^N(0) - \alpha\big)$. It is therefore sensible to estimate $E\overline{X}_L^N(T)$ using some global average $\overline{X}^{\text{IND}}(T)$ extracted from other available data. If we now substitute $\widehat{C_{L+1}^{N,d}}$ for C_{L+1}^N in the last equation, we propose the following procedure: use

$$(9) \qquad \overline{X}^{\text{IND}}(T) = \widehat{\alpha} + e^{-\widehat{C_{L+1}^{N,d}}T}\big(\overline{X}^{\text{IND}}(0) - \widehat{\alpha}\big)$$

in conjunction with formula (7) to first extract both $\widehat{\alpha}$ and $\widehat{C_{L+1}^{N,d}}$ from the nonlinear system thus obtained (possibly using some numerical scheme there); then use this value for $\widehat{C_{L+1}^{N,d}}$ in formula (5) so as to complete the iterative estimation procedure.

Using results of Ibragimov and Hasminskii [4], one can finally derive the fluctuations of the MLE around its intended target θ. We assume anew that α is known.

CENTRAL LIMIT THEOREM 2.3. *The MLE* $\widehat{\theta} = (\widehat{C_1^N}, \widehat{C_2^N}, \ldots, \widehat{C_{L+1}^N})^t$ *given by formulae (4) and (5) is asymptotically normal, uniformly over compact time sets, in the sense that under true law* P_θ *the distribution of* $N^{L/2}I(N,t)^{-1/2}(\gamma^N)^{-1} \times [\widehat{\theta} - \theta]$ *is aymptotically multivariate gaussian with zero mean and the identity for covariance matrix, uniformly in* $t \in [0,T]$ *as* N *goes to infinity. Here* $I(N,t) = [I_{qp}(N,t)]$ *is defined by* $I_{qp} = G - H_{\min(p,q)}$ *with*

$$G(t) = N^{-L} \int_0^t \sum_{i_1,\ldots,i_L=1}^{N} [X_i^N(s)]^2 \, ds,$$

$H_k(t) = N^{k-L} \int_0^t \sum_{i_{k+1},\ldots,i_L=1}^{N} [\overline{X}_{i,k}^N(s)]^2 \, ds$ *for each* $k \in \{1, 2, \ldots, L\}$ *and finally* $H_{L+1}(t) = \int_0^t 2\alpha[\overline{X}_L^N(s)] - \alpha^2 \, ds$.

REFERENCES

1. O. E. Barndorff-Nielsen and M. Sorensen, *A review of some aspects of asympotics likelihood theory for stochastic processes.* Internat. Statist. Rev. (1) **62**(1994), 133–165.
2. D. A. Dawson and A. Greven, *Multiple time scale analysis of interacting diffusions.* Probab. Theory Related Fields **95**(1993), 467–508.
3. G. Gauthier, *Multilevel bilinear system of stochastic differential equations.* Stochastic Process. Appl. **67**(1997), 117–138.
4. I. A. Ibragimov and R. Z. Khasminskii, *Statistical Estimation, asymptotic theory.* Heldermann, 1984.
5. R. A. Kasonga, *Estimation of parameter of stochastics differential equations.* PhD thesis, Carleton University, 1986.
6. P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations.* Springer-Verlag, Berlin, 1992.
7. R. S. Lipster and A. N. Shiryayev, *Statistics of random processes.* 2 volumes. Springer Verlag, Berlin, 1977–78.

Service de l'enseignement
des méthodes quantitatives de gestion
École des Hautes Études Commerciales
3000, chemin de la Côte-Sainte-Catherine
Montréal, Québec
H3T 2A7
email: genevieve.gauthier@hec.ca

Revenu Canada
410, avenue Laurier ouest
Ottawa, Ontario
K1A 0L5
email: danokako@hotmail.com

Décanat de la recherche
Université du Québec en Outaouais
283, boulevard Alexandre-Taché, AT-E-2600
Hull, Québec
J8X 3X7
email: jean.vaillancourt@uqo.ca