Paul-André Meyer

Le théorème de convergence des martingales dans les variétés riemanniennes d'après R.W. Darling et W.A. Zheng

Séminaire de probabilités (Strasbourg), tome 17 (1983), p. 187-193

<http://www.numdam.org/item?id=SPS_1983__17__187_0>

LE THEOREME DE CONVERGENCE DES MARTINGALES DANS LES VARIETES RIEMANNIENNES

d'après R.W. DARLING et W.A. ZHENG

par P.A. Meyer

I. INTRODUCTION

Considérons une martingale continue réelle (X_t), nulle en 0 pour fixer les idées, et le processus croissant associé $\langle X,X\rangle_t$. Il est bien connu que les deux ensembles

$$C_1 = \{ \omega : X_\infty(\omega) = \lim_t X_t(\omega) \text{ existe dans } \mathbb{R} \}$$
$$C_2 = \{ \omega : \langle X,X\rangle_\infty < \infty \}$$

ne diffèrent que par un ensemble négligeable. On se propose d'établir ce résultat pour des martingales (locales) continues à valeurs dans une variété riemannienne V . Voici les énoncés obtenus par Darling et Zheng ; les termes exigeant une définition seront soulignés, et expliqués ensuite.

Soit (X_t) une martingale (locale) à valeurs dans V, à trajectoires continues, et soit $\langle X,X\rangle_t$ son processus croissant associé. Alors :

Théorème 1 (Darling). Sur l'ensemble où $\langle X,X\rangle_\infty < \infty$, X_∞ existe p.s. dans le compactifié d'Alexandrov de V.

Théorème 2 (Zheng). Sur l'ensemble où X_∞ existe et appartient à V, on a p.s. $\langle X,X\rangle_\infty < \infty$.

Je voudrais ici rendre ces énoncés compréhensibles pour les probabilistes, en laissant de côté certains détails (que l'on trouvera bien sûr dans les articles définitifs de R.W.R. Darling et W.A. Zheng, à paraître). Les éléments de géométrie différentielle nécessaire sont donnés dans le Sém. Prob. XV, ou dans le Sém. Prob. XVI (article de L. Schwartz). On va revoir dans un cas particulier les notions indispensables.

EXPLICATION DES TERMES : UN CAS PARTICULIER

Supposons que V soit $\mathbb{R}^n$; nous voulons exprimer les résultats usuels sur les martingales locales continues, non pas dans le système usuel de coordonnées linéaires $(u^1,\dots,u^n)$, mais dans un système quelconque de coordonnées curvilignes, $(x^1,\dots,x^n)$. Soit donc (X_t) une semimartingale à valeurs dans V, à trajectoires continues. Nous posons $U_t^i=u^i\circ X_t$, $X_t^i = x^i\circ X_t$ (coordonnées du même processus dans deux cartes différentes).

1) Le « principe de Schwartz » . Soit f une fonction C^∞ sur V. La

formule d Ito nous dit que le processus $(f\circ X_t)$ est une semimartingale réelle (en particulier, les U_t^i, X_t^i sont des semimartingales réelles), et nous avons (en écrivant Ito, puis prenant un compensateur prévisible)

$$\begin{aligned} d(f\circ X_t)^{\sim} &= \Sigma_i\ D_i f(X_t)d\tilde{X}_t^i + \tfrac{1}{2}\Sigma_{ij}\ D_{ij}f(X_t)d\langle X^i,X^j\rangle_t \quad \text{(carte curviligne)} \\ &= \Sigma_i\ \overline{D}_i f(X_t)d\tilde{U}_t^i + \tfrac{1}{2}\Sigma_{ij}\ \overline{D}_{ij}f(X_t)d\langle U^i,U^j\rangle_t \quad \text{(carte linéaire)} \end{aligned}$$

avec bien sûr $D_i f=\partial f/\partial x^i$, $\overline{D}_i f=\partial f/\partial u^i$... Puisque cela est vrai pour toute fonction f, nous pouvons écrire formellement

$$(1)\qquad \Sigma\ d\tilde{X}_t^i D_i + \tfrac{1}{2}d\langle X^i,X^j\rangle_t D_{ij} = \Sigma\ d\tilde{U}_t^i\ \overline{D}_i + \tfrac{1}{2}d\langle U^i,U^j\rangle_t\overline{D}_{ij}$$

les deux membres étant considérés comme des opérateurs différentiels du second ordre au point X_t . Je noterai $d^2\tilde{X}_t$ cet opérateur d'ordre 2.

EXEMPLE. Dire que (X_t) est un mouvement brownien revient à dire que cet opérateur différentiel formel est égal à $\frac{1}{2}\Delta dt$ (au point X_t). Plus généralement, dire que (X_t) est une diffusion gouvernée par un opérateur du second ordre

$$(2)\qquad \begin{aligned} L_t f &= \Sigma\ \ \overline{a}^i(t,u)\overline{D}_i f+ \overline{a}^{ij}(t,u)\overline{D}_{ij}f \quad \text{(carte linéaire)} \\ &= \Sigma\ \ a^i(t,x)D_i f+ a^{ij}(t,x)D_{ij}f \quad \text{(carte curviligne)} \end{aligned}$$

revient à dire que $d^2\tilde{X}_t=L(t,X_t)dt$. Cette écriture ne dépend pas de la carte.

2) <u>Le processus croissant associé à X.</u>

Dans la carte linéaire, il s'écrit $d\langle X,X\rangle_t = \Sigma_i\ d\langle U^i,U^i\rangle_t$. On interprète cela comme une métrique riemannienne en tout point de $\mathbb{R}^n$

$$\Sigma_i\ (du^i)^2 = \Sigma_{ij}\ g_{ij}(x)dx^i dx^j$$

et alors on a l'expression invariante pour le même processus croissant

$$(3)\qquad d\langle X,X\rangle_t = \Sigma_{ij}\ g_{ij}(X_t)d\langle X^i,X^j\rangle_t \quad .$$

3) <u>Définition intrinsèque des martingales locales</u>. Dans la carte linéaire, dire que (X_t) est une martingale locale signifie que $d\tilde{U}_t^i$ =0. Par exemple, si (X_t) est une diffusion gouvernée par L_t, cela revient à dire que l'opérateur L_t, écrit dans la carte linéaire, est <u>purement d'ordre</u> 2. Mais cela n'est pas invariant par changement de cartes. Il faut interpréter ainsi l'opération << prendre la partie d'ordre un d'un opérateur >> :

En tout point x, on se donne une <u>application linéaire</u> Γ des opérateur d'ordre 2 en x dans les opérateurs d'ordre 1, qui est <u>l'identité</u> sur les opérateurs d'ordre 1. Dans la carte linéaire, Γ s'écrit

$$\Gamma(\overline{D}_i)=\overline{D}_i \quad , \quad \Gamma(\overline{D}_{ij})=0$$

mais dans la carte curviligne, elle s'écrit avec des coefficients

$$(4)\qquad \Gamma(D_i) = D_i \quad , \quad \Gamma(D_{ij}) = \Sigma_k\ \Gamma_{ij}^k(x)D_k$$

Dans ces conditions, dire que (X_t) est une martingale locale au sens usuel s'écrit $\Gamma(d^2\tilde{X}_t)=0$, soit dans la carte curviligne

(5) $$\forall\ k\ ,\ d\tilde{X}_t^k + \Sigma_{ij}\ \frac{1}{2}\Gamma_{ij}^k(X_t)d\langle X^i,X^j\rangle_t = 0\ .$$

EXTENSION AUX VARIETES

On voit que pour donner un sens aux énoncés des théorèmes 1 et 2, il faut s'être donné sur la variété V :

- Une métrique riemannienne, qui permettra de définir $d\langle X,X\rangle_t$,
- Une application Γ, donnant en tout point la « partie d'ordre 1 » d'un opérateur d'ordre 2. Une telle application est appelée, en géométrie différentielle, une connexion (plus précisément, une connexion linéaire, sans torsion) et les coefficients Γ_{ij}^k sont appelés les symboles de Christoffel de la connexion (dans la carte (x^i)). On les supposera C^∞. La donnée de la connexion permet de définir les martingales locales (continues). La condition reste $\Gamma(d^2\tilde{X}_t)=0$.

[Le mot « locale » est souvent omis lorsqu'on travaille dans les variétés, car la notion « globale » correspondante n'a pas de sens].

Soulignons que la validité des théorèmes 1 et 2 n'exige aucune relation entre la donnée de la métrique riemannienne et celle de la connexion Γ. Les théorèmes 1 et 2 ne sont donc pas des résultats fins, mais ils montrent où se trouvent les vrais problèmes : ceux-ci consistent à exclure l'ambiguïté du théorème 1, i.e. à trouver des critères pour que $\langle X,X\rangle_\infty<\infty$ entraîne $X_\infty \in V$.

II. LE THEOREME DE DARLING (TH. 1)

Pour simplifier, nous allons continuer à raisonner sur $\mathbb{R}^n$ muni de sa carte curviligne (x^i). La seule différence, c'est que la structure riemannienne (g_{ij}) et la connexion (Γ_{jk}^i) sont quelconques - à coefficients C^∞. Pour le cas général, se reporter à l'article à paraître de Darling et Zheng.

Il nous faut d'abord rappeler un fait sur les semimartingales, et énoncer un petit lemme d'algèbre linéaire.

Les processus $\langle X^i,X^j\rangle_t$ sont à variation finie, continus, nuls en 0. Il existe un processus croissant scalaire A_t, continu, nul en 0, tel que tous les $d\langle X^i,X^j\rangle_t$ soient absolument continus par rapport à dA_t . Nous désignerons par λ_t^{ij} une densité prévisible $d\langle X^i,X^j\rangle_t/dA_t$. Il est bien connu que ces coefficients peuvent être choisis de telle sorte que les formes quadratiques $\Sigma_{ij}\ \lambda_t^{ij}(\omega)\xi_i\xi_j$ soient positives pour tout (t,ω).

LEMME 1. Soient E un e.v. de dimension finie, E' son dual. On se donne

1) Une forme quadratique g sur E, strictement positive (une base (e_i) de E étant choisie, g s'écrit $\Sigma\ g_{ij}x^ix^j$) .

2) Une forme quadratique positive (mais peut être dégénérée) λ sur E'.

3) Une forme quadratique c sur E (sans autre restriction).

(on écrira $\lambda(\xi)=\Sigma_{ij}\ \lambda^{ij}\xi_i\xi_j$, et de même pour c). Alors on a

(6) $$|\ \Sigma_{ij}\ c_{ij}\lambda^{ij}| \leqq |\Sigma_{ij}\ g_{ij}\lambda^{ij}|.\|c\|_g$$ ($\|c\|_g$ est le sup de $|c(x)|$ sur la g-boule unité)

Si de plus c est positive, $\Sigma\ c_{ij}\lambda^{ij}$ est positif, et l'on peut remplacer $\|c\|_g$ par $Tr_g(c) = \Sigma_{ij}\ g^{ij}c_{ij}$ ((g^{ij}) est la forme duale de g).

DEMONSTRATION. Nous supposons bien connu le fait que des expressions comme $\Sigma_{ij}\ g_{ij}\lambda^{ij}$, $\Sigma_{ij}\ g^{ij}c_{ij}$ ne dépendent pas de la base utilisée. Alors nous choisissons une base de E, g-orthonormale , et dans laquelle λ est diagonale (à coefficients positifs). Donc

$$g(x) = \Sigma_i\ x^{i2} \quad , \quad g_{ij}=\delta_{ij}\ ,\ g^{ij}=\delta^{ij}$$

$$\lambda(\xi) = \Sigma_i\ \lambda^{ii}\xi_i^2\ ,\ \xi\in E'$$

et (6) provient des inégalités triviales

$$|\Sigma_i\ c_{ii}\lambda^{ii}| \leqq (\ \Sigma_i\ \lambda^{ii}).\sup_i|c_{ii}|\ ,\ \text{or}\ \Sigma_i\ \lambda^{ii}=\Sigma_i\ g_{ij}\lambda^{ij}$$
$$|c_{ii}|=|c(e_i,e_i)|\leqq\|c\|_g\ .$$

Pour la seconde inégalité, on majore plutôt par l'inégalité de Schwarz Sans supposer la positivité de c

$$|\Sigma_i\ c_{ii}\lambda^{ii}|\leqq (\Sigma_i\ c_{ii}^2)^{1/2}(\Sigma_i(\lambda^{ii})^2)^{1/2} \leqq (\Sigma_i\ |c_{ii}|)(\Sigma_i\ \lambda^{ii})$$

Si c est positive, les c_{ii} le sont, et $\Sigma_i\ c_{ii} = Tr_g(c)$.

DEMONSTRATION DU THEOREME 1 . Il suffit de démontrer que, pour toute fonction f sur V, C^∞ à support compact, $f(X_t)$ a une limite à l'infini sur l'ensemble où $\langle X,X\rangle_\infty < \infty$. Pour cela, on écrit la formule d'Ito sous la forme

$$f(X_t) = f(X_0) + \Sigma_i \int_0^t\ D_if(X_s)(dX_s^i + \tfrac{1}{2}\Sigma_{jk}\ \Gamma^i_{jk}(X_s)d\langle X^j,X^k\rangle_s)$$
$$+\tfrac{1}{2}\Sigma_{jk}\int_0^t\ d\langle X^j,X^k\rangle_s\ (\ D_{jk}f\ -\Gamma^i_{jk}D_if)(X_s)\ \ = f(X_0)+M_t+C_t\ ,$$

où (M_t) est une martingale locale, et (C_t) un processus à variation finie. On va montrer que $\langle M,M\rangle_\infty < \infty$, $\int_0^\infty |dC_s|<\infty$, sur l'ensemble où $\langle M,M\rangle_\infty < \infty$, et on aura terminé.

Premier terme : on a $d\langle M,M\rangle_t = \Sigma_{ij}\ D_if(X_s)D_jf(X_s)d\langle X^i,X^j\rangle_s$, et $d\langle X,X\rangle_s = \Sigma_{ij}\ g_{ij}(X_s)d\langle X^i,X^j\rangle_s$. Remplaçons $d\langle X^i,X^j\rangle_s$ par $\lambda_s^{ij}dA_s$, et appliquons le lemme avec $c_{ij}= D_if(x)D_jf(x)$: comme f est à support compact, $\|c\|_g$ est une fonction bornée de x, et nous pouvons écrire

$$d\langle M,M\rangle_t \leqq k\ d\langle X,X\rangle_t \quad ,\ \text{d'où la convergence.}$$

[En fait, la forme c est ici positive, et de rang 1 : il est plus intéressant d'appliquer l'autre majoration, qui fait apparaître la norme riemannienne de la forme df . Nous n'insistons pas].

Second terme. On applique l'inégalité du lemme 1 avec

$$c_{jk} = D_{jk}f(x) - \Sigma_i \Gamma^i_{jk}(x)D_if(x)$$

et $\|c\|_g$ est encore une fonction continue et bornée de x . Ici non plus, nous n'insistons pas sur le cas intéressant où cette forme est positive (f est alors dite convexe : cela ne peut s'appliquer à une fonction à support compact, mais permet de préciser l'étude à l'infini). Voir le travail définitif de Darling et Zheng.

III. LE THEOREME DE ZHENG (TH. 2)

La démonstration du théorème 2 fait intervenir des idées probabilistes intéressantes (assez proches, me semble-t-il, de certaines idées de Lenglart). Pour la clarté, nous présenterons le lemme de Zheng d'abord dans $\mathbb{R}$.

L'idée qui me semble la plus intéressante consiste à introduire la classe - invariante par changement de loi - des semimartingales continues réelles de la forme

(7) $$X=X_0+M+A \quad , \text{ telles que } dA_t \ll d\langle M,M\rangle_t$$

Il s'agit ici de la décomposition canonique : M est une martingale locale continue, nulle en 0, A est continu à variation finie, nul en 0. Nous pouvons alors écrire $A_t = \int_0^t H_s d\langle M,M\rangle_s$, avec H prévisible.

LEMME 2 . Soient λ,μ,K trois nombres >0. On a si $X_0=0$

(8) $$P\{X^*_\infty \leqq \lambda \ , \ M^*_\infty > \lambda+\mu \ , \ H^*_\infty \leqq K\} \leq K(\lambda+\mu)^2/\mu$$

[Note : le processus (H_s) n'étant défini qu'à un processus négligeable pour $d\langle M,M\rangle$ près, le sup H^* peut être interprété comme un sup ess : nous omettrons désormais ce genre de petites améliorations].

DEMONSTRATION. Soit U l'ensemble à gauche de (8), et soit S le temps d'arrêt $\inf\{t : |M_t|>\lambda+\mu\}$; sur U on a $S<\infty$. D'autre part, M^S est une martingale locale nulle en 0 et bornée, donc une vraie martingale, et l'on a

$$(\lambda+\mu)^2 \geqq E[M_S^2] = E[\langle M,M\rangle_S] \geqq E[I_U\langle M,M\rangle_S] \geqq \tfrac{1}{K}E[I_U\int_0^S |H_t|\, d\langle M,M\rangle_t]$$
$$\geqq \tfrac{1}{K}E[I_U|A_S|] \geqq \tfrac{1}{K}E[I_U(|M_S|-|X_S|) \geqq \tfrac{1}{K}E[\mu I_U]$$

car sur U on a $|M_S|=\lambda+\mu$ et $|X_S|\leqq\lambda$. [En fait, l'inégalité vaut aussi sans la condition $X_0=0$: pour le voir, se ramener par conditionnement à $X_0=x$, et faire entrer X_0 dans M : la seule modification est $(\lambda+\mu)^2 \geqq E[M_S^2 I_{\{S>0\}}] \geqq E[\langle M,M\rangle_S]$. Nous n'aurons pas besoin de cette extension.]

Le résultat suivant a vraiment un intérêt probabiliste !

LEMME 3. <u>Avec les notations</u> (7), $dA_t=H_t d\langle M,M\rangle_t$, <u>sur l'ensemble</u>

(9) $$C = \{ X_\infty \text{ existe et est fini, } \limsup_t |H_t|<\infty \}$$

<u>les limites</u> M_∞ <u>et</u> $\int_0^\infty |dA_s|$ <u>existent et sont finies p.s.</u>.

DEMONSTRATION. Il nous suffit de démontrer que, sur C, on a p.s. $M^*_\infty<\infty$. En effet, il est bien connu que sur l'ensemble $\{M^*_\infty<\infty\}$, M_∞ existe et est fini, ainsi que $\langle M,M\rangle_\infty$; comme $dA_t = H_t d\langle M,M\rangle_t$ et $|H_t|$ est borné sur C pour t assez grand, cela entraîne enfin $\int_0^\infty |dA_s|<\infty$ sur C.

Tout revient donc à montrer que, pour tout K>0, l'ensemble

$$D=\{X_\infty \text{ existe et est fini, } M^*_\infty=+\infty\ ,\ \limsup_t |H_t|<K/2\}$$

est négligeable. Supposons au contraire P(D)>0, et soit $\varepsilon<P(D)$. Soit $t_o\geqq 0$; posons

$$X'_t = X_{t_o+t}-X_{t_o}\ ,\ M'_t = M_{t_o+t}-M_{t_o}\ ,\ H'_t = H_{t_o+t}$$

Nous avons encore, pour t_o assez grand

$$P\{X'^*_\infty \leqq \varepsilon^2,\ M'^*_\infty = +\infty\ ,\ H'^*_\infty \leqq K\} > \varepsilon$$

et le lemme 2, appliqué à X' (nulle en 0) avec $\lambda=\mu=\varepsilon^2$ nous donne $\varepsilon \leqq K(2\varepsilon^2)^2/\varepsilon^2$ pour $0<\varepsilon<P(D)$, qui est absurde pour ε petit.

Passons aux semimartingales vectorielles. Nous supposons que X est à valeurs dans $\mathbb{R}^m$, nous écrivons chaque composante sous la forme $X^i= X^i_0+M^i+A^i$, et nous faisons l'hypothèse

(10) $$dA^i_t = \Sigma_{jk}\, H^i_{jkt} d\langle M^j,M^k\rangle_t \qquad (H^i_{jk}=H^i_{kj})$$

Pour appliquer le raisonnement du lemme 2, puis du lemme 3, nous devons interpréter $|M_t|$, $|X_t|$, $|A_t|$ comme les normes euclidiennes de M,X,A , $\langle M,M\rangle$ comme le processus croissant scalaire $\Sigma_i \langle M^i,M^i\rangle_t$, et il nous faut interpréter $|H|$ de manière à avoir

$$|dA_t| \leqq |H_t| d\langle M,M\rangle_t$$

Le lemme 1 nous permet de choisir de telles normes. En voici une, relativement grossière :

$$|H| = \Sigma_i \sup_{|x|\leqq 1} \Sigma_{jk}|H^i_{jk}x^i x^k|$$

Le lemme 3 subsiste alors sans modification : <u>sur l'ensemble</u> $\{\limsup |H_t|<\infty\}$, <u>l'existence de</u> X_∞ <u>dans</u> $\mathbb{R}^n$ <u>entraîne</u> $\langle M,M\rangle_\infty<\infty$.

APPLICATION AUX VARIETES

Nous présentons d'abord le raisonnement sous sa forme la plus simple, dans $\mathbb{R}^n$ avec ses coordonnées curvilignes $(x^1,\dots,x^n)$ comme au début : comme X est une martingale dans la variété, nous avons

$$dA^i_t = d\tilde{X}^i_t = -\Sigma_{jk}\ \Gamma^i_{jk}(X_t)d\langle X^i,X^j\rangle_t$$

Comme $\langle X^i,X^j\rangle_t = \langle M^i,M^j\rangle_t$, l'hypothèse (10) est satisfaite. D'autre part, si X_∞ existe dans V , la trajectoire $X_.(\omega)$ reste dans un compact $\mathcal{K}$ de V , et les fonctions $\Gamma^i_{jk}(x)$ sont bornées sur $\mathcal{K}$, donc la condition $|H|^*_\infty < \infty$ est automatiquement satisfaite. Par conséquent, on a $\Sigma_i\ \langle X^i,X^i\rangle_\infty < \infty$. Mais d'autre part, sur $\mathcal{K}$ on a une inégalité de la forme

$$|\Sigma_{ij}\ g_{ij}(x)\xi^i\xi^j\ | \leqq C\Sigma_i\ |\xi^i|^2$$

et par conséquent, pour une trajectoire qui reste dans $\mathcal{K}$, on peut affirmer que le processus croissant scalaire $\langle X,X\rangle_t$ est majoré par $C\Sigma_i\langle X^i,X^i\rangle_t$. Finalement, on voit que l'existence de X_∞ dans V entraîne $\langle X,X\rangle_\infty < \infty$.

Le cas d'une << vraie >> variété est un peu plus compliqué, et nous ne le traiterons pas : un moyen consiste à plonger V dans $\mathbb{R}^m$, avec m suffisamment grand, et à appliquer le raisonnement précédent dans $\mathbb{R}^m$ au lieu de $\mathbb{R}^n$: le principe est exactement le même, mais il faut un peu de travail pour écrire explicitement les H^i_{jk} avec ces m-n coordonnées << de trop >>.