JOURNAL OF APPROXIMATION THEORY 72, 329–345 (1993)

Multi-scale Analysis and Random Walks Boundaries

J. P. CONZE AND A. RAUGI

Université de Rennes I, Campus de Beaulieu, 35042 Rennes Cedex, France

Communicated by S. Riemenschneider

Received April 24, 1991; accepted in revised form December 6, 1991

Using martingales and random walks boundaries, we build approximation units in the framework of a multi-scale analysis close to the theory of wavelets. © 1993 Academic Press, Inc.

INTRODUCTION

The recently developed theory of wavelets and multi-scale analysis can be related to methods used in probability in the theory of μ-boundaries. In multi-scale analysis, one considers a space X and a group or a semi-group G acting on the space X and satisfying a contraction property. The main example is given by the affine group acting on $\mathbb{R}$. The wavelet bases are derived from a function which is a solution of a convolution equation for the action of G on X. In the same way, the μ-boundaries are associated to a group or a semi-group acting on X and to a measure ν on X solution of a convolution equation of the form $\mu * \nu = \nu$, where μ is a probability measure on G.

The purpose of this article is to present a method for the analysis of functions defined on the support of the invariant measure ν of a μ-boundary, in a way which is analogous to the analysis of functions on $\mathbb{R}$, or $\mathbb{R}^d$, in wavelet bases.

In a first part, we briefly recall the methods used in multi-scale analysis, then we introduce the main ideas of μ-boundary analysis. The second part is devoted to a more formal presentation of μ-boundary analysis, and to the proof of the results. The main tool is the martingale theory, which is known to be related to the Calderón–Zygmund methods in real analysis underlying the wavelet analysis.

A note announcing the results of this paper has appeared in [1].

0021-9045/93 $5.00

1. Wavelet Analysis and μ-Boundaries

In recent years, the wavelet analysis has been developed in two directions (cf. [3, 6]), namely, continuous time analysis and expansion in orthonormal bases of wavelets. We recall briefly these two methods in the case of the real line.

We denote by G the affine group $\{g=(a,b), a>0, b\in\mathbb{R}\}$ acting on $\mathbb{R}$ by $x\to g\cdot x=ax+b$.

1.1. *Wavelet Transforms*

In the first method, we consider a fixed test function $\psi\in L^2(\mathbb{R})$, satisfying the admissibility condition

$$\int_0^\infty \frac{|\hat{\psi}(-\lambda)|^2}{\lambda}\,d\lambda=\int_0^\infty \frac{|\hat{\psi}(\lambda)|^2}{\lambda}\,d\lambda<\infty,$$

where $\hat{\psi}$ denotes the Fourier transform of ψ. To every function f of $L^2(\mathbb{R})$, one can associate the function H_f defined on G, or equivalently on the half-plane $\{(a,b), a>0, b\in\mathbb{R}\}$, by

$$H_f(a,b)=\int_{\mathbb{R}} f(x)\,a^{-1/2}\psi\left(\frac{x-b}{a}\right)dx. \tag{1}$$

The admissibility condition satisfied by ψ implies, for some constant c_ψ, an isometry formula

$$\int |f(x)|^2\,dx=\frac{1}{c_\psi}\iint |H_f(a,b)|^2\,\frac{da}{a^2}\,db, \tag{2}$$

and an inversion formula which can be written formally,

$$f(x)=\frac{1}{c_\psi}\int H_f(a,b)\,a^{-1/2}\psi\left(\frac{x-b}{a}\right)\frac{da}{a^2}\,db.$$

1.2. *Orthogonal Wavelets*

In the orthogonal wavelets method, we avoid redundancy and use only the coefficients $H_f(2^{-j}, 2^{-j}k)$, $j,k\in\mathbb{Z}$, from formula (1). We select for ψ an orthogonal wavelet, i.e., a function such that the family $\{\psi_{j,k}(\cdot)\}=\{2^{j/2}\psi(2^j\cdot-k), j,k\in\mathbb{Z}\}$ forms an orthonormal basis of $L^2(\mathbb{R})$. The coefficients $H_f(2^{-j}, 2^{-j}k)$ are, in that case, the coefficients of the expansion of f in this orthonormal basis $\{\psi_{j,k}\}$.

The Haar basis give a simple example of a family $\{\psi_{j,k}\}$ which forms an orthonormal, but not regular, basis of $L^2(\mathbb{R})$. The construction of regular orthonormal wavelets, which is mainly due to recents works of Yves Meyer and Ingrid Daubechies (cf. [6]), can be described within the framework of multi-scale analysis introduced by Stéphane Mallat [5]. We shall briefly describe in the sequel this approach.

Let V_0 be a closed sub-space of $L^2(\mathbb{R})$ which is invariant under the action of the translations by integers, and such that

$$f(\cdot) \in V_0 \Rightarrow f(\cdot/2) \in V_0. \tag{3}$$

The sequence $(V_n = \{f(2^n \cdot), f \in V_0\}, n \in \mathbb{Z})$ forms an increasing sequence of closed sub-spaces of $L^2(\mathbb{R})$. If this sequence fulfills the condition

$$\bigcap_n V_n = \{0\} \qquad \text{and} \qquad \overline{\bigcup_n V_n} = L^2(\mathbb{R}), \tag{4}$$

one can use the sequence of sub-spaces $(V_n)_{n \in \mathbb{Z}}$ to make a multi-scale or multi-resolution analysis of L^2 $(\mathbb{R})$. The sequence (V_n) is to be compared to a filtration in martingale models.

Our aim is to construct a sub-space V_0 generated by the functions $\{\phi(\cdot - n), n \in \mathbb{Z}\}$, ϕ being in $L^2(\mathbb{R})$. More precisely, we try to construct a function ϕ such that the family $\{\phi(\cdot - n), n \in \mathbb{Z}\}$ forms a Riesz basis of the closed sub-space spanned in $L^2(\mathbb{R})$. The condition (3) can be written

$$\phi(x) = 2 \sum_k h_k \phi(2x - k), \tag{5}$$

where (h_k) is an array of coefficients such that $\sum |h_k|^2 < \infty$.

One recognizes a convolution equation for the action of affine group G on $\mathbb{R}$. We consider to this effect the group G and the discrete measure supported by G giving the mass h_k to $g_k = (\frac{1}{2}, k/2)$, for $k \in \mathbb{Z}$. Equation (5) can be seen as a convolution equation of the form $\mu * \nu = \nu$, where the measure ν has density ϕ with respect to the Lebesgue measure.

For some particular choices of the array h, it is possible to show the existence of a regular solution for Equation (5). In that way, we obtain a function ϕ whose translates span a space V_0 verifying the condition (3). Moreover, it is not difficult to show (4).

It still remains to construct a ψ such that the family $\{\psi_{j,k}\}$ provides an orthonormal basis of $L^2(\mathbb{R})$. If the functions $(\phi(\cdot - n), n \in \mathbb{Z})$ form not only a Riesz basis but an orthonormal basis of V_0, then ψ is given by the formula

$$\psi(x) = 2 \sum_n (-1)^n h_{1-n} \phi(2x - n).$$

It is easy to check that the translates of ψ by the integers form, in that case, an orthonormal basis of the orthogonal supplementary of V_0 in V_1.

The orthogonality property requires that the array (h_k) satisfies some algebraic conditions. In some situations, it could be interesting to relax the orthogonality requirement, and replace it by other properties, as it is the case, for instance, in the dyadic interpolation.

Another possibility consists in taking for the array (h_k) a probability vector ($h_k \geqslant 0$, $\sum_k h_k = 1$). Such a choice destroys orthogonality (except in the particular case of the Haar basis), but provides us with several advantages: it leads to positive approximation operators, it can be generalized to various geometric situations, and finally, it allows the use of tools of probability theory.

It is this point of view which led us to the analysis of μ-boundaries presented here. In the remainder of this section, we give a short sketch for the case of $\mathbb{R}$, before the general case and the proof of the results in the further sections.

1.3. *Analysis of a μ-Boundary, Case of the Real Line*

In this section we remain within the framework of the affine group G acting on $E = \mathbb{R}$. For an element g of G, we denote by $(a(g), b(g))$ the coefficients of g.

Let μ be a probability measure on G which satisfies the conditions

$$\int \log^+ a(g)\, \mu(dg) < +\infty, \int \log^+ |b(g)|\, \mu(dg) < +\infty,$$

and

$$\int \log a(g)\, \mu(dg) < 0.$$

Let $(Y_n)_{n \geqslant 1}$ be a sequence of independant random variables defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, with values in G and distribution μ. Let $(X_n = Y_1 \cdots Y_n)_{n \geqslant 1}$ be the associated random walk.

The assumptions imply, using the law of large number,

$$\lim_n\, [a(Y_1) \cdots a(Y_n)]^{1/n} = \exp\left(\int \log a(g)\, \mu(dg)\right) < 1,$$

and

$$\limsup_n |b(Y_n)|^{1/n} \leqslant 1.$$

For every $x \in \mathbb{R}$, it follows that the process

$$Y_1 \cdots Y_n \cdot x = a(Y_1) \cdots a(Y_n)\, x + \sum_{k=0}^{n-1} a(Y_1) \cdots a(Y_k)\, b(Y_{k+1}),$$

converges $\mathbb{P}$-a.s. to a random variable Z which does not depend on x.

From that, we deduce that the distribution ν of Z is the unique probability measure on $\mathbb{R}$ which satisfies the convolution equation $\mu * \nu = \nu$, and that the sequence $(X_n \cdot \nu)_{n \geqslant 1}$ converges to the Dirac measure δ_Z.

We will say that the pair (E, ν), i.e., the real line $\mathbb{R}$ endowed with the measure ν, is a *μ-boundary* of G.

Let us assume that, for μ-almost every $g \in G$, the measure $g \cdot \nu$ is absolutely continuous with respect to ν, a condition which is satisfied, for instance, when μ is a discrete measure. For each integer $n \geqslant 1$, we can define

$$T_n f(x) = \int_G \left\langle f, \frac{dg\nu}{d\nu} \right\rangle_\nu \frac{dg\nu}{d\nu}(x)\, \mu^n(dg) = \int_{\mathbb{R}} K_n(x, y)\, f(y)\, \nu(dy), \tag{6}$$

with

$$K_n(x, y) = \int_G \frac{dg\nu}{d\nu}(x) \frac{dg\nu}{d\nu}(y)\, \mu^n(dg).$$

The previous formula is analogous to a decomposition formula of wavelet type. The measure ν is a test measure, in the same way the function ψ in Section 1.1 is a test function in the multi-scale analysis. The integrals of f with respect to the measures $g\nu$ give the expansion coefficients of f.

The following result (proved for general μ boundaries of groups in Section 3) shows that the sequence $(T_n f(x))_{n \in \mathbb{N}}$ gives an approximation of f:

Let (E, ν) be a μ-boundary of G. Then, for every function f in $L^p(\nu)$, we have $\lim_n T_n f = f$, *the convergence being in the L^p norm whenever $p \geqslant 1$, and pointwise whenever $p > 1$.*

In this way, we have defined a "unit approximation" in $L^p(E, \nu)$. Eq. (6) can be rewritten, using the random walk $(X_n)_{n \geqslant 1}$,

$$T_n f(x) = \mathbb{E}\left[f(Z) \frac{dX_n \nu}{d\nu}(x) \right],$$

or, in approximate form,

$$T_n f(x) \sim \mathbb{E}\left[f(X_n \cdot x_0) \frac{dX_n \nu}{d\nu}(x) \right],$$

where x_0 is any starting point in E. This last equation is an interpolation formula which provides a reconstruction scheme for f from the points $\{f(X_n \cdot x_0)\}_{n > 0}$, the "information" that the random walk reads on the space (E, ν). The approximation of a function on (E, ν) given by the previous scheme fits the hierarchical structure of (E, ν) defined as a μ-boundary of G and therefore can be interpreted as a multi-scale method.

2. Definitions, Notations, Examples

We present now the general framework of μ-boundaries analysis. We will use classical results on conditional expectation and martingales, (see, for instance, [7]). For the theory of μ-boundaries, see H. Furstenberg [2].

2.1. Definition. In the following, we denote by G a topological group (or semi-group), and by μ a probability measure on the Borel sets of G. The law of G is denoted multiplicatively, and we assume that G has an identity element e.

A topological space E on which G acts in a continuous way is called a *G-space*. That means that there exists a continuous application $(g, x) \to g \cdot x$ from $G \times E$ to E such that $e \cdot x = x$ and $g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot x$, $\forall x \in E$, $\forall g_1, g_2 \in G$.

2.2. *Notation.* If ν is a probability measure on a G-space E, we denote by $\mu * \nu$ the image of the product measure $\mu \otimes \nu$ on $G \times E$ by the application $(g, x) \to g \cdot x$. In other words we have

$$\int_E f(x)\, \mu * \nu(dx) = \int_G \int_E f(g \cdot x)\, \mu(dg)\, \nu(dx).$$

Considering in particular G itself as a G-space, we find the usual convolution of measures. When μ is a Dirac measure δ_g on G, the measure $\delta_g * \nu$ will be simply denoted by $g \cdot \nu$.

2.3. Definition. We call *(G, μ)-space* a pair (E, ν), where E is a G-space and ν a μ-invariant probability measure (i.e., such that $\mu * \nu = \nu$).

2.4. Examples. (1) Let $G = \{(a, b) : a > 0, b \in \mathbb{R}\}$ be the affine group acting on $\mathbb{R}$ by: $g \cdot x = ax + b$, for $g = (a, b) \in G$ and $x \in \mathbb{R}$. For every non zero natural integer r, consider the probability measure μ_r on G defined by

$$\mu_r = \frac{1}{2^r} \sum_{k=0}^{r} C_r^k \delta_{(1/2, k/2)}.$$

The distribution ν_r of the sum of r independant random variables, uniformly distributed on $[0, 1]$, is the unique probability measure μ_r-invariant on $\mathbb{R}$. The measure ν_r has a density φ_r with respect to the Lebesgue measure, supported by $[0, r]$. For $r \geqslant 2$, φ_r is a function of class C^{r-2}, piecewise polynomial of degree $r - 1$.

(2) Let $D = \{z \in \mathbb{C} : |z| \leqslant 1\}$ be the unit disk of the complex plane. Let $G = \{(\rho, \alpha) \in \mathbb{C}^2 : |\rho| = 1, |\alpha| < 1\}$ be the group of homographic

transformations on the disk D: if $g=(\rho, \alpha)\in G$ and $z\in D$, $g\cdot z=\rho((z+\alpha)/(1+\bar{\alpha}z))$. Let $K=\{(\rho, 0) : |\rho|=1\}$ be the sub-group of rotations of G. The group G acts continuously on the unit circle $E=\{z\in\mathbb{C} : |z|=1\}$. If μ is a probability measure on G, left-invariant under the rotations (i.e., such that $k\cdot\mu=\mu$, $\forall k\in K$), then the Lebesgue measure on the unit circle E is the unique μ-invariant probability measure.

2.5. *Harmonic Functions and (G, μ)-Spaces*

In the following, we will assume that we have a (G, μ)-space (E, ν), such that E is locally compact with a countable basis.

A function H on G, with values in $\bar{\mathbb{R}}$, is said to be μ-harmonic if, for every $g\in G$, the integral $\int_G H(gh)\,\mu(dh)$ exists in $\bar{\mathbb{R}}$ and is equal to $H(g)$.

A function H on G, with values in $\bar{\mathbb{R}}$, is said to be μ-harmonic in the wide sense if, for μ-almost every $g\in G$, the integral $\int_G H(gh)\,\mu(dh)$ exists in $\bar{\mathbb{R}}$ and is equal to $H(g)$.

Let f be a Borel function on E, Let us define

$$H_f(g)=\int_E f(g\cdot x)\,\nu(dx) \qquad (g\in G).$$

The function H_f is μ-harmonic bounded on G if the function f is bounded. It is μ-harmonic on G, with values in $[0, +\infty]$, if f is positive. If f is ν-integrable, $H_f(g)$ is defined for μ-almost every $g\in G$. We have thus defined a function which is μ-harmonic in the wide sense on G. The application $f\to H_f$ is a contraction from $L^p(E, \nu)$ to $L^p(G, \mu)$, for every real p, $1\leqslant p\leqslant +\infty$.

2.6. Example. We come back to the second example from (2.4). Since ν is the Lebesgue measure of the unit circle, the function H_f is right-invariant under the rotations, i.e., $H_f(gk)=H_f(g)$, $\forall g\in G$, $\forall k\in K$. The origin is left invariant by the sub-group of rotations K in G; $K=\{g\in G : g\cdot 0=0\}$. Setting

$$h_f(g\cdot 0)=H_f(g) \qquad (g\in G),$$

we define a function h_f on the unit disk D. If $g\cdot 0$ has the polar decomposition $re^{i\theta}$, we have

$$\begin{aligned} h_f(g\cdot 0)&=\int_E f(g\cdot x)\,\nu(dx)\\ &=\int_E f(x)\,\frac{dg\nu}{d\nu}(x)\,\nu(dx)\\ &=\frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\,\frac{1-r^2}{1+r^2-2r\cos(\theta-t)}\,dt. \end{aligned}$$

One recognizes in the last equality, the Poisson kernel; therefore the function h_f is an harmonic function (in the classical sense) for the Laplacian. In this example, the transform $f \to H_f$ can be interpreted as the construction of the harmonic function (in the sense of the Laplacian) on D with f as a trace on the unit circle E.

3. Convergence Theorems

3.1. We consider the product space $\Omega = G^{\mathbb{N}^*}$, equipped with the σ-algebra $\mathscr{F}$ of its Borel sets and the product measure $\mathbb{P} = \bigotimes_{\mathbb{N}^*} \mu$. We denote by $(Y_k)_{k \geq 1}$ the sequence of coordinates on Ω and we set

$$X_0 = e,$$
$$X_n = Y_1 \cdots Y_n, \qquad \forall n \geq 1.$$

Let us denote by $\mathscr{F}_0$ the trivial σ-algebra and, for $n \geq 1$, by $\mathscr{F}_n$ the σ-algebra which is generated by the variables $\{Y_k : 1 \leq k \leq n\}$.

We use the martingale theory to prove the following result.

3.2. Proposition. (i) *For $\mathbb{P}$-almost every $\omega \in \Omega$, the sequence of probability measures $(X_n(\omega) \cdot \nu)_{n \geq 0}$ converges vaguely to a probability measure $P(\omega, \cdot)$ on E such that*

$$\nu(dx) = \int_\Omega P(\omega, dx)\, \mathbb{P}(d\omega).$$

(ii) *The transition operator, denoted by P, defined by*

$$Pf(\omega) = \int_E f(x)\, P(\omega, dx) \qquad (\omega \in \Omega),$$

is a contraction from $L^p(E, \nu)$ into $L^p(\Omega, \mathbb{P})$, for every $p \in [1, +\infty]$.

(iii) *For every f in $L^1(E, \nu)$, we have*

$$H_f(X_n) \overset{\mathbb{P}\text{-a.s.}}{=} \mathbb{E}_{\mathbb{P}}[Pf \mid \mathscr{F}_n].$$

(iv) *The process $\{H_f(X_n) : n \geq 0\}$ converges $\mathbb{P}$-a.s. and in $L^p(\Omega, \mathbb{P})$ norm to Pf, for $f \in L^p(E, \nu)$, for every $p \in [1, +\infty[$.*

Proof. From the invariance relation $\mu * \nu = \nu$, it follows that, for every Borel bounded function f, the process $\{H_f(X_n), n \geq 0\}$ is a bounded martingale. From martingale theory, this process converges therefore $\mathbb{P}$-a.s. and in the sense of $L^p(\Omega, \mathscr{F}, \mathbb{P})$ $(p \geq 1)$ to a random variable W_f which closes the martingale

$$H_f(X_n) = \mathbb{E}_{\mathbb{P}}[W_f \mid \mathscr{F}_n].$$

The space $C_0(E)$ is separable, E being assumed to be locally compact with a countable basis. Let $(f_p)_{p \geqslant 0}$ be a dense sequence in $C_0(E)$. Applying the previous results to each f_p, we obtain a measurable subset Ω_0 of Ω with $\mathbb{P}$-measure 1 such that, for every $\omega \in \Omega_0$ and every $p \geqslant 0$, the sequence $(H_{f_p}(X_n(\omega)) = X_n(\omega) \cdot \nu(f_p))_{n \geqslant 0}$ converges. The convergence of $(X_n(\omega) \cdot \nu(f))_{n \geqslant 0}$ for every $f \in C_0(E)$, and for $\omega \in \Omega_0$, follows. The sequence of probability measures $(X_n(\omega) \cdot \nu)_{n \geqslant 0}$ converges therefore vaguely to a positive measure $P(\omega, \cdot)$, which satisfies, a.s.,

$$H_f(X_n) = \mathbb{E}_{\mathbb{P}}\left[\int_E f(x)\, P(\cdot, dx) \middle| \mathscr{F}_n\right], \qquad \forall f \in C_0(E). \tag{7}$$

Taking expectations, we obtain

$$\int_E f(x)\, \nu(dx) = \mathbb{E}_{\mathbb{P}}\left[\int_E f(x)\, P(\cdot, dx)\right], \qquad \forall f \in C_0(E).$$

One deduces that $\nu(dx) = \int_\Omega P(\omega, dx)\, \mathbb{P}(d\omega)$, on one hand and, taking an increasing sequence of elements in $C_0(E)$ which converges to the constant 1 function, that, for $\mathbb{P}$-almost every $\omega \in \Omega$, $P(\omega, \cdot)$ is a probability measure on E, on the other hand.

The relation (7) can be extended to elements in $L^1(E, \nu)$, using a density argument. The last assertion (iv) follows from (iii) using standard results of martingale theory. ∎

3.3. Let us consider the product $\tilde{\Omega} = \Omega \times E$ provided with its Borel sets and with the probability measure $\tilde{\mathbb{P}}$ defined by

$$\tilde{\mathbb{P}}(d\omega, dx) = \mathbb{P}(d\omega)\, P(\omega, dx).$$

We call W (resp. U) the projection map from $\tilde{\Omega}$ to Ω (resp. E).

For every $f \in L^p(E, \nu)$ and $F \in L^q(\Omega, \mathbb{P})(1/p + 1/q = 1, 1 \leqslant p \leqslant +\infty)$, the equality

$$\mathbb{E}_{\tilde{\mathbb{P}}}[f \circ U \;\; F \circ W] = \mathbb{E}_{\mathbb{P}}[Pf\, F]$$

shows that

$$(Pf) \circ W = \mathbb{E}_{\tilde{\mathbb{P}}}[f \circ U \mid W].$$

In other words, P is a conditional distribution for U knowing W.

Let P^*F be the element of $L^q(E, \nu)$ defined by $(P^*F) \circ U = \mathbb{E}_{\tilde{\mathbb{P}}}[F \circ W \mid U]$. This defines a contraction P^* from $L^q(\Omega, \mathbb{P})$ into $L^q(E, \nu)$ which is characterized by the relations

$$\int_\Omega Pf(\omega)\, F(\omega)\, \mathbb{P}(d\omega) = \int_E f(x)\, P^*F(x)\, \nu(dx),$$

for every $f \in L^p(E, \nu)$ and $F \in L^q(\Omega, \mathbb{P})$.

The spaces $G^{\mathbb{N}^*}$ and E being Polish spaces, there exists a "transition operator" version of P^* (cf., for example, [7, Chap. 5.4]). This version gives a conditional distribution of U knowing W.

3.4. For every $n \geqslant 0$, we denote by T_n the contraction of $L^p(E, \nu)$ $(1 \leqslant p \leqslant +\infty)$, defined by

$$T_n = P^* \mathbb{E}_{\mathbb{P}}^{\mathscr{F}_n} P,$$

where $\mathbb{E}_{\mathbb{P}}^{\mathscr{F}_n}$ denotes the conditional expectation operator with respect to the σ-algebra $\mathscr{F}_n$. We have

$$T_n f = P^*[H_f(X_n(\cdot))].$$

From the martingale theory and the properties of the conditional expectation, if follows immediately that for every $f \in L^p(E, \nu)$, $1 \leqslant p < +\infty$, the sequence of functions $(T_n f)_{n \geqslant 0}$ converges in the sense of $L^p(E, \nu)$ to P^*Pf. Moreover this convergence holds ν-a.s. when

$$\mathbb{E}_{\mathbb{P}}[\sup_{n \geqslant 0} |H_f(X_n)|] < +\infty,$$

which turns out to be the case for $p > 1$, according to Doob's theorem.

3.5. We have shown that, for $f \in L^p(E, \nu)$, $1 \leqslant p < +\infty$, the sequence $(T_n f)_{n \geqslant 0}$ converges in $L^p(E, \nu)$ to the element P^*Pf of $L^p(E, \nu)$. An interesting case is when the operator P^*P is the identity.

If $U = Z \circ W$, $\mathbb{P}$-a.s., for some random variable Z, the operator P^*P is clearly the identity. Conversely, assume that P^*P is the identity. For every $f \in L^2(E, \nu)$, we have

$$\mathbb{E}_{\mathbb{P}}[(Pf)^2] = \langle Pf, Pf \rangle_{\mathbb{P}} = \langle f, P^*Pf \rangle_{\nu} = \langle f, f \rangle_{\nu} = \mathbb{E}_{\mathbb{P}}[P(f^2)].$$

There exists then a subset Ω_0 in Ω with $\mathbb{P}$-measure 1 such that, for $\omega \in \Omega_0$, we have

$$\forall f \in C_0(E), \qquad P(f^2)(\omega) = (Pf)^2 (\omega),$$

i.e., f is $P(\omega, \cdot)$-a.s. constant. It follows that $P(\omega, \cdot)$ is a point measure, for every ω in Ω_0.

Therefore, we have shown that the operator P^*P is the identity if and only if, for $\mathbb{P}$-almost every $\omega \in \Omega$, the measure $P(\omega, \cdot)$ is a point measure, of the form $\delta_{Z(\omega)}$, where Z is a random variable with values in E and distribution ν.

3.6. Definition [2]. A (G, μ)-space (E, ν) is called a *μ-boundary* of G if, for $\mathbb{P}$-almost every $\omega \in \Omega$, the sequence of probability measures $(X_n \nu)_{n \geqslant 0}$ converges vaguely to a Dirac measure $\delta_{Z(\omega)}$.

In the case of a μ-boundary, we have the relations

$$P^*P = I \qquad \text{and} \qquad PP^* = \mathbb{E}_{\mathbb{P}}^{Z},$$

where $\mathbb{E}_{\mathbb{P}}^{Z}$ denotes the conditional expectation projector with respect to Z.

Therefore we have shown the following result:

3.7. THEOREM. *Let (E, ν) be a μ-boundary. For every f in $L^p(E, \nu)$, $1 \leqslant p < +\infty$, the sequence $(T_n f)_{n \geqslant 0}$ converges in $L^p(E, \nu)$ to f and this convergence holds also ν-a.s. for $p > 1$.*

4. EXPRESSION OF $T_n f$, SERIES REPRESENTATION

4.1. HYPOTHESIS. *We assume now that for μ-almost every $g \in G$, and therefore for $\sum_{r \geqslant 1}(1/2^r)\mu^r$-almost every $g \in G$, the measure $g\nu$ is absolutely continuous with respect to ν, i.e.,*

$$g\nu(dx) = \frac{dg\nu}{d\nu}(x)\,\nu(dx),$$

where we have denoted by $dg\nu/d\nu$ the Radon–Nikodým–Lebesgue derivative of the measure $g\nu$ with respect to ν.

This hypothesis is clearly satisfied if the measure μ is discrete.

4.2. *Expression of T_n*

For $\varphi \in L^q(E, \nu)$ and $f \in L^p(E, \nu)$, we have

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}[P\varphi \cdot H_f(X_n)] &= \mathbb{E}_{\mathbb{P}}[H_f(X_n)\,\mathbb{E}_{\mathbb{P}}^{\mathscr{F}_n}[P\varphi]] \\
&= \mathbb{E}_{\mathbb{P}}[H_f(X_n)\,H_\varphi(X_n)] \\
&= \int_G H_f(g)\,H_\varphi(g)\,\mu^n(dg) \\
&= \int_G H_f(g)\left(\int_E \varphi(g\cdot x)\,\nu(dx)\right)\mu^n(dg) \\
&= \int_G H_f(g)\left(\int_E \varphi(u)\,\frac{dg\nu}{d\nu}(u)\,\nu(du)\right)\mu^n(dg) \\
&= \int_E \varphi(u)\left[\int_G H_f(g)\,\frac{dg\nu}{d\nu}(u)\,\mu^n(dg)\right]\nu(du).
\end{aligned}$$

It follows that

$$T_n f(x) = \int_G H_f(g) \frac{dg\nu}{d\nu}(x)\, \mu^n(dg) \qquad (x \in E).$$

From now on (E, ν) is assumed to be a μ-boundary of G for which the hypothesis (4.1) is satisfied.

4.3. *Expression of the "Innovation"*

For $f \in L^1(E, \nu)$, $T_n f$ can be written

$$T_n f(\cdot) = \mathbb{E}_{\mathbb{P}}\left[H_f(X_n) \frac{dX_n \nu}{d\nu}(\cdot)\right] = \int K_n(\cdot, y)\, f(y)\, \nu(dy),$$

where

$$K_n(x, y) = \mathbb{E}\left[\frac{dX_n \nu}{d\nu}(x) \frac{dX_n \nu}{d\nu}(y)\right].$$

We remark that the kernel $K_n(x, y)$ is not necessarily defined for $x = y$. Since $\mu * \nu = \nu$, for ν-almost every $x \in E$, the function $g \to (dg\nu/d\nu)(x)$ is μ-harmonic in the wide sense. It follows that, for ν-almost every $x \in E$, the process $(dX_n \nu/d\nu)(x))$ is a martingale with respect to the filtration $(\mathscr{F}_n)_{n \geqslant 0}$; from which we deduce the relations,

$$\begin{aligned}(T_{n+1} f - T_n f)(\cdot) &= \mathbb{E}_{\mathbb{P}}\left[[H_f(X_{n+1}) - H_f(X_n)] \frac{dX_{n+1}\nu}{d\nu}(\cdot)\right] \\ &= \mathbb{E}_{\mathbb{P}}\left[(H_f(X_{n+1}) - H_f(X_n)) \left(\frac{dX_{n+1}\nu}{d\nu} - \frac{dX_n \nu}{d\nu}\right)(\cdot)\right].\end{aligned}$$

4.4. For $g, y \in G$, define

$$\psi_{g,y}(x) = \frac{dgy\nu}{d\nu}(x) - \frac{dg\nu}{d\nu}(x) \qquad (x \in E).$$

The function ψ is in $L^1(E, \nu)$.

From the previous relations, it follows, for every function f in $L^1(E, \nu)$, the equality (in $L^1(E, \nu)$),

$$f(\cdot) = \langle f, 1\rangle_\nu + \sum_{n \geqslant 1} \int_G \int_G \langle f, \psi_{g,y}\rangle_\nu\, \psi_{g,y}(\cdot)\, \mu^n(dg)\, \mu(dy).$$

If $f \in L^p(E, \nu)$, with $p > 1$, this equality is valid in $L^p(E, \nu)$ and even ν-a.s.

4.5. EXAMPLE. Coming back to the first example in (2.4), we obtain the formulae

$$\varphi_r(x)\, T_n f(x) = 2^{2n} \sum_{k=0}^{(2^n-1)r} \left[\int_{\mathbb{R}} f(u)\, \varphi_r(2^n u - k)\, du \right] \varphi_r(2^n x - k)\, \rho_n(k),$$

$$\varphi_r(\cdot) f(\cdot) = \int_E f(u)\, \nu(du) + \sum_{n \geqslant 1} 2^{2n} \sum_{k=0}^{(2^n-1)r} \sum_{l=0}^{r} \rho_n(k)\, \rho_1(l)$$

$$\times \left[\int_{\mathbb{R}} f(u)\, \psi_{k,l}^{(n)}(u)\, du \right] \psi_{k,l}^{(n)}(\cdot),$$

with

$$\psi_{k,l}^{(n)}(x) = 2\varphi_r(2^{n+1}x - 2k - l) - \varphi_r(2^n x - k),$$

$$\rho_n(k) = \frac{1}{2^{nr}} \sum_{i=0}^{[k/2^n]} (-1)^i C_r^i C_{k-2^n i+r-1}^{r-1} \qquad 0 \leqslant k \leqslant (2^n - 1)\, r,$$

and, for $l - 1 \leqslant x < l$, $l = 1, ..., r$,

$$\varphi_r(x) = \frac{1}{(r-1)!} \sum_{i=0}^{l-1} (-1)^i C_r^i (x - i)^{r-1} \qquad \forall r \geqslant 2.$$

The previous formula giving $\rho_n(k)$ follows from

$$\rho_n(k) = \mathbb{P}[X_1 + 2X_2 + \cdots + 2^{n-1} X_n = k],$$

where $X_1, X_2, ..., X_n$ are independant random variables with binomial $B(r, \frac{1}{2})$ distribution.

4.6. *Remarks.* (1) Let us assume, for simplicity, that μ is discrete. Let μ^0 be the Dirac measure at the identity element e of G. For every integer $n \geqslant 1$, we call V_n the closed vector subspace of $L^1(E, \nu)$ generated by the functions $\{T_n f, f \in L^1(E, \nu)\}$. The sub-space V_0 contains only the constant functions.

The sub-space V_n is equal to the closed vector sub-space W_n generated by the functions $\{dg\nu/d\nu,\ g \in \operatorname{supp}(\mu^n)\}$. (It is clear that V_n is contained in W_n. On the other hand, if $\phi \in L^\infty(E, \nu)$ is orthogonal to V_n, we have

$$\langle T_n f, \phi \rangle_\nu = 0, \qquad \forall f \in L^1(E, \nu).$$

In particular, for $f = \phi$, we obtain $\mathbb{E}[H_\phi^2(X_n)] = 0$, i.e., $\langle \phi, dg\nu/d\nu \rangle = 0$, for μ^n-almost every $g \in G$. From which the inclusion of W_n in V_n.)

The equality

$$\sum_{\gamma \in G} \frac{dg\gamma\nu}{d\nu}(\cdot)\,\mu(\gamma) = \frac{dg\nu}{d\nu}(\cdot) \qquad (g \in G)$$

implies the inclusion $V_n \subset V_{n+1}$.

We have the properties (to be compared with those of a multi-scale analysis)

$$\bigcap_{n \geq 0} V_n = V_0$$

and, from what we have seen earlier,

$$\overline{\bigcup_n V_n} = L^1(E, \nu).$$

(2) There are numerous examples of μ-boundaries. We have already given examples when G is the affine group. If we consider now the case of $G = Gl(d, \mathbb{R})$ or $G = Sl(d, \mathbb{R})$ we can take for a G-space E the projective space P^{d-1} or more generally a space of flags. For a large class of probability measures μ, there exists a unique probability measure ν on E such that (E, ν) is a (G, μ) boundary. See [2, 4, 8].

(3) We have a "Plancherel's formula" (to be compared with the isometry formula (2))

$$\int_E f^2(x)\,\nu(dx) = \sum_{n > 0} \int (H_f(g))^2\,(\mu^{n+1} - \mu^n)(dg) + H_f^2(e).$$

5. From Local to Global

5.1. In the previous formulae, the pair (μ, ν) can be replaced by the pair $(\tau\mu\tau^{-1}, \tau\nu)$ for any element τ of G. It follows that, for $f \in L^p(E, \tau\nu)$, the sequence of functions,

$$T_n f^\tau(\tau^{-1} \cdot x) = \int_G H_f(\tau g) \frac{dg\nu}{d\nu}(\tau^{-1} \cdot x)\,\mu^n(dg),$$

converges in $L^p(E, \tau\nu)$, and even $\tau\nu$-a.s. if $p > 1$, to f. (We denote by f^τ the function $x \in E \to f(\tau \cdot x)$.)

This allows us the reconstruction of f on the support of $\tau\nu$ and not only on the support of ν.

5.2. Assume that there exists a discrete subgroup Γ of G such that $\sum_{\tau \in \Gamma} \tau v$ defines a positive Radon measure λ on the Borel sets of E, and define

$$S_n f(x) = \sum_{\tau \in \Gamma} \left(\int_G H_f(\tau g) \frac{dgv}{d\lambda} (\tau^{-1} x)\, \mu^n(dg) \right) \qquad (x \in E).$$

We have then the following result.

5.3. Theorem. *For every integer $n \geqslant 0$, the operator S_n is a contraction of the spaces $L^p(E, \lambda)$ $(p \geqslant 1)$. For every $f \in L^p(E, \lambda)$, the sequence of functions $(S_n f)_{n \geqslant 0}$ converges in $L^p(E, \lambda)$ to f and this convergence takes place even λ-a.s. if $p > 1$.*

Proof. For any subset Λ of Γ and any integer $n \geqslant 0$, define

$$S_n^\Lambda f(x) = \sum_{\tau \in \Lambda} \left(\int_G H_f(\tau g) \frac{dgv}{d\lambda} (\tau^{-1} x)\, \mu^n(dg) \right) \qquad (x \in E).$$

We note that

$$\left(\sum_{\tau_0 \in \Lambda} \frac{d\tau_0 v}{d\lambda}(x) \right)^{-1} \sum_{\tau_0 \in \Lambda} \frac{dgv}{d\lambda} (\tau_0^{-1} x)\, \delta_{\tau_0}(d\tau)\, \mu^n(dg),$$

defines, for λ-almost every $x \in E$, a probability measure on the Borel sets of $\Gamma \times G$. It follows, by the convexity inequality,

$$\| S_n^\Lambda f \|_{L^p(E,\lambda)}^p \leqslant \int_E \left(\sum_{\tau \in \Lambda} \frac{d\tau v}{d\lambda}(x) \right)^p S_n^\Lambda(|f|^p)(x)\, \lambda(dx).$$

As $\sum_{\tau \in \Lambda} d\tau v / d\lambda(\cdot) \leqslant 1$, λ-a.s., we obtain

$$\| S_n^\Lambda f \|_{L^p(E,\lambda)}^p \leqslant \sum_{\tau \in \Lambda} H_{|f|^p}(\tau) = \| f \|_{L^p(E, \sum_{\tau \in \Lambda} \tau v)}^p.$$

We deduce that the operator $S_n = S_n^\Gamma$ is a contraction in the $L^p(E, \lambda)$ spaces and that, for every f in $L^p(E, \lambda)$, the sequence $(S_n f)_{n \geqslant 0}$ converges in $L^p(E, \lambda)$ to f.

For $p > 1$, we show that this convergence holds even λ-a.s. We know that, for $\tau \in \Gamma$, the sequence of functions

$$\int_G H_f(\tau g) \frac{dgv}{d\lambda} (\tau^{-1} \cdot)\, \mu^n(dg) = \left(T_n f^\tau(\tau^{-1} \cdot) \frac{d\tau v}{d\lambda}(\cdot) \right)_{n \geqslant 0}$$

converges τv-a.s. to $f(\cdot)(d\tau v/d\lambda)(\cdot)$. From the Lebesgue dominated convergence theorem, it is sufficient to show that, for every positive element f of $L^p(E, \lambda)$,

$$\sum_{\tau \in \Gamma} \sup_n \left(T_n f^\tau(\tau^{-1} \cdot) \right) \frac{d\tau v}{d\lambda}(\cdot) < +\infty, \qquad \lambda\text{-a.s.}$$

We show in fact that this function is in $L_p(\lambda)$. Actually, we have

$$\left[\sum_{\tau\in\Gamma}\sup_n (T_n f^{\tau}(\tau^{-1}\cdot))\frac{d\tau\nu}{d\lambda}(\cdot)\right]^p \leqslant \sum_{\tau\in\Gamma}\left[\sup_n (T_n f^{\tau}(\tau^{-1}\cdot))^p \frac{d\tau\nu}{d\lambda}(\cdot)\right], \lambda\text{-a.s.},$$

because

$$\sum_{\tau\in\Gamma}\frac{d\tau\nu}{d\lambda} = 1, \lambda\text{-a.s.}$$

On the other hand, using Doob's inequality on p-integrable martingales, we have

$$\begin{aligned}
\int_E [\sup_n (T_n f^{\tau}(\tau^{-1}x))]^p \frac{d\tau\nu}{d\lambda}(x)\,\lambda(dx) &= \int_E \sup_n (T_n f^{\tau}(x))^p\,\nu(dx)\\
&= \mathbb{E}[\sup_n (\mathbb{E}[\mathbb{E}[f(\tau Z)\,|\,\mathscr{F}_n]\,|\,Z])^p]\\
&\leqslant \mathbb{E}[\sup_n \mathbb{E}[f(\tau Z)\,|\,\mathscr{F}_n]^p]\\
&\leqslant \mathbb{E}[\sup_n H_f(\tau X_n)^p]\\
&\leqslant \|\sup_n H_f(\tau X_n)\|^p_{L^p(\Omega,\mathbb{P})}\\
&\leqslant \left(\frac{p}{p-1}\right)^p \sup_n \|H_f(\tau X_n)\|^p_{L^p(\Omega,\mathbb{P})}\\
&\leqslant \left(\frac{p}{p-1}\right)^p \int_E f^p(x)\,\tau\nu(dx).
\end{aligned}$$

Consequently, the theorem follows. ∎

5.4. Example. In Example 1 of (2.4), we take for Γ the group of integral translations. We obtain for λ the Lebesgue measure on $\mathbb{R}$.

The approximation of a function f on $\mathbb{R}$ is given by

$$S_n f(x) = \sum_{m\in\mathbb{Z}} \gamma_{n,m}\left(\int_{\mathbb{R}} f(u)\,\varphi_r(2^n u - m)\,du\right)\varphi_r(2^n x - m),$$

with

$$\gamma_{n,m} = 2^{2n}\sum_{\{l:0\leqslant m-2^n l\leqslant(2^n-1)r\}} \rho_n(m-2^n l),$$

and the expansion of f is

$$f(x) = \sum_{m \in \mathbb{Z}} \left\{ \left(\int f(u)\, \varphi_r(u-m)\, du \right) \varphi_r(x-m) + \sum_{n \geq 1} \sum_{l=0}^{r} \gamma_{n,m} \rho_1(l) \left[\int_{\mathbb{R}} f(u)\, \psi_{m,l}^{(n)}(u)\, du \right] \psi_{m,l}^{(n)}(x) \right\}.$$

References

1. J.-P. Conze and A. Raugi, Fonctions μ-harmoniques et méthodes probabilistes d'approximation, *C.R. Acad. Sci. Paris Sér. I* **308** (1989), 185–188.
2. H. Furstenberg, Boundary theory and stochastic processes on homogenous spaces, *in* "Proceedings Sympos. Pure Math.," Vol. 26, pp. 193–229, Amer. Math. Soc., Providence, RI, 1972.
3. A. Grossmann, J. Morlet, and T. Paul, Integral transforms associated to square integrable representation, I, *J. Math. Phys.* **26** (1985), 2473–2479.
4. G. Letac and V. Seshadri, A characterisation of the generalized inverse Gaussian distribution by continuous fractions, *Zeit. Wahr.* **62** (1983), 485–489.
5. S. Mallat, Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$, *Trans. Amer. Math. Soc.* **315**, No. 1 (1989), 69–88.
6. Y. Meyer, "Ondelettes et Opérateurs," Vols. I, II, and III, Hermann, Paris, 1990.
7. J. Neveu, "Bases Mathématiques du Calcul des Probabilités," Masson, Paris, 1964.
8. A. Raugi, Fonctions harmoniques et théorèmes limites pour les marches aléatoires sur les groupes, *Bull. Soc. Math. France Mémoire* **54** (1977).