Stochastic Processes and their Applications 79 (1999) 117–134

stochastic processes and their applications

Kac's moment formula and the Feynman–Kac formula for additive functionals of a Markov process

P.J. Fitzsimmons, Jim Pitman *,1

Department of Statistics, University of California, 367 Evans Hall #3860, Berkeley, CA 94720-3860, USA

Abstract

Mark Kac introduced a method for calculating the distribution of the integral $A_v = \int_0^T v(X_t)\,dt$ for a function v of a Markov process $(X_t,\ t \geqslant 0)$ and a suitable random time T, which yields the Feynman–Kac formula for the moment-generating function of A_v. We review Kac's method, with emphasis on an aspect often overlooked. This is Kac's formula for moments of A_v, which may be stated as follows. For any random time T such that the killed process $(X_t,\ 0 \leqslant t < T)$ is Markov with substochastic semi-group $K_t(x, dy) = \boldsymbol{P}_x\ (X_t \in dy,\ T > t)$, any non-negative measurable function v, and any initial distribution λ, the nth moment of A_v is $\boldsymbol{P}_\lambda A_v^n = n!\lambda(GM_v)^n\mathbf{1}$ where $G = \int_0^\infty K_t\,dt$ is the Green's operator of the killed process, M_v is the operator of multiplication by v, and $\mathbf{1}$ is the function that is identically 1. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Occupation time; Local time; Resolvent; Killed process; Terminal time; Green's operator

1. Introduction

Kac (1949, 1951), and Darling and Kac (1957) introduced a method for calculating the distribution of the integral

$$A_v = \int_0^T v(X_t)\,dt \tag{1}$$

for a function v defined on the state space E of a Markov process $X = (X_t,\ t \geqslant 0)$, and a time T that may be fixed or random. Kac (1949, 1951) considered the case when X is a Brownian motion (BM), but his method leading to the Feynman–Kac formula in that setting has since been developed and applied much more generally. See Chung and Williams (1990); Durrett (1984); Karatzas and Shreve (1988); Simon (1979); Stroock (1993) for textbook treatments of the F-K formula for BM, Section III.19 of Rogers and Williams (1994) for a modern treatment of the F-K formula for a Feller–Dynkin process, and Section 5 of Kesten (1986) for a survey with further

* Corresponding author. E-mail: Pitman@stat.berkeley.edu.

[1] Research supported in part by N.S.F. Grants DMS94-04345 and DMS9703691.

PII: S0304-4149(98)00081-7

references. In this paper we review an aspect of Kac's method not mentioned in these treatments. This is his formula for moments of A_v for suitable T, first derived in Kac (1951) for BM on the line, then generalized in Darling and Kac (1957) to a Markov process with abstract state space. The reader is not assumed to be acquainted with the modern theory of Markov processes beyond what can be found, for example, in Chapter III of Rogers and Williams (1994).

To state the basic form of Kac's formula in some generality, let $(\boldsymbol{P}_x,\ x\in E)$ be the family of probability measures governing a Markov process (X_t) set up on a suitable probability space $(\Omega,\mathscr{F})$; $\boldsymbol{P}_x$ is the law of X under the initial condition $X_0=x$. We assume that there is a σ-algebra $\mathscr{E}$ on E such that (i) $x\mapsto P^x(F)$ is $\mathscr{E}$-measurable for each $F\in\mathscr{F}$, and (ii) $(t,\omega)\mapsto X_t(\omega)$ is a $\mathscr{B}\otimes\mathscr{F}/\mathscr{E}$-measurable mapping of $[0,\infty)\times\Omega$ into E, where $\mathscr{B}$ is the Borel σ-algebra on $[0,\infty)$. It is convenient to assume that $(\Omega,\mathscr{F})$ accommodates a random variable T_α $(\alpha>0)$ which (under $\boldsymbol{P}_x$ for all $x\in E$) is independent of X, and has the exponential distribution with parameter α. Other random times T involving extra randomization may also be assumed to be defined on the same basic setup.

Call T a *Markov killing time* of X if under each $\boldsymbol{P}_x$ the killed process $(X_t,\ 0\leqslant t<T)$ is Markovian with (sub-Markovian) semigroup $(K_t,\ t\geqslant 0)$:

$$K_t f(x)=\boldsymbol{P}_x[f(X_t)1(t<T)]. \tag{2}$$

In addition we assume that $K_t f$ is $\mathscr{E}$-measurable for all $t>0$ and all positive $\mathscr{E}$-measurable f. In formula (2) (and elsewhere in the paper), $1(B)$ is the indicator of the event B and $\boldsymbol{P}_x$ serves double duty as the expectation operator for the probability measure $\boldsymbol{P}_x$. Define the *Green's operator* or *potential kernel* G associated with T by

$$Gf(x)=\boldsymbol{P}_x\int_0^T f(X_t)\,\mathrm{d}t=\int_0^\infty K_t f(x)\,\mathrm{d}t \tag{3}$$

for non-negative $\mathscr{E}$-measurable f. For example, T_α is a Markov killing time with $K_t=\mathrm{e}^{-\alpha t}P_t$, where $(P_t,\ t\geqslant 0)$ is the semigroup of X, in which case $G=R_\alpha=\int_0^\infty \mathrm{e}^{-\alpha t}P_t\,\mathrm{d}t$ is the *resolvent* or *α-potential* operator associated with (P_t). Other Markov killing times are $\infty=\lim_{\alpha\to 0}T_\alpha$, and T the first entrance or last exit time of a suitable subset B of the state space of X. A finite fixed time T is typically not a Markov killing time unless X is set up as a space–time process, so T becomes a hitting time. Other Markov killing times can be constructed (i) by killing the process at state-dependent rate $k(X_t)$ for some *killing rate function* k defined on E, (ii) by killing according to a multiplicative functional, and (iii) by combinations of these kinds of operations. See Blumenthal and Getoor (1968). As shown by the example of last exit times, a Markov killing time of X is not necessarily a stopping time. See Meyer et al. (1972), Sharpe (1981) for further examples in this vein.

Kac's moment formula (Darling and Kac (1957), Kac (1951)). *Let T be a Markov killing time for X, let λ be an arbitrary initial distribution on E, and let v be a non-negative measurable function on E. Then the nth moment of $A_v=\int_0^T v(X_t)\,\mathrm{d}t$ under*

$\boldsymbol{P}_\lambda$ is given by

$$\boldsymbol{P}_\lambda A_v^n = n!\,\lambda G_v^n \mathbf{1} \quad (n=1,2,\dots) \tag{4}$$

where $G_v(x,\mathrm{d}y) = G(x,\mathrm{d}y)v(y)$, G is the potential kernel for the killed process as in (3), *and* $\mathbf{1}$ *stands for the function that is identically* 1.

In terms of operators, $G_v = GM_v$ where M_v is the operator of multiplication by v. For $n=1$, formula (4) just restates definition (3) of the Green's operator G. For $n=2$ the formula reads

$$\boldsymbol{P}_\lambda A_v^2 = 2\lambda G_v^2 \mathbf{1} = 2\int_E \lambda(\mathrm{d}x)\int_E G(x,\mathrm{d}y)v(y)\int_E G(y,\mathrm{d}z)v(z). \tag{5}$$

Note the special case $v=\mathbf{1}$ in Eq. (4): $A_v = T$, $G_v = G$, so Eq. (4) becomes

$$\boldsymbol{P}_\lambda T^n = n!\,\lambda G^n \mathbf{1}. \tag{6}$$

The first appearance of formula (4) seems to be (3.5) in Kac (1951). There $X_t = (B_t, t)$ is a space–time BM derived from a one-dimensional BM B, and T is a fixed time. Darling–Kac [1957, p. 445, line 4] give the Laplace transformed version of the same formula for B a two-dimensional BM, which amounts to the present formula (4) for $X=B$ and $T=T_\alpha$. Formula (4) for general X and v, and $T=T_\alpha$, is implicit in the discussion on p. 446 of Darling and Kac (1957) and is used there for an asymptotic calculation of moments which identifies the limit distribution of $\int_0^T v(X_t)\,\mathrm{d}t$ as $T\to\infty$ for a large class of Markov processes. See Athreya (1986), Bingham (1973) for more recent developments in this vein. To illustrate with three more examples from the literature, Exercise 4.11.10 of Itô–McKean (1965) is (6) for X a one-dimensional diffusion and T the first exit time from an interval; Nagylaki (1974) gives the more general formula (4) in the same setting; Propositions 8.6 and 8.7 of Iosifescu (1980) are Eq. (4) for X a Markov chain, $T=\infty$, and v either the indicator of the set of all transient states, or the indicator of a single transient state.

As noted by Kac (1951) in the Brownian setting, summing the moment formula (4) weighted by $1/n!$ yields the

Feynman–Kac formula (Feynman (1948), Kac (1949))**.** *For $v \geqslant 0$,*

$$\boldsymbol{P}_\lambda \exp(A_v) = \lambda \sum_{n=0}^{\infty} G_v^n \mathbf{1} = \lambda f_v \tag{7}$$

where f_v is the minimal positive solution f of

$$f = \mathbf{1} + G_v f. \tag{8}$$

Informally, we may write

$$f_v = (I - G_v)^{-1}\mathbf{1}. \tag{9}$$

The meaning of $(I-G_v)^{-1}$ has just been precisely defined for $v\geqslant 0$, but this expression also makes sense for signed v under appropriate conditions.

Note that replacing A_v in Eq. (7) by $A_{\theta v} = \theta A_v$ for $\theta > 0$ gives an expression for the $\boldsymbol{P}_\lambda$ moment generating function of A_v, which may however diverge for all $\theta > 0$. If the m.g.f. does converge for some $\theta > 0$, it of course determines the $\boldsymbol{P}_\lambda$ distribution of A_v. But even if not, Kac's moment formula still allows evaluation of whatever moments of A_v are finite.

Khas'minskii (1959) found Eqs. (4) and (7) for a general X and T the exit time of a domain. He also noted the following immediate consequence of Eq. (4) which has found numerous applications in the theory of Schrödinger semigroups (Aizenman and Simon, 1982; Simon, 1982). See also Berthier and Gaveau (1978), Carmona (1979) and Pinsky (1986) for various refinements and further references.

Khas'minskii's condition. *If $G_v\mathbf{1}$ is bounded then for all x the moment generating function $\boldsymbol{P}_x[\exp(\theta A_v)]$ converges for $\theta < 1/\|G_v\mathbf{1}\|_\infty$.*

If the infinitesimal generator $\mathscr{G}$ of the semigroup (P_t) is a differential operator (such as $\frac{1}{2}\Delta$ for Brownian motion), then integral equation (8) can be recast as a differential equation subject to suitable boundary conditions depending on the nature of T. For details in various settings see Durrett (1984), Karatzas and Shreve (1988), Stroock (1993) and Section 13.4 of Dynkin (1965). Ciesielski and Taylor (1962) used Eq. (7) to derive the distribution of A_v for X a BM in $\mathbb{R}^k$ for $k \geqslant 3$, $T = \infty$ and v the indicator function of a solid sphere in $\mathbb{R}^k$, in which case A_v represents the total time spent by B in the sphere. See also Rogers and Williams (1994) Section III.20 for a different treatment.

The rest of this paper is organized as follows. Kac's moment formula as stated above is proved in Section 2. Some variations and corollaries are presented in Section 3. In Section 4 we explain how these results relate to the more customary statement of the F-K formula that the semi-group of the process obtained by killing X at rate $v(x)$ has infinitesimal generator $\mathscr{G} - M_v$. In Section 5 the general results are specialized to the context of a Markov chain with finite state space, where the F-K formula can be understood with almost no calculation by direct probabilistic argument. In Section 6 we point out how the F-K formula for occupation times of Markov chains applies to local times of more general Markov processes. Such formulae were the basis of Ray's (1963) derivation of the Ray–Knight description of the local time field of a one-dimensional diffusion evaluated at a Markov killing time T, and of calculations by Williams (1967, 1969) for Markov chains.

2. Proof of Kac's moment formula

The proof is essentially just a formalization of Kac's (1951) original argument for space–time BM. Variations appear in the proofs of similar results in Dynkin (1984), Khas'minskii (1959), Pitman (1974, 1977). Let Y denote the killed process with state space $E \cup \partial$ defined by $Y_t = X_t$ if $t < T$ and $Y_t = \partial$ if $t \geqslant T$, where $\partial \in E$ is a cemetery state. In terms of Y, the killing time T is just the hitting time of ∂. It is therefore enough to prove the result for T the hitting time of a point in the state space E of X.

We assume, without any real loss of generality, that the sample space Ω is equipped with a family of shift operators $(\theta_t,\ t\geqslant 0)$, such that $X_s\circ\theta_t = X_{s+t}$ for all $s,t\geqslant 0$. Furthermore, we assume that there is a filtration $(\mathscr{F}_t,\ t\geqslant 0)$ on $(\Omega,\mathscr{F})$ to which X is adapted and with respect to which X has the simple Markov property:

$$\boldsymbol{P}_x(F\Phi(\theta_t)) = \boldsymbol{P}_x(FP_{X_t}\Phi) \quad (x\in E), \tag{10}$$

for all $t\geqslant 0$, all non-negative $\mathscr{F}_t$-measurable functions F, and all non-negative functions Φ on Ω that are measurable with respect to $\mathscr{H} := \sigma\{X_s,\ s\geqslant 0\}$.

The key property of a hitting time T of X is that it is a *terminal time* (Blumenthal and Getoor, 1968; Sharpe, 1988); that is, an $(\mathscr{F}_t)$-stopping time T with the property $T\circ\theta_t = T-t$ on the event $\{T>t\}$. The basic inductive step which allows Kac's moment formula (4) to be pushed from n to $n+1$ involves the following identity, which holds for an arbitrary $(\mathscr{F}_t)$-stopping time T. Let $G=G_T$ be the *pre-T occupation kernel* defined by

$$Gv(x) = \int_E G(x,\mathrm{d}y)v(y) = \boldsymbol{P}_x A_v \tag{11}$$

for an arbitrary non-negative measurable v. If λ is an initial distribution, then $\lambda G = \int_E \lambda(\mathrm{d}x)G(x,\cdot)$ is the measure

$$\lambda G(F) = \boldsymbol{P}_\lambda \int_0^T 1_F(X_s)\,\mathrm{d}s \quad (F\in\mathscr{E}), \tag{12}$$

which describes the $\boldsymbol{P}_\lambda$ expected amount of time X spends in various subsets F of E up to time T. Call λG the $\boldsymbol{P}_\lambda$ *pre-T occupation measure* for $(X_t,\ t\geqslant 0)$. In case T is a Markov killing time of X, G is the potential kernel derived from the killed process, as discussed in Section 1. But the above definition (11) of G makes sense, and the following identity is valid, for an arbitrary stopping time T:

Occupation measure identity (Khas'minskii, 1959, Pitman, 1974, 1977)**.** *For each initial distribution λ on E, each non-negative $\mathscr{H}$-measurable Φ, and each non-negative $\mathscr{E}$-measurable f,*

$$\boldsymbol{P}_\lambda \int_0^T f(X_t)\Phi(\theta_t)\,\mathrm{d}t = \int_E \lambda G(\mathrm{d}y)f(y)P_y\Phi. \tag{13}$$

The assumed measurability of $(t,\omega)\mapsto X_t(\omega)$ implies that $(t,\omega)\mapsto\Phi(\theta_t\omega)$ is $\mathscr{B}\otimes\mathscr{F}$-measurable, because Φ is $\mathscr{H}$-measurable. Thus the left side of Eq. (13) is well defined. Now Fubini's theorem shows that

$$\boldsymbol{P}_\lambda \int_0^T f(X_t)\Phi(\theta_t)\,\mathrm{d}t = \int_0^\infty \boldsymbol{P}_\lambda[f(X_t)\Phi(\theta_t)\,1(t<T)]\,\mathrm{d}t, \tag{14}$$

which by the Markov property (10) and the $\mathscr{F}_t$-measurability of $\{t<T\}$ is equal to

$$\int_0^\infty \boldsymbol{P}_\lambda[f(X_t)P_{X_t}(\Phi)\,1(t<T)]\,\mathrm{d}t = \boldsymbol{P}_\lambda \int_0^T f(X_t)P_{X_t}(\Phi)\,\mathrm{d}t. \tag{15}$$

Taken together, Eqs. (12), (14), and (15) yield Eq. (13).

Notice that the proof of Eq. (13) required no strong Markov property of X. So the occupation measure identity holds without any assumptions about the state space of X or path properties of X beyond the joint measurability of $X_t(\omega)$ as a function of t and ω.

Turning to the proof of Eq. (4), observe that

$$A_v^n = \left(\int_0^T v(X_t)\,\mathrm{d}t \right)^n = n!\, I_n \tag{16}$$

where

$$I_n = \int_{0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_n \leqslant T} v(X_{t_1})\,\mathrm{d}t_1\, v(X_{t_2})\,\mathrm{d}t_2 \cdots v(X_{t_n})\,\mathrm{d}t_n. \tag{17}$$

Because T is a terminal time, the obvious change of variables in Eq. (17) leads to

$$I_{n+1}(\omega) = \int_0^T v(X_t(\omega))\, I_n(\theta_t\omega)\,\mathrm{d}t. \tag{18}$$

Thus Eq. (4) follows by induction from the occupation measure identity (13).

3. Corollaries of Kac's moment formula

The basic notation and assumptions regarding X, T and G are as for Eq. (4).

3.1. Positive continuous additive functionals

To this point our study has focused on the random variable A_v, which is the value at time $t=T$ of the additive functional $A_v(t) = \int_0^t v(X_s)\,\mathrm{d}s$. In the abstract, a *positive continuous additive functional* (PCAF) is an $(\mathscr{F}_t)$-adapted family $A = (A(t),\ t \geqslant 0)$ of positive finite random variables satisfying the additivity condition

$$A(t+s) = A(t) + A(s)\theta_t \quad (s, t \geqslant 0). \tag{19}$$

We can define an operator G_A analogous to the operator G_v by the formula

$$G_A f(x) = \boldsymbol{P}_x \int_0^T f(X_t)\,\mathrm{d}A(t), \tag{20}$$

for positive $\mathscr{E}$-measurable f. Note that $G_A = G_v$ when $A(t) = \int_0^t v(X_s)\,\mathrm{d}s$. The validity of the analog of Eq. (13), namely

$$\boldsymbol{P}_x \int_0^T f(X_t)\Phi(\theta_t)\,\mathrm{d}A(t) = \int_E G_A(x, \mathrm{d}y) f(y)\, \boldsymbol{P}_y(\Phi), \tag{21}$$

requires a mild additional hypothesis. For instance, if E is a complete separable metric space (with Borel σ-algebra $\mathscr{E}$) and X has right-continuous sample paths, then Eq. (21) is valid with f and Φ as for Eq. (13). The proof of this assertion involves Ray–Knight compactification methods found in Sharpe (1988), and is well beyond the scope of this article. Assuming the validity of Eq. (21), we can repeat the earlier argument to show that *Kac's moment formula* (4) *and the F-K formula* (7) *hold for any PCAF A, provided the operator G_A is substituted for G_v*. Ray (1963), Section 2, used this version of the F-K formula for finite linear combinations of local times.

3.2. Signed additive functionals

For A_v derived from a function v that takes both positive and negative values, or, more generally, for a CAF $A=B-C$ that is the difference of PCAFs, it is easily seen that Kac's moment formula remains valid provided at each of the n successive integrations involved in computing $(G_A\mathbf{1})(x)$, $(G_A(G_A\mathbf{1}))(x),\dots$, the integral is absolutely convergent for each $x\in E$, as is the final integration with respect to λ. Such formulae are used, for example, in Marcus and Rosen (1992a,b,c, 1995), Rosen (1991) to study the asymptotics of differences $L_t^y - L_t^x$ of local times, as $y\to x$.

3.3. Covariances

For positive measurable functions v and w, applying Kac's formula to A_v, A_w and A_{v+w} and examining the result yields

$$\boldsymbol{P}_\lambda(A_vA_w)=\lambda(G_vG_w+G_wG_v)\mathbf{1} \tag{22}$$

where the terms in the decomposition can be understood using Eq. (16) and the occupation measure identity:

$$\lambda G_vG_w\mathbf{1}=\boldsymbol{P}_\lambda\int_0^T \mathrm{d}A_v(t)\int_t^T \mathrm{d}A_w(s). \tag{23}$$

These formulae are related to an energy form associated with v and w (Glover, 1981; Kemeny and Snell, 1961; Stroock, 1993). The $\boldsymbol{P}_\lambda$ covariance of A_v and A_w is

$$\lambda G_vG_w\mathbf{1}+\lambda G_wG_v\mathbf{1}-(\lambda G_v\mathbf{1})(\lambda G_w\mathbf{1}). \tag{24}$$

For each initial distribution λ, Eq. (24) is a symmetric bilinear non-negative definite function of pairs of non-negative functions chosen from $\{u\colon \lambda G_u^2\mathbf{1}<\infty\}$. Similar remarks apply to the more general CAFs considered in Section 3.2.

3.4. Additive functionals with jumps

In principle, the moments of an additive functional with jumps can be found by the same method. The formulae are not as simple however, because diagonal terms (which vanish in the continuous case) now appear in Eq. (17). The same thing happens in discrete time analogs of Kac's formulae discussed in the next subsection.

3.5. Discrete time analogs

To illustrate, in the discrete time version of Eq. (22) (see Kemeny and Snell (1961) p. 212) there is another term which must be subtracted due to double counting on the diagonal in Eq. (17): For T a Markov killing time of the discrete time Markov chain, $X_0, X_1, \ldots,$

$$\boldsymbol{P}_\lambda \left[\left(\sum_{n=0}^{T-1} v(X_n) \right) \left(\sum_{n=0}^{T-1} w(X_n) \right) \right] = \lambda(\tilde{G}_v \tilde{G}_w + \tilde{G}_w \tilde{G}_v - \tilde{G}_{vw})\mathbf{1} \tag{25}$$

where $\tilde{G}$ is the discrete time potential kernel $\tilde{G} = \sum_{n=0}^{\infty} K_1^n$ for K_1 as in Eq. (2). Because of diagonal terms like $\lambda \tilde{G}_{vw}\mathbf{1}$ above, the discrete time moment formulae do not iterate neatly except when v is the indicator of some subset B of E. In this case $A_v = N_B$ (say) is the number of hits on B before time T, and there is the following analog of Kac's formula for the rising factorial moments of N_B (Pitman, 1974; 1977)

$$\boldsymbol{P}_\lambda[N_B(N_B+1)\cdots(N_B+n-1)] = \lambda(\tilde{G}M_B)^n\mathbf{1} \tag{26}$$

where M_B is the operator of multiplication by the indicator of B. See Pitman (1974; 1977) and Section 3.2 of Iosifescu (1980) for further moment formulae in discrete time.

3.6. *The general product moment formula*

Returning to the setup of Kac's moment formula, or, more generally, the setting of Section 3.1, by iterated application of the occupation measure identity there is the following generalization of Eqs. (22) and (4) to a product of n additive functionals $A^{(i)}$, $1 \leqslant i \leqslant n$:

$$\boldsymbol{P}_\lambda \left(\prod_{i=1}^{n} A^{(i)} \right) = \lambda \left(\sum_{\pi} G_{A^{(\pi(1))}} G_{A^{(\pi(2))}} \cdots G_{A^{(\pi(n))}} \right) \mathbf{1} \tag{27}$$

where the sum extends over all permutations $\pi = (\pi(1), \ldots, \pi(n))$ of $\{1, \ldots, n\}$. Theorem 5.2 of Dynkin (1984) is this result in a slightly different framework. Dynkin assumes a symmetric potential density $g(x,y)$, but Eq. (27) applies nonetheless without such symmetry, and without assuming the existence of a potential density provided Eq. (21) is valid. Dynkin has shown how in the symmetric case the moment formula (27) underlies a far-reaching isomorphism between the distribution of functionals of the occupation field of a symmetric Markov process, and the distribution of the square of a Gaussian field with covariance derived from the positive definite kernel $g(x,y)$. See Dynkin (1983, 1984), Marcus and Rosen (1992a,b,c, 1996) for further developments, and Rogers and Williams (1994) I.27 for an elementary proof of Dynkin's isomorphism formula for a Markov chain, which is closely related to the discussion in Section 5 below.

Question. Is there any interesting connection between the occupation field of a Markov process that is not necessarily symmetric and the Gaussian process with covariance structure defined by the non-negative definite function (24)? It seems not, since it is not this positive kernel but the one derived more directly from $g(x,y)$ that works in the symmetric case.

3.7. Conditioning on X_{T-}

Assume now that E is a complete separable metric space, and that the sample paths of X are right-continuous with left limits; in particular Eq. (21) is valid. In this situation all of the previously displayed formulae have versions involving a conditioning on X_{T-}, as indicated in various settings by Kac (1951), Ray (1963) and Dynkin (1983, 1984). Such conditioning can be achieved in great generality using h-processes. To illustrate, the h-process version of Eq. (27), as formulated in Dynkin (1984), takes a simple form due to cancellation of the h-factors in the product of Green's kernels. See also Proposition 5.14 of Pitman (1977) for a discrete time example, and Aase (1977) for the formula obtained this way for the mean exit time of a diffusion on an interval conditioned to exit at a specified boundary point. The effect of conditioning on X_{T-} is simplest for the special class of Markov killing times introduced in the following definition:

Definition 1. Say a Markov killing time T is *killing with state-dependent rate* k, where k is a non-negative measurable function on E, if given $(X_s,\ 0\leqslant s\leqslant t)$ and the event $\{T>t\}$, the killing rate is $k(X_t)$:

$$\boldsymbol{P}_x(T\in(t,t+\,\mathrm{d}t)\,|\,X_s, 0\leqslant s\leqslant t, T>t)=k(X_t)\,\mathrm{d}t. \tag{28}$$

More formally, assuming T has been set up as a stopping time relative to a suitable enlargement $(\overline{\mathscr{F}_t})$ of the filtration $(\mathscr{F}_t)$, the assumption is that, under $\boldsymbol{P}_x$ for every $x\in E$, the process $(1_{(t\geqslant T)}-A_k(t\wedge T),\ t\geqslant 0)$ is an $(\overline{\mathscr{F}_t})$-martingale. Equivalently,

$$\boldsymbol{P}_x(Z_{\mathrm{T}}\,f(X_{T-}); T<\infty)=\boldsymbol{P}_x\int_0^T Z_t\,f(X_{t-})\,k(X_t)\,\mathrm{d}t \tag{29}$$

for every positive $(\overline{\mathscr{F}_t})$-predictable process Z and every positive $\mathscr{E}$-measurable function f.

The above definition makes sense, and the obvious analog of Eq. (29) is valid, if a general PCAF is substituted for $A_k(t)$. For example, the last exit time from a subset B of E will be "killing with rate $\mathrm{d}A(t)$" (for a suitable PCAF A) provided X is a strong Markov process with quasi-left-continuous sample paths (a "Hunt process"). In this context the PCAF A is naturally associated with the so-called "equilibrium distribution" on B; see Chung (1973), Getoor and Sharpe (1973a,b) and Glover (1982). Note that a first passage time into a set B will be of this form only if the first passage occurs at the time of a jump of X. In particular, a predictable Markov killing time T, such as the hitting time of a set for a process with continuous paths, will not be of this form.

Proposition 1. *For a Markov killing time T that is killing with rate function $k\geqslant 0$, and arbitrary non-negative measurable v and f,*

$$\boldsymbol{P}_\lambda[(A_v)^n f(X_{T-}); T<\infty]=n!\,\lambda G_v^n G_k f \quad (n=0,1,2,\ldots) \tag{30}$$

where $G_v=GM_v$ and $G_k=GM_k$.

Proof. For $n=0$ the result is the special case $Z\equiv 1$ in Eq. (29):

$$P_\lambda(f(X_{T-});T<\infty)=\lambda G_k f. \tag{31}$$

For general n we proceed by induction, as in the proof of Eq. (4). Thus, define

$$I_n(t)=\int_{0\leqslant t_1\leqslant t_2\leqslant\cdots\leqslant t_n\leqslant t} v(X_{t_1})\,\mathrm{d}t_1\, v(X_{t_2})\,\mathrm{d}t_2\cdots v(X_{t_n})\,\mathrm{d}t_n$$

and notice that $I_{n+1}(t)=\int_0^t I_n(t-s)\theta_s\, v(X_s)\,\mathrm{d}s$. Using $\varphi_n(x)$ as an abbreviation of $\boldsymbol{P}_x[(A_v)^n f(X_{T-});T<\infty]$, we have

$$\begin{aligned}
\varphi_{n+1}(x) &= (n+1)!\boldsymbol{P}_x\int_0^T I_{n+1}(t)f(X_{t-})\,k(X_t)\,\mathrm{d}t\\
&= (n+1)!\boldsymbol{P}_x\int_0^T\int_0^t [I_n(t-s)f(X_{t-s})k(X_{t-s})]\theta_s\, v(X_s)\,\mathrm{d}s\,\mathrm{d}t\\
&= (n+1)!\boldsymbol{P}_x\int_0^T\left[\int_0^T I_n(u)f(X_u)k(X_u)\,\mathrm{d}u\right]\theta_s\, v(X_s)\,\mathrm{d}s\\
&= (n+1)\boldsymbol{P}_x\int_0^T \varphi_n(X_s)\,v(X_s)\,\mathrm{d}s,
\end{aligned}$$

the final equality following from Eqs. (13) and (29). Thus, $\varphi_{n+1}=(n+1)G_v\varphi_n$, and Eq. (30) follows by induction on n.

Remark. Formula (30) holds also with either $A_v(t)$ or $A_k(t)$ or both replaced by general PCAFs. A formula could also be obtained with $f(X_{T-},X_T)$ instead of $f(X_{T-})$, assuming the existence of a Lévy system for the jumps of X and that T is a jump time. See e.g. Benveniste and Jacod (1973) and Pitman (1981).

Example. The special case of the above proposition for $T=T_\alpha$ an independent exponential time, when $k(x)=\alpha$ for all x, is already evident in Kac (1951). Then $G=R_\alpha=\int_0^\infty \mathrm{e}^{-\alpha t}P_t\,\mathrm{d}t$ is the resolvent operator of the semi-group of X. After cancelling the common factor of α on both sides, the result is as follows: for arbitrary $f\geqslant 0$,

$$\boldsymbol{P}_\lambda\int_0^\infty \mathrm{e}^{-\alpha t}\,(A_v)^n(t)f(X_t)\,\mathrm{d}t=n!\lambda(R_\alpha M_v)^n R_\alpha f. \tag{32}$$

While the existence of left limits was assumed in the previous proposition, it is easily shown that no such hypothesis is required for Eq. (32).

4. The Feynman–Kac formula

To recover more standard expressions of the F-K formula, as presented in Section III.19 of Rogers and Williams (1994), let (P_t^v) be the semigroup derived from (P_t) by killing X with state-dependent rate $v(x)$. So if T is the associated Markov killing time then

$$P_t^v f(x) = \boldsymbol{P}_x[f(X_t)1(T>t)] = \boldsymbol{P}_x[\exp(-A_v(t))f(X_t)]. \tag{33}$$

Summing formula (32) weighted by $(-1)^n/n!$ yields an expression for the resolvent of this semigroup:

$$\lambda R_\alpha^v f = \int_0^\infty \boldsymbol{P}_\lambda[\exp(-A_v(t))f(X_t)]\mathrm{e}^{-\alpha t}\,\mathrm{d}t = \lambda[I + R_\alpha M_v]^{-1} R_\alpha f \tag{34}$$

where $R_\alpha = \int_0^\infty \mathrm{e}^{-\alpha t} P_t\,\mathrm{d}t$ is the resolvent of the semi-group of X. Some mild regularity on f and v are required to make sense of the second equality in Eq. (34), but when rearranged as

$$R_\alpha^v + R_\alpha M_v R_\alpha^v = R_\alpha \tag{35}$$

the formula holds as an identity of bounded positive kernels for arbitrary $\alpha > 0$ and non-negative measurable v. This is formula (III.19.5) of Rogers and Williams (1994), which is one of the three forms of the F-K formula presented by Rogers-Williams. As they remark, and as shown by the above argument, Eq. (35) is a robust form of the F-K formula which is valid with no hypotheses on the underlying Markov process X beyond jointly-measurable paths.

To interpret the F-K formula as a statement relating the infinitesimal generators of (P_t) and (P_t^v), let us recall that the *weak infinitesimal generator* of (P_t) (say) is the operator $\mathscr{G}$ defined by

$$\mathscr{G}f(x) = \lim_{t\downarrow 0} \frac{P_t f(x) - f(x)}{t} \tag{36}$$

on the domain $D(\mathscr{G})$ comprising those functions f for which the pointwise limit indicated in Eq. (36) exists and $\sup_{t>0} t^{-1}\|P_t f - f\|_\infty < \infty$. See Dynkin (1956), where it is shown that for each $\alpha > 0$, the resolvent operator R_α is an injective mapping of $b\mathscr{E}$ (the class of bounded $\mathscr{E}$-measurable functions) onto $D(\mathscr{G})$, and $\mathscr{G}R_\alpha f = \alpha R_\alpha f - f$. Briefly, $R_\alpha = (\alpha - \mathscr{G})^{-1}$. Viewed in these terms, Eq. (35) amounts to the most common presentation of the F-K formula:

$$\textit{the killed semigroup } (P_t^v,\ t\geqslant 0) \textit{ has generator } \mathscr{G}^v = \mathscr{G} - M_v, \tag{37}$$

the domain $D(\mathscr{G}^v)$ consisting of those functions $u \in D(\mathscr{G})$ for which uv is a bounded function. In the context of symmetric Markov processes, Eq. (37) can be reformulated in terms of "Dirichlet forms"; see Section 6.1 of Fukushima et al. (1994).

A third form of the F-K formula noted in Rogers and Williams (1994) (III.19.7) is the following variant of Eq. (35):

$$R_\alpha^v + R_\alpha^v M_v R_\alpha = R_\alpha \tag{38}$$

which can be understood with almost no calculation due to the following:

Probabilistic interpretation of Eq. (38). Consider $T \wedge T_\alpha$, the minimum of T and an independent exponential time with rate α. From Eq. (38), the Green's operator for X killed at time $T \wedge T_\alpha$ is R_α^v. Since killing at time $T \wedge T_\alpha$ is the same as killing with rate function $v(x)+\alpha$, the $\boldsymbol{P}_\lambda$ distribution of $X_{T\wedge T_\alpha}$ on $\{T<T_\alpha\}$ is found by an easy variation of the "last exit" formula (31):

$$P_\lambda(X_{T\wedge T_\alpha}\in \mathrm{d}x, T_\alpha>T)=\lambda R^v_\alpha M_v(\mathrm{d}x). \tag{39}$$

Thus Eq. (38) comes from integrating with respect to $\boldsymbol{P}_\lambda$ the decomposition

$$A_f(T\wedge T_\alpha)+(A_f(T_\alpha)-A_f(T\wedge T_\alpha))1(T_\alpha>T)=A_f(T_\alpha) \tag{40}$$

for an arbitrary additive functional $A_f=\int_0^t f(X_s)\,\mathrm{d}s$.

Another probabilistic interpretation of Eq. (38). Multiplication of both sides of Eq. (38) by α yields an identity of Markov kernels which may be understood in another way. By the companion of Eq. (39) with $\{T_\alpha\leqslant T\}$ instead of $\{T_\alpha>T\}$,

$$\text{the } \boldsymbol{P}_\lambda \text{ distribution of } X_{T_\alpha} \text{ on } \{T_\alpha\leqslant T\} \text{ is } \alpha\lambda R^v_\alpha \tag{41}$$

and by Eq. (39) and the memoryless property of T_α,

$$\text{the } \boldsymbol{P}_\lambda \text{ distribution of } X_{T_\alpha} \text{ on } \{T_\alpha>T\} \text{ is } \alpha R^v_\alpha M_v R_\alpha. \tag{42}$$

Adding these two distributions we arrive at $\alpha\lambda R_\alpha$, which is the $\boldsymbol{P}_\lambda$ distribution of X_{T_α}.

Remark. For a constant function v, say $v(x)=\beta$ for all x, Eq. (38) reduces to the *resolvent identity*:

$$R_{\alpha+\beta}+\beta R_{\alpha+\beta}R_\alpha=R_\alpha \quad (\alpha,\beta>0). \tag{43}$$

So the above arguments give two simple probabilistic interpretations of this identity involving the minimum $T_{\alpha+\beta}\overset{\mathrm{d}}{=}T_\alpha\wedge T_\beta$ of two independent exponential variables T_α and T_β and analysis of $(X_t,\ 0\leqslant t\leqslant T_\alpha)$ according to whether $T_\beta<T_\alpha$ or $T_\beta\geqslant T_\alpha$. Since it is obvious that $R_{\alpha+\beta}R_\alpha=R_\alpha R_{\alpha+\beta}$, the resolvent identity also yields Eq. (35) in this special case. For a general v, comparison of Eqs. (35) and (38) establishes the identity of kernels

$$R_\alpha-R^v_\alpha=R^v_\alpha M_v R_\alpha=R_\alpha M_v R^v_\alpha \quad (\alpha>0) \tag{44}$$

where the first identity was interpreted probabilistically above.

There is also a probabilistic interpretation of the second identity in Eq. (44), involving the idea of *resurrection* of X after the killing time T, as considered in Fitzsimmons (1991) and Meyer (1975). Note that Eq. (44) is the Laplace transformed version of the following identity:

$$P_t-P^v_t=\int_0^t P^v_s M_v P_{t-s}\,\mathrm{d}s=\int_0^t P_s M_v P^v_{t-s}\,\mathrm{d}s, \tag{45}$$

and that if λ is an initial distribution, then

$$\lambda(P_t-P^v_t)f=\boldsymbol{P}_\lambda[f(X_t);\ T\leqslant t] \tag{46}$$

for any measurable $f\geqslant 0$. The first equality in Eq. (45) is evident from evaluation of this expectation by conditioning on (T,X_T). The second equality in Eq. (45) is obtained by the following construction. Assume for simplicity that v is bounded. Given

X, let $T=T_{(1)}$ where $T_{(1)}<T_{(2)}<\cdots$ are the points of a Poisson process on $(0,\infty)$ with intensity $v(X_t)\,dt$, $t>0$. The right-hand expression in Eq. (45) arises from evaluating expectation (46) by conditioning on $(T_{(N)}, X_{T_{(N)}})$, where $N=\max\{n: T_{(n)}\leqslant t\}$. Replacing the fixed time t by T_α gives a similar interpretation of the second equality in Eq. (44). From this perspective, the middle and right-hand expressions in Eqs. (44) and (45) are seen to be typical "first entrance" and "last exit" decompositions.

5. Application to Markov chains

Suppose now that X is a Markov chain with finite state space and T is a *finite* Markov killing time for X. By obvious reductions there is no loss of generality involved in the following:

Assumption. *The state space of X is $E\cup\{\partial\}$ where E is finite, ∂ is an absorbing state, $T=\inf\{t: X_t=\partial\}$ and $\boldsymbol{P}_x(0<T<\infty)=1$ for all $x\in E$.*

Let K_t denote the substochastic semi-group of X restricted to E, and view K_t and all other operators as matrices indexed by E, for example $K_t(x,y)=\boldsymbol{P}_x(X_t=y)$. Then $G=\int_0^\infty K_t\,dt$ is just

$$G=-Q^{-1}\quad\text{where } Q=\lim_{t\to 0} t^{-1}(K_t-I) \tag{47}$$

is the usual Q-matrix of the chain killed at time T. Recall that $G_v=GM_v$ where M_v is the operator of multiplication by v. It is clear that there exists $b>0$ such that for all v with $|v|\leqslant b$ the matrix $(I-G_v)$ is invertible. So $(I-G_v)^{-1}=(I+Q^{-1}M_v)^{-1}=(Q^{-1}(Q+M_v))^{-1}=(Q+M_v)^{-1}Q$ Eq. (47) and the F-K formula (7) can be restated as follows: *There exists $b>0$ such that for all v with $|v|\leqslant b$*

$$\boldsymbol{P}_\lambda \exp(A_v)=\lambda(I-G_v)^{-1}\mathbf{1}=\lambda(Q+M_v)^{-1}Q\mathbf{1}. \tag{48}$$

For $x\in E$ let $L_t^x=\int_0^t 1(X_s=x)\,ds$. Since $A_v=\sum_{x\in E} v(x)L_T^x$, formula Eq. (48) determines the joint moment generating function of the L_T^x, $x\in E$. The second expression in Eq. (48) for this m.g.f. appears as formula (4) in Kingman (1968), and again in Puri (1972). Kingman noted as a consequence that the joint m.g.f. is a ratio of two multilinear forms in $v(x)$, $x\in E$, and that the marginal distribution of each L_T^x is a mixture of a point mass at zero and an exponential distribution on $(0,\infty)$. Kingman raised the problem, which is apparently still open, of characterizing which joint distributions can appear as the joint distributions of such occupation times of a transient finite state chain. For some study of particular examples see Kent (1983) and Longford (1991).

Every killing time of a finite state chain is easily seen to be of the form assumed in Proposition 1 for the killing rate function $k(x)=(-Q\mathbf{1})(x)$ where Q is the Q-matrix of the killed chain. From Eq. (31)

$$\boldsymbol{P}_\lambda(X_{T-}=x)=(\lambda GM_k)(x) \tag{49}$$

so the assumption that $\boldsymbol{P}_x(T<\infty)=1$ for all x implies $GM_k\mathbf{1}=\mathbf{1}$. Formula (30) now yields expressions for the $\boldsymbol{P}_\lambda$ conditional moments of A_v given X_{T-}. This leads to the following sharper form of the F-K formula for chains. Formula (51) is a variant of Theorem 2.1 of Dynkin (1983). See also Sections I.27 of Rogers and Williams (1994) and IV.22 of Rogers and Williams (1987) for related presentations.

Proposition 2. *Let $k(x)=(-Q\mathbf{1})(x)$ where Q is the Q-matrix of X killed at time T. Then for v with $|v|\leqslant b$ for some $b>0$, and all f,*

$$\boldsymbol{P}_\lambda[\exp(-A_v)f(X_{T-})]=\lambda(I+GM_v)^{-1}GM_kf=\lambda(M_v-Q)^{-1}M_kf. \tag{50}$$

Furthermore, for all $x,y\in E$ such that $\boldsymbol{P}_x(X_{T-}=y)>0$,

$$\boldsymbol{P}_x[\exp(-A_v)\,|\,X_{T-}=y]=\frac{\tilde{G}_{xy}}{G_{xy}} \tag{51}$$

where $G=(-Q)^{-1}$ is the Green's matrix for X killed at time T, and $\tilde{G}=(M_v-Q)^{-1}$, which for $v\geqslant 0$ is the Green's matrix for X killed at time $\tilde{T}\leqslant T$ where $\tilde{T}$ is defined by additional killing with rate v.

Proof. Formula (50) results from summing Eq. (30) weighted by $-1/n!$, and using $QG=-I$. To derive Eq. (51) from Eq. (50), take λ to be a point mass at x and f to be the indicator of a single point y, and cancel the common factor of $k(y)$. □

Formula (51) for $v\geqslant 0$ can also be understood probabilistically by consideration of $\boldsymbol{P}_x(\tilde{T}=T,\ X_{T-}=y)$; that is, the probability that the chain starting at x ends with left limit y at time T having survived the additional killing at rate v. By conditioning on $\tilde{T}$,

$$\begin{aligned}\boldsymbol{P}_x(\tilde{T}=T,\ X_{T-}=y)&=\int_0^\infty \boldsymbol{P}_x(\tilde{T}\in \mathrm{d}t,\ X_{t-}=y,\tilde{T}=T)\,\mathrm{d}t\\&=\int_0^\infty \boldsymbol{P}_x(X_{t-}=y,\ \tilde{T}>t)k(y)\,\mathrm{d}t=\tilde{G}_{xy}\,k(y).\end{aligned}$$

On the other hand, by conditioning on X_{T-} the same probability equals

$$\boldsymbol{P}_x(X_{T-}=y)\boldsymbol{P}_x(\tilde{T}=T\,|\,X_{T-}=y)=G_{xy}\,k(y)\boldsymbol{P}_x\exp(-A_v)\,|\,X_{T-}=x).$$

Comparing the two results yields Eq. (51). Note the parallel between Eq. (51) and the more obvious formula

$$\boldsymbol{P}_x(\exp(-A_v)\,|\,T>t,\ X_t=y)=\tilde{K}(t,x,y)/K(t,x,y) \tag{52}$$

where $K(t,x,y)=\boldsymbol{P}_x(T>t,\ X_t=y)$ is the transition function of the chain killed at time T, and $\tilde{K}$ is the same for $\tilde{T}$ instead of T, namely the semigroup with Q-matrix $\tilde{Q}$ instead of Q, where $\tilde{Q}=Q-M_v$. Compare with formula (2.6.6) of Itô-McKean (1965) in Kac's original Brownian setting. A common generalization of these formulae in an abstract setting could be given using h-processes, but this is left to the reader.

6. Application to local times

Suppose as in Blumenthal and Getoor (1968) Eq. (3.41) that X with state space E admits a jointly measurable local time process $(L_t^x,\ t\geqslant 0,\ x\in E)$ relative to a reference measure dx on E, so that for $v\geqslant 0$

$$\int_0^t v(X_s)\,\mathrm{d}s = \int_E L_t^x\, v(x)\,\mathrm{d}x \quad \boldsymbol{P}_x \text{ a.s. for all } t\geqslant 0,\ x\in E. \tag{53}$$

For example, X could be a Markov chain with countable state space, with $L_t^x = \int_0^t 1(X_s = x)\,\mathrm{d}s$, or a one-dimensional diffusion (Rogers and Williams, 1987).

Fix a Markov killing time T for X. Then $g(x,y) = \boldsymbol{P}_x L_T^y$ serves as a density for the Green's kernel of X killed at time T. From Eq. (22), for all $a,x,y\in E$ there is the formula

$$\boldsymbol{P}_a(L_T^x L_T^y) = g(a,x)g(x,y) + g(a,y)g(y,x). \tag{54}$$

So for each a such that $P_a(T<\infty)=1$ the covariance

$$g(a,x)g(x,y) + g(a,y)g(y,x) - g(a,x)g(a,y) \tag{55}$$

is a symmetric non-negative definite function of $(x,y)\in E\times E$. Suppose now that a finite subset F of E is such that $\boldsymbol{P}_x(T<\infty)=1$, for all $x\in F$. The results of Section 3.6 show that for any initial distribution λ on F the $\boldsymbol{P}_\lambda$ joint distribution of $(L_T^y, y\in F)$ is determined by the values of the Green's function $g(x,y)$ for $x,y\in F$. In particular, all product moments $\boldsymbol{P}_\lambda[\prod_{y\in F}(L_T^y)^{n(y)}]$ for non-negative integer $n(y)$ have finite values which can be read from Eq. (27). And the $\boldsymbol{P}_\lambda$ joint moment-generating function of the $(L_T^y, y\in F)$ converges in a neighborhood of the origin and is given there by the formula

$$\boldsymbol{P}_\lambda \exp\left(\sum_{y\in F} v(y) L_T^y\right) = \lambda(I - G_v)^{-1}\mathbf{1} \tag{56}$$

where $G_v(x,y) = g(x,y)v(y),\ x,y\in F$. Put another way, Eq. (56) states that for v in a neighborhood of $\mathbf{0}$, the function

$$f_v(x) = \boldsymbol{P}_x \exp\left(\sum_{y\in F} v(y) L_T^y\right)$$

is the unique solution f of the system of equations

$$f(x) = 1 + \sum_{y\in F} g(x,y)v(y)f(y).$$

For X a one-dimensional diffusion, Ray (1963) Eq. (2.1) derived this system of equations for an h-process obtained conditioning on X_{T}, and went on to show that these equations imply the Ray–Knight descriptions for the distribution of local times of one-dimensional diffusions stopped at a Markov killing time. See also Sheppard (1985) who recovered most of Ray's results with the help of Dynkin's isomorphism theorem.

Note that the class of possible finite-dimensional distributions for $(L_T^y, y \in F)$ as above is precisely the class of joint distributions of total occupation times of various states in a finite state Markov chain. This can be understood by considering the time-changed Markov chain $(X_{\tau_l},\ l \geqslant 0)$ where $(\tau_l,\ l \geqslant 0)$ is the inverse of $(\sum_{y \in F} L_t^y,\ t \geqslant 0)$. Williams (1967, 1969) used a similar time change argument to derive variations of formula (54) for local time processes associated with both ordinary and fictitious states of a countable state Markov chain. See Theorem 6.1 of Williams (1969).

An altogether different application of the F-K formula to the local times of one-dimensional Lévy processes can be found in Bertoin (1995). This work concerns the law of the Hilbert transform of L_T^x with respect to x (for certain random T); it extends and simplifies (Fitzsimmons and Getoor, 1992), in which Kac's moment formula (4) is used. See also Jeanblanc et al. (1997) regarding connections between the F-K formula and path decompositions for one-dimensional BM.

Acknowledgements

We are grateful to Jay Rosen for his illuminating comments on the manuscript.

References

Aase, K., 1977. A conditional expectation formula for diffusion processes. J. Appl. Probab. 14, 626–629.

Aizenman, M., Simon, B., 1982. Brownian motion and Harnack inequality for Schrödinger operators. Comm. Pure Appl. Math. 35, 209–273.

Athreya, K.B., 1986. Darling and Kac revisited. Sankya: Indian J. Statist. 48, 255–266.

Benveniste, A., Jacod, J., 1973. Systèmes de Lévy des processus de Markov. Invent. Math. 21, 183–198.

Berthier, A., Gaveau, B., 1978. Critère de convergence des fonctionelles de Kac et applications en mechanique et en géometrie. J. Funct. Anal. 29, 416–424.

Bertoin, J., 1995. On the Hilbert transform of the local times of a Lévy process. Bull. Sci. Math. 119 (2), 147–156.

Bingham, N.H., 1973. Limit theorems in fluctuation theory. Adv. Appl. Probab. 5, 554–569.

Blumenthal, R.M., Getoor, R.K., 1968. Markov Processes and Potential Theory. Academic Press, New York.

Carmona, R., 1979. Regularity properties of Schrödinger and Dirichlet semigroups, J. Funct. Anal. 33 (3), 259–296.

Chung, K.L., 1973. A probabilistic approach to the equilibrium problem in potential theory. Ann. Inst. Fourier 23, 313–322.

Chung, K.L., Williams, R.J., 1990. Introduction to Stochastic Integration. 2nd ed. Birkhäuser, Basel.

Ciesielski, Z., Taylor, S.J., 1962. First passage times and sojourn density for Brownian motion in space and the exact Hausdorff measure of the sample path. Trans. Amer. Math. Soc. 103, 434–450.

Darling, D.A., Kac, M., 1957. On occupation times for Markoff processes. Trans. Amer. Math. Soc. 84, 444–458.

Durrett, R., 1984. Brownian Motion and Martingales in Analysis. Wadsworth, Belmont, CA.

Dynkin, E.B., 1956. Infinitesimal operators of Markov processes. Teor. Veroyatnost. i Primenen. 1, 38–60.

Dynkin, E.B., 1965. Markov Processes, I,II. Springer, Berlin, Heidelberg, New York, Toronto (translated from the Russian edition of 1962).

Dynkin, E.B., 1983. Local times and quantum fields. In: Huber, P., Rosenblatt, M. (Eds.), Seminar on Stochastic Processes. Birkhäuser, Basel, pp. 69–84.

Dynkin, E.B., 1983. Markov processes as a tool in field theory. J. Funct. Anal. 50, 167–187.

Dynkin, E.B., 1984. Gaussian and non-Gaussian random fields associated with Markov processes. J. Funct. Anal. 55, 344–376.

Dynkin, E.B., 1984. Polynomials of the occupation field and related random fields. J. Funct. Anal. 58, 20–52.

Feynman, R.J., 1948. Space-time approach to nonrelativistic quantum mechanics. Rev. Mod. Phys. 20, 367–387.

Fitzsimmons, P.J., 1991. Harmonic morphisms and the resurrection of Markov processes. In: Barlow, M., Bingham, N. (Eds.), Stochastic Analysis, London Mathematical Society Lecture Note Series, vol. 167. Cambridge University Press, Cambridge, pp. 71–90.

Fitzsimmons, P.J., Getoor, R.K., 1992. On the distribution of the Hilbert transform of the local time of a symmetric Lévy process. Ann. Probab. 20(3), 1484–1497.

Fukushima, M., Ōshima, Y., Takedam, M., 1994. Dirichlet forms and symmetric Markov processes, vol. 19 of de Gruyter Studies in Mathematics. Walter de Gruyter & Co., Berlin.

Getoor, R.K., Sharpe, M.J., 1973a. Last exit decompositions and distributions. Indiana Univ. Math. J. 23, 377–404.

Getoor, R.K., Sharpe, M.J., 1973b. Last exit times and additive functionals. Ann. Probab. 1, 550–569.

Glover, J., 1981. Energy and the maximum principle for nonsymmetric Hunt processes. Theory Probab. Appl. 26, 745–757.

Glover, J., 1982. Representing last exit potentials as potentials of measures. Z. Wahrsch. Verw. Gebiete. 61, 17–30.

Iosifescu, M., 1980. Finite Markov processes and their applications. Wiley, New York, Chichester.

Itô, K., McKean, H.P., Jr., 1965. Diffusion Processes and their Sample Paths. Springer, Berlin.

Jeanblanc, M., Pitman, J., Yor, M., 1997. The Feynman–Kac formula and decomposition of Brownian paths. Comput. Appl. Math. 16, 27–52.

Kac, M., 1949. On the distribution of certain Wiener functionals. Trans. Amer. Math. Soc. 65, 1–13.

Kac, M., 1951. On some connections between probability theory and differential and integral equations. In: Neyman, J. (Ed.), Proc. 2nd Berkeley Symp. Math. Stat. Prob., Univ. of California Press, Berkeley, CA, pp. 189–215.

Karatzas, I., Shreve, S., 1988. Brownian Motion and Stochastic Calculus. Springer, Berlin.

Kemeny, J.G., Snell, J.L., 1961. Potentials for denumerable Markov chains. J. Math. Anal. and Appl. 3, 196–260.

Kent, J.T., 1983. The appearance of a multivariate exponential distribution in sojurn times for birth–death and diffusion processes. In: Probability, Statistics and Analysis. London Mathematical Society Lecture Notes. Cambridge University Press, Cambridge, pp. 161–179.

Kesten, H., 1986. The influence of Mark Kac on probability theory. Ann. Probab. 14, 1103–1128.

Khas'minskii, R.Z., 1959. On positive solutions of the equation $au+vu=0$. Theory Probab. Appl. 4, 309–318.

Kingman, J.F.C., 1968. Markov transition probabilities. IV: Recurrence time distributions. Z. Wahrsch. Verw. Gebiete 11, 9–17.

Longford, N.T., 1991. Classes of multivariate exponential and multivariate geometric distributions derived from Markov processes. In: Sampson, A.R., Block, H.W., Savits, T.H. (Eds.), Topics in Statistical Dependence, vol. 16 of Lecture Notes-Monograph Series, I. M. S., pp. 359–369.

Marcus, M.B., Rosen, J., 1992a. Moduli of continuity of local times of strongly symmetric Markov processes via Gaussian processes. J. Theoret. Probab. 5, 791–825.

Marcus, M.B., Rosen, J., 1992b. p-variation of the local times of symmetric stable processes and of Gaussian processes with stationary increments. Ann. Probab. 20, 1685–1713.

Marcus, M.B., Rosen, J., 1992c. Sample path properties of the local times of strongly symmetric Markov processes via Gaussian processes. Ann. Probab. 20, 1603–1684.

Marcus, M.B., Rosen, J., 1995. Logarithmic averages for the local times of recurrent random walks and Lévy processes. Stoch. Proc. Appl. 59, 175–184.

Marcus, M.B., Rosen, J., 1996. Gaussian chaos and sample path properties of additive functionals of symmetric Markov processes. Ann. Probab. 24, 1130–1177.

Meyer, P.A., 1975. Renaissance, recollements, mélanges, ralentissement de processus de Markov. Ann. Inst. Fourier Grenoble 25, 465–497.

Meyer, P.A., Smythe, R.T., Walsh, J.B., 1972. Birth and death of Markov processes. Proc. 6th Berk. Symp. Math. Stat. Prob., vol. 3, pp. 295–305.

Nagylaki, T., 1974. The moments of stochastic integrals and the distribution of sojourn times Proc. Nat. Acad. Sci. USA 71, 746–749.

Pinsky, R., 1986. A spectral criterion for the finiteness or infiniteness of stopped Feynman–Kac functionals of diffusion processes. Ann. Probab. 14, 1180–1187.

Pitman, J., 1974. An identity for stopping times of a Markov process. In Studies in Probability and Statistics, Jerusalem Academic Press, pp. 41–57.

Pitman, J., 1977. Occupation measures for Markov chains. Adv. Appl. Probab. 9, 69–86.

Pitman, J., 1981. Lévy systems and path decompositions. Seminar on Stochastic Processes, 1981, Birkhäuser, Boston, pp. 79–110.
Puri, P.S., 1972. A method for studying the integral functionals of stochastic processes with applications. II: Sojourn time distributions for Markov chains. Z. Wahrsch. Verw. Gebiete, 23, 85–96.
Ray, D.B., 1963. Sojourn times of a diffusion process. Ill. J. Math. 7, 615–630.
Rogers, L.C.G., Williams, D., 1987. Diffusions, Markov Processes and Martingales, vol. II, Itô Calculus. Wiley, New York.
Rogers, L.C.G., Williams, D., 1994. Diffusions, Markov Processes and Martingales, 2nd ed., vol. I, Foundations, Wiley, New York.
Rosen, J., 1991. Second order limit laws for the local times of stable processes, Séminaire de Probabilités, XXV, Lecture Notes in Mathematics, vol. 1485. Springer, Berlin, pp. 407–424.
Sharpe, M.J., 1988. General theory of Markov processes. Academic Press, New York, London.
Sharpe, M.J., 1981. Killing times for Markov processes. Z. Wahrsch. Verw. Gebiete 58, 223–230.
Sheppard, P., 1985. On the Ray–Knight Markov property of local times. J. London Math. Soc. 31, 377–384.
Simon, B., 1979. Functional Integration and Quantum Physics, vol. 86 of Pure and Applied Mathematics. Academic Press, New York.
Simon, B., 1982. Schrödinger semigroups. Bull. A. M. S. 7, 447–526.
Stroock, D.W., 1993. Probability Theory, an Analytic View. Cambridge University Press, Cambridge.
Williams, D., 1967. On local time for Markov chains. Bull. A. M. S. 73, 432–433.
Williams, D., 1969. Fictitious states, coupled laws and local time. Z. Wahrsch. Verw. Gebiete 11, 288–310.