Proc. Nat. Acad. Sci. USA
Vol. 72, No. 3, pp. 780–783, March 1975

On a Variational Formula for the Principal Eigenvalue for Operators with Maximum Principle

(semigroups/Rayleigh–Ritz formula/Feynman–Kac formula)

MONROE D. DONSKER AND S. R. S. VARADHAN

Courant Institute of Mathematical Sciences, New York University, New York, N.Y. 10012

Communicated by Louis Nirenberg, December 11, 1974

ABSTRACT In this paper a variational formula is obtained for the principal eigenvalue for operators with maximum principle. This variational formula does not require the operators to be self-adjoint. But if they are self-adjoint this formula reduces to the classical Rayleigh-Ritz formula.

Let X be a compact metric space, and for $t \geq 0$ let T_t be a strongly continuous semigroup mapping $C(X) \to C(X)$ having the properties that if $f \geq 0$, then $T_t f \geq 0$ and $T_t 1 = 1$. Under these hypotheses the infinitesimal generator L of T_t will have domain $\mathfrak{D}$ dense in $C(X)$, and L will satisfy the maximum principle. Examples of operators L arising from such semigroups are

(*a*) X is a compact manifold and L is a second order elliptic operator with reasonable coefficients.

(*b*) $(Lf)(x) = \int_X [f(y) - f(x)]\pi(x,dy)$ where $\pi(x,dy)$ is a nonnegative measure on X for each $x \in X$ and is weakly continuous in x.

In fact, the most general L satisfying our assumptions is a limit of examples of type (*b*).

Let $V(x) \in C(X)$ and let λ_V be the principal eigenvalue of the operator $L+V$. If the operator L is self-adjoint with respect to a measure ν on X, then there is a classical variational formula (Rayleigh-Ritz) for λ_V, namely,

$$\lambda_V = \sup_{\substack{f \in L^2(\nu) \\ \|f\|_2 = 1}} \left[\int_X V(x) f^2(x)\nu(dx) + \langle Lf, f\rangle\right]. \qquad [1]$$

In this note we obtain a variational formula for λ_V for all L considered above whether L is self-adjoint or not. In the self-adjoint case the new variational formula reduces to [1].

Let $\mathfrak{D}^+$ denote the functions $u \in \mathfrak{D}$ that are positive and let $\mathfrak{M}$ be the space of all probability measures on X. For each $\mu \in \mathfrak{M}$ we define

$$I(\mu) = -\inf_{u \in \mathfrak{D}^+} \int_X \left(\frac{Lu}{u}\right)(x)\mu(dx). \qquad [2]$$

It is easy to see that $I(\mu)$ is nonnegative, lower semicontinuous, and convex.

The operator $L+V$ is the infinitesimal generator of a semigroup T_t^V given by a family of measures $p_V(t,x,dy)$, i.e.,

$$T_t^V f(x) = \int_X f(y) p_V(t,x,dy)$$

and we note

$$\|T_t^V\| = \sup_{x \in X} \int_X p_V(t,x,dy).$$

Moreover, if we let $\phi(t,V) = \log \|T_t^V\|$, we see that $\phi(t,V)$ is subadditive in t and therefore

$$\phi(V) = \lim_{t\to\infty} \frac{\phi(t,V)}{t} = \lim_{t\to\infty} \frac{1}{t} \log \sup_{x\in X} \int_X p_V(t,x,dy) = \lambda_V$$

exists. Because T_t is a positive semigroup, λ_V is in the spectrum of $L+V$ and is in fact the principal eigenvalue. We will prove

THEOREM. *The principal eigenvalue* λ_V of $L+V$ is given by*

$$\lambda_V = \sup_{\mu \in \mathfrak{M}} \left[\int_X V(x)\mu(dx) - I(\mu)\right]$$

where $I(\mu)$ is defined by [2]. *The theorem follows from Lemmas 1 and 2 below. Let us define*

$$\phi(V) = \lim_{t\to\infty} \frac{1}{t} \log \sup_{x\in X} \int_X p_V(t,x,dy),$$

$$\psi_1(V) = \sup_{\mu\in\mathfrak{M}} \left[\int_X V(x)\mu(dx) - I(\mu)\right],$$

$$\psi_2(V) = \inf_{u\in\mathfrak{D}^+} \sup_{x\in X} \left[V(x) + \left(\frac{Lu}{u}\right)(x)\right].$$

The theorem states that $\phi(V) = \psi_1(V)$. We will first prove in Lemma 1 that $\psi_1(V) \leq \phi(V) \leq \psi_2(V)$ and then in Lemma 2 that $\psi_2(V) = \psi_1(V)$.

* In general $\phi(V)$ is not an eigenvalue. However, it belongs to the spectrum of $L+V$ which is contained in $[z{:}\mathrm{Re}\, z \leq \phi(V)]$. We therefore call it the principal eigenvalue λ_V.

LEMMA 1.

$$\psi_1(V) \le \phi(V) \le \psi_2(V).$$

Proof: Let $\psi_2(V) = l$ and $\epsilon > 0$ be given. Then there exists $u \in \mathfrak{D}^+$ such that for all $x \in X$

$$V(x) + \left(\frac{Lu}{u}\right)(x) \le l+\epsilon.$$

For this u, let $v(t,x) = u(x)e^{(l+\epsilon)t}$ and we have

$$(L+V)v(t,x) = \left[\left(\frac{Lu}{u}\right)(x) + V(x)\right]v(t,x) \le (l+\epsilon)u(x)e^{(l+\epsilon)t}.$$

Thus, for all x and t,

$$\frac{\partial v(t,x)}{\partial t} \ge (L+V)v(t,x)$$

so that from the maximum principle we conclude

$$v(t,x) \ge \int_X v(0,y)p_V(t,x,dy) = \int_X u(y)p_V(t,x,dy).$$

Since X is compact and $u \in \mathfrak{D}^+$ we have $\inf_{x\in X} u(x) > 0$, and therefore for all x and t

$$\int_X p_V(t,x,dy) \le \frac{v(t,x)}{\inf_{x\in X} u(x)} \le \frac{\sup_{x\in X} u(x)}{\inf_{x\in X} u(x)} e^{(l+\epsilon)t}.$$

Thus,

$$\phi(V) = \lim_{t\to\infty} \frac{1}{t} \log \sup_{x\in X} \int_X p_V(t,x,dy) \le l+\epsilon,$$

and hence $\phi(V) \le \psi_2(V)$.

Now suppose $\phi(V) = h$. Let $u(t,x)$ be the solution for $t \ge 0$ of

$$\frac{\partial u}{\partial t} = Lu + Vu,$$

$$u(0,x) = 1,$$

i.e.,

$$u(t,x) = \int_X p_V(t,x,dy).$$

Since $\phi(V) = h$, corresponding to any $\epsilon > 0$ there is a t_0 such that $t \ge t_0$ implies

$$\sup_{x\in X} u(t,x) \le e^{(h+\epsilon)t}. \qquad [3]$$

Let $V^\epsilon(x) = V(x) - h - \epsilon$ and $u_\epsilon(t,x)$ be the solution of

$$\frac{\partial u_\epsilon}{\partial t} = Lu_\epsilon + V^\epsilon u_\epsilon,$$

$$u_\epsilon(0,x) = 1.$$

We see that $u_\epsilon(t,x) = u(t,x)e^{-(h+\epsilon)t}$ and from [3] we have for all $t \ge t_0$

$$\sup_{x\in X} u_\epsilon(t,x) \le 1.$$

In particular, $u_\epsilon(t_0,x) \le 1$ for all $x \in X$. If we consider then $u_\epsilon(t,x)$ for $x \in X$ and $0 \le t \le t_0$, we have

$$\frac{\partial u_\epsilon}{\partial t} = Lu_\epsilon + V^\epsilon u_\epsilon,$$

$$u_\epsilon(0,x) = 1,$$

$$u_\epsilon(t_0,x) \le 1,$$

and $u_\epsilon(t,x)$ is bounded below. Thus,

$$\frac{\partial \log u_\epsilon}{\partial t} = \frac{Lu_\epsilon}{u_\epsilon} + V^\epsilon,$$

so that for any $\mu \in \mathfrak{M}$ and $0 \le t \le t_0$

$$\int_X \frac{\partial \log u_\epsilon(t,x)}{\partial t}\mu(dx) = \int_X \left(\frac{Lu_\epsilon}{u_\epsilon}\right)(t,x)\mu(dx) + \int_X V^\epsilon(x)\mu(dx).$$

Therefore,

$$\int_0^{t_0} dt\left[\int_X \left(\frac{Lu_\epsilon}{u_\epsilon}\right)(t,x)\mu(dx) + \int_X V^\epsilon(x)\mu(dx)\right] = \int_X \log u_\epsilon(t_0,x)\mu(dx) - \int_X \log u_\epsilon(0,x)\mu(dx) = \int_X \log u_\epsilon(t_0,x)\mu(dx) \le 0.$$

This implies that there is some point $0 \le t' \le t_0$ such that

$$\int_X \left(\frac{Lu_\epsilon}{u_\epsilon}\right)(t',x)\mu(dx) + \int_X V^\epsilon(x)\mu(dx) \le 0,$$

and hence, letting $u(x) = u_\epsilon(t',x)$, we see that for each $\mu \in \mathfrak{M}$ there is a $u \in \mathfrak{D}^+$ such that

$$\int_X \left(\frac{Lu}{u}\right)(x)\mu(dx) + \int_X V^\epsilon(x)\mu(dx) \le 0.$$

This means

$$\sup_{\mu\in\mathfrak{M}}\left[\inf_{u\in\mathfrak{D}^+}\int_X \left(\frac{Lu}{u}\right)(x)\mu(dx) + \int_X V^\epsilon(x)\mu(dx)\right] \le 0,$$

and thus, from the definition of $I(\mu)$, we have

$$\sup_{\mu\in\mathfrak{M}}\left[-I(\mu) + \int_X V^\epsilon(x)\mu(dx)\right] \le 0,$$

i.e., $\psi_1(V^\epsilon) \le 0$. But obviously $\psi_1(V^\epsilon) = \psi_1(V) - h - \epsilon$, so that $\psi_1(V) \le h+\epsilon$ which implies $\psi_1(V) \le \phi(V)$.

LEMMA 2.

$$\psi_1(V) = \psi_2(V).$$

Proof: We must show $\psi_2(V) \le \psi_1(V)$. Let $\psi_1(V) = l$ and $\epsilon > 0$ be given. We then want to find $u_\epsilon \in \mathfrak{D}^+$ such that for all $x \in X$

$$\left(\frac{Lu_\epsilon}{u_\epsilon}\right)(x) + V(x) \leq l + \epsilon.$$

Since $\psi_1(V) = l$, we have from the definition of $I(\mu)$

$$\sup_{\mu \in \mathfrak{M}} \inf_{u \in \mathfrak{D}^+} \left[\int_X \left(\frac{Lu}{u}\right)(x)\mu(dx) + \int_X V(x)\mu(dx)\right] = l.$$

Thus, for any $\mu \in \mathfrak{M}$ and our given $\epsilon > 0$, there exists $u_{\epsilon,\mu} \in \mathfrak{D}^+$ such that

$$\int_X \left(\frac{Lu_{\epsilon,\mu}}{u_{\epsilon,\mu}}\right)(x)\mu(dx) + \int_X V(x)\mu(dx) \leq l + \frac{\epsilon}{8}.$$

Since the left side of this last inequality is continuous in the weak topology on $\mathfrak{M}$, there exists a neighborhood N_μ of μ such that if $\lambda \in N_\mu$,

$$\int_X \left(\frac{Lu_{\epsilon,\mu}}{u_{\epsilon,\mu}}\right)(x)\lambda(dx) + \int_X V(x)\lambda(dx) \leq l + \frac{\epsilon}{4}.$$

Thus $\bigcup_{\mu \in \mathfrak{M}} N_\mu$ is an open covering of the compact space $\mathfrak{M}$. Hence, there exists a finite covering $N_{\mu_1}, N_{\mu_2}, \ldots, N_{\mu_k}$. Let $u_i = u_{\epsilon,\mu_i}$ for $i = 1,2,\ldots,k$ so that we now have

$$\sup_{\mu \in \mathfrak{M}} \inf_{1 \leq i \leq k} \left[\int_X \frac{Lu_i}{u_i}(x)\mu(dx) + \int_X V(x)\mu(dx)\right] \leq l + \frac{\epsilon}{4}. \quad [4]$$

From the definition of the infinitesimal generator, $\frac{T_h u_i - u_i}{h} \to Lu_i$ uniformly in x as $h \to 0$. Since u_i has a positive lower bound, and

$$\int_X \left(\frac{T_h u_i - u_i}{hu_i}\right)(x)\mu(dx) \to \int_X \left(\frac{Lu_i}{u_i}\right)(x)\mu(dx)$$

uniformly in μ, we conclude from [4] that there exists an h_0 such that if $h \leq h_0$

$$\sup_{\mu \in \mathfrak{M}} \inf_{1 \leq i \leq k} \left[\int_X \left(\frac{T_h u_i - u_i}{hu_i}\right)(x)\mu(dx) + \int_X V(x)\mu(dx)\right] \leq l + \frac{\epsilon}{3}. \quad [5]$$

Since u_i is continuous and is bounded above and bounded below by a positive constant, we can write $u_i = e^{g_i}$ where $g_i \in C(X)$. Let G be the convex hull of $\{g_1, g_2, \ldots, g_k\}$. From [5] we obtain for $h \leq h_0$

$$\sup_{\mu \in \mathfrak{M}} \inf_{g \in G} \left[\int_X \left(\frac{T_h e^g - e^g}{he^g}\right)(x)\mu(dx) + \int_X V(x)\mu(dx)\right] \leq l + \frac{\epsilon}{3}. \quad [6]$$

Since T_h is bounded and nonnegative it is easy to see that $\frac{T_h e^g - e^g}{he^g}$ is a convex functional of g for each $x \in X$. Hence, we have from one of the mini-max theorems (see e.g., ref. 1, Thm. 3.4) that

$$\inf_{g \in G} \sup_{\mu \in \mathfrak{M}} \left[\int_X \left(\frac{T_h e^g - e^g}{he^g}\right)(x)\mu(dx) + \int_X V(x)\mu(dx)\right] \leq l + \frac{\epsilon}{3}, \quad [7]$$

which implies

$$\inf_{g \in G} \sup_{x \in X} \left[\left(\frac{T_h e^g - e^g}{he^g}\right)(x) + V(x)\right] \leq l + \frac{\epsilon}{3}.$$

Since G is compact the infimum is attained at some point of G, call it g_h, i.e., for all $x \in X$

$$\left[\left(\frac{T_h e^{g_h} - e^{g_h}}{he^{g_h}}\right)(x) + V(x)\right] \leq l + \frac{\epsilon}{3}.$$

Letting $u_h = e^{g_h}$ we have for all $x \in X$ and $h \leq h_0$

$$\left[\left(\frac{T_h u_h - u_h}{hu_h}\right)(x) + V(x)\right] \leq l + \frac{\epsilon}{3}. \quad [8]$$

Since the family $\{u_h\}$ for $h \leq h_0$ belongs to a compact set, we can assume without loss of generality that $u_h \to u_0$ as $h \to 0$ where $u_0 \in C^+(X)$ (continuous functions on X bounded below by a positive number) and $u_0 = e^g$ where $g \in G$. Without loss of generality we can assume $u_h \in \mathfrak{D}^+$ because $\mathfrak{D}^+$ is dense in $C^+(X)$. Hence, we get from [8]

$$\left[\left(\frac{T_h u_h - u_h}{hu_h}\right)(x) + V(x)\right] \leq l + \frac{\epsilon}{2}$$

for all $x \in X$, $h \leq h_0$, and where $u_h \to u_0$ as $h \to 0$ with $u_h \in \mathfrak{D}^+$. Since $u_h \in \mathfrak{D}^+$ we can rewrite this last as

$$\left[\frac{L\left(\frac{1}{h}\int_0^h T_s u_h ds\right)}{u_h}\right](x) + V(x) \leq l + \frac{\epsilon}{2} \quad [9]$$

for all $x \in X$ and $h \leq h_0$. Let

$$v_h = \frac{1}{h}\int_0^h T_s u_h ds \in \mathfrak{D}^+.$$

Since $u_h \to u_0$ and T_s is a strongly continuous semigroup, we have $v_h \to u_0$ as $h \to 0$. Now from [9] we get for all $x \in X$ and $h \leq h_0$

$$\left(\frac{Lv_h}{v_h u_h} \cdot v_h\right)(x) + V(x) \leq l + \frac{\epsilon}{2}$$

or

$$\left(\frac{Lv_h}{v_h}\right)(x) + V(x) \leq \left(l + \frac{\epsilon}{2}\right) + \left(\frac{u_h(x)}{v_h(x)} - 1\right) \times \left(l + \frac{\epsilon}{2}\right) + V(x)\left(1 - \frac{u_h(x)}{v_h(x)}\right)$$

$$\leq \left(l + \frac{\epsilon}{2}\right) + \left[\left(l + \frac{\epsilon}{2}\right) + \|V\|\right]\left(\left\|\frac{u_h}{v_h} - 1\right\|\right).$$

But $u_h \to u_0$ uniformly and, as noted earlier, $v_h \to u_0$ so that $\frac{u_h}{v_h} \to 1$ uniformly, i.e., for h sufficiently small

$$\left\|\frac{u_h}{v_h} - 1\right\| \left[\left(l + \frac{\epsilon}{2}\right) + \|V\|\right] \leq \frac{\epsilon}{2}.$$

Thus, there is an h_1 such that for all $x \in X$

$$\left(\frac{Lv_{h_1}}{v_{h_1}}\right)(x) + V(x) \leq l + \epsilon$$

which gives $\psi_2(V) \leq \psi_1(V)$. This completes the proof of *Lemma* 2 and the theorem.

Now in ref. 2 the authors proved (Section 4 of ref. 2) that if L is self-adjoint with respect to a reference measure ν on X and if $I(\mu) < \infty$ for some $\mu \in \mathfrak{M}$, then under mild assumptions $\mu \ll \nu$ and letting $g = \frac{d\mu}{d\nu}$ and and $f = g^{1/2}$ we have

$$I(\mu) = -\langle Lf,f\rangle.$$

This shows that in the case where L is self-adjoint the variational expression in the theorem reduces to the classical formula [1].

The relation between $\phi(V)$ and $I(\mu)$ is clearly that they are conjugate convex functionals. The relation $\phi(V) = \psi_2(V)$ has been noticed before. See, for instance, ref. 3.

The work reported in this paper was supported by the Office of Naval Research, Contract N00014-67-A-L467-0031 and the National Science Foundation, Grant NSF-GP-41985.

1. Sion, M. (1958) "On general minimax theorem," *Pac. J. Math.* **8,** 171–176.
2. Donsker, M. D. & Varadhan, S. R. S. (1975) "Asymptotic evaluation of certain Markov process expectations for large time—I," *Comm. Pure Appl. Math.*, in press.
3. Protter, M. H. & Weinberger, H. F. (1966) "On the spectrum of general second order operators," *Bull. Amer. Math. Soc.* **72,** 251–255.