June 1986 LIDS-P-1569

UPPER BOUNDS FOR SYMMETRIC MARKOV TRANSITION FUNCTIONS

by

E. A. Carlen*
Department of Mathematics,
Massachusettes Institute of Technology
Cambridge, MA 02139, U.S.A.

S. Kusuoka*,**
Department of Mathematics
Faculty of Science, University of Tokyo
Hongo, Tokyo 113, Japan

D. W. Stroock*,**
Department of Mathematics,
Massachusettes Institute of Technology
Cambridge, MA 02139, U.S.A.

(*): Research partially supported by N.S.F. grant DMS 8415211
(**): Research partially supported by A.R.O. grant DAAG29-84-K-0005

Introduction:

A large number of properties which are peculiar to symmetric Markov semigroups stem from the fact that such semigroups can be analyzed simultaneously by Hilbert space techniques as well as techniques coming from maximum principle considerations. The feature of symmetric Markov semigroups in which this fact is most dramatically manifested is the central role played by the Dirichlet form. In particular, the Dirichlet form is a remarkably powerful tool with which to compare symmetric Markov semigroups. The present paper consists of a number of examples which illustrate this point. What we will be showing is that there exist tight relationships between uniform decay estimates on the semigroup and certain Sobolev-like inequalities involving the Dirichlet form.

Because of their interest to both analysts and probabilists, such relationships have been the subject of a good deal of reserch. So far as we can tell, much of what has been done here-to-fore, and much of what we will be doing here, has its origins in the famous paper by J. Nash [N]. More recently, Nash's theme has been taken up by, among others, E. B. Davies [D] and N. Th. Varopoulos [V-1] and [V-2]; and, in a sense, much of what we do here is simply unify and extend some of the results of these authors. In particular, we have shown that many of their ideas apply to the general setting of symmetric Markov semigroups.

Before describing the content of the paper, we briefly set forth some terminology and notation. Careful definitions can be found in the main body of the paper.

Let E be a complete separable metric space, $\mathcal{B}$ its Borel field, and m a (σ-finite, positive) Borel measure on E. Let $\{\bar{P}_t: t > 0\}$ be a strongly continuous symmetric Markov semigroup on $L^2(m)$. The semigroup $\{\bar{P}_t: t > 0\}$ determines a quadratic form $\mathcal{E}$ on $L^2(m)$ through the definition

$$(0.1) \qquad \mathcal{E}(f,f) = \lim_{t\to 0} \frac{1}{t}((f,f) - (f,\bar{P}_t f)) .$$

(Here $(\cdot,\cdot)$ denotes the inner produuct in $L^2(m)$, and we are postponing all domain questions to the main body of the paper.) $\mathcal{E}(f,g)$ is then defined by polarization. $\mathcal{E}$ is called the Dirichlet form associated with the semigroup $\{\bar{P}_t: t > 0\}$. It is closed and non-negative, and therefore it determines a non-negative self adjoint operator $\bar{A}$ so that $\mathcal{E}(f,f) = (f,\bar{A}f)$.

One easily sees that $\bar{P}_t = e^{-t\bar{A}}$, and so the semigroup is in principle determined by its Dirichlet form. Our aim here is to show that at least as far as upper bounds are concerned, this is also true in practice; the Dirichlet form $\mathcal{E}$ provides a particularlly useful infintessimal description of the semigroup $\{\bar{P}_t: t > 0\}$.

Finally, to facilitate the description of our results, we assume in this introduction that the semigroup $\{\bar{P}_t: t > 0\}$ posseses a nice kernel $p(t,x,y)$.

In section 1) we carefully define the objects introduced above and spell out their relations to one another.

In section 2) we begin by characterizing the semigroups for which one has uniform estimates such as

$$(0.2) \qquad p(t,x,y) \le C/t^{\nu/2}$$

in terms of Dirichlet form inequalities of a type first considered by J. Nash [N]:

$$(0.3) \qquad \|f\|_2^{2+4/\nu} \le B\mathcal{E}(f,f)\|f\|_1^{4/\nu} \ ;$$

and indeed, our method of passing from (0.3) to (0.2) is taken directly from the work of Nash. (Our own contribution is that (0.2) and (0.3) are actually equivalent. Several applications here and elsewhere [K-S] turn on this equivalence.)

Once these basic facts have been established, the rest of section 2) is devoted to Dirichlet form characterizations -- again involving Nash type inequalities -- of cases when $p(t,x,y)$ decays differently for small times and large times. The characterizations again have a pleasantly simple form. (Theorem (2.9) and Corollary (2.12) are the main new results here.) Some applications of these results are given in section 2), others are described in section 5).

At the end of section 2), we discuss Varopoulos' result [V-2] characterizing (0.2) when $\nu > 2$ in terms of a Sobolev inequality

$$(0.4) \qquad \|f\|_{2\nu/(\nu-2)}^2 \le B'\mathcal{E}(f,f) \ .$$

Together the two characterizations yield the suprising result that (0.3) and (0.4) are equvalent for $\nu > 2$. However, because (0.2) and (0.3) are equivalent for all $\nu > 0$, and because (0.4) either does not make sense or is not correct for $\nu \le 2$, we find it more natural to characterize decay of $p(t,x,y)$, as we have throughout this paper, in terms of Nash type inequalities.

The uniform estimate (0.2) and all the estimates in section 2) are really only on-diagonal estimates for the kernel $p(t,x,y)$. Indeed, a simple application of the semigroup law and Schwarz's inequality yields $p(t,x,y) \le (p(t,x,x)p(t,y,y))^{1/2}$. In section 3) we take up an idea of Davies [D] to obtain off-diagonal decay estimates.

Davies' idea is to consider the semigroup $\{\bar{P}_t^{\psi}: t > 0\}$ defined by

$$(0.5) \qquad \bar{P}_t^{\psi} f(x) = e^{\psi}[\bar{P}_t(e^{-\psi} f)](x)$$

for some nice function ψ. Clearly this semigroup has a kernel $p^{\psi}(t,x,y)$ which is just $e^{\psi(x)} p(t,x,y) e^{-\psi(y)}$. In general, $\bar{P}_t^{\psi}$ will not be symmetric, or even contractive, on $L^2(m)$. Nonetheless, when $p(t,x,y)$ satisfies (0.2), one might still hope that for some number $N(\psi)$ and some number C independent of ψ,

$$(0.6) \qquad p^{\psi}(t,x,y) \le C t^{-\nu/2} e^{tN(\psi)} .$$

It would follow immediately that

$$(0.7) \qquad p(t,x,y) \le C t^{-\nu/2} e^{(\psi(y) - \psi(x) + tN(\psi))} ,$$

and one would then vary ψ to make the exponent as negative as possible.

Davies worked this strategy out for symmetric Markov semigroups coming from second order elliptic operators. In this case, the associated Dirichlet form $\mathcal{E}(f,f)$ is an integral whose integrand is a quadratic form in the grandient of f. Davies used the the classical Leibniz rule to, in effect, split the multiplication operators $e^{-\psi}$ and e^{ψ} off from $\bar{P}_t^{\psi}$ so that symmetric semigroup methods could be applied to $\{\bar{P}_t^{\psi}: t > 0\}$.

Here we develop Davies' strategy in a general setting, treating also the non-local case. (That is, the case when $\{\bar{P}_t: t > 0\}$ is not generated by a differential operator.) We are able to do this because, under very mild domain assumptions, a generic Dirichlet form $\mathcal{E}$ behaves as if $\mathcal{E}(f,f)$ were given by the integral of a quadratec from in ∇f. In particular, $\mathcal{E}$ satisffies a kind of Leibniz rule. (Of course, there is no "chain rule" in the non-local setting, and so it is somewhat suprising that there is a Leibniz rule, even in the absence of any differentiable structure.) We develop this Leibniz rule at the beginning of section 3); where we use ideas coming from Fukushima [F] and Bakry and Emery [B-E]. Even though a good deal of further input must be supplied to prove our generalization of Davies' result, it is this Leibniz rule which allows us to take apart the product structure of $\bar{P}_t^{\psi}$. Thus the principle underlying our generalization is really the same as the one which he used.

At the end of section 3) we give a brief example of the application of our result to a non-local case.

In section 4) we develop analogs of the results of section 2) in the discrete time case. In places this involves considerable modification of our earlier arguments. In fact, we do not know how to extend the results of section 3) to the discrete time case. Our direct treatment of the discrete time case appears to be both new and useful. In a recent paper [V-1], Varopoulos gave a very interesting application of continuous time decay estimates to determine the transcience or recurrance of a Markov chain. He was able to apply continuous time methods to this particular discrete time problem essentialy because it is a question about Green's functions. Other problems, however, seem to require a more direct approach.

In section 5) we give an assortment of applications and further illustrations of the results described above. For example, Theorem (5.20) discusses a discrete-time situation for which the results of section 4) appear to be essential.

1. Background Material:

Let E be a locally compact separable metric space, denote by $\mathscr{B} = \mathscr{B}_E$ the Borel field over E, and let m be a locally finite measure on E. Given a transition probability function $P(t,x,\cdot)$ on $(E,\mathscr{B})$, we say that $P(t,x,\cdot)$ is m-symmetric if, for each $t > 0$, the measure $m_t(dx\times dy) \equiv P(t,x,dy)m(dx)$ is symmetric on $(E\times E, \mathscr{B}\times\mathscr{B})$. We will always be assuming that our transition probability functions are continuous at 0 in the sense that $P(t,x,\cdot)$ tends weakly to δ_x as t decreases to 0. Note that if $\{P_t: t > 0\}$ denotes the semigroup on B(E) (the space of bounded $\mathscr{B}$-measurable functions on E into $\mathbb{R}$) associated with $P(t,x,\cdot)$ (i.e. $P_t f(x) = \int f(y)P(t,x,dy)$ for $t > 0$ and $f \in B(E)$), then for all $f \in B_o(E)$ (the elements of B(E) with compact support):

(1.1) $$\|P_t f\|_{L^p(m)} \leq \|f\|_{L^p(m)}, \quad t > 0 \text{ and } p \in [1,\infty].$$

Thus, for each $p \in [1,\infty)$, $\{P_t: t > 0\}$ determines a unique strongly continuous contraction semigroup $\{\overline{P}_t^p: t > 0\}$ on $L^p(m)$.

In particular, when $p = 2$ we write $\overline{P}_t$ in place of $\overline{P}_t^2$ and observe that $\{\overline{P}_t: t > 0\}$ is a strongly continuous semigroup of self-adjoint contractions. Then the spectral theorem provides a resolution of the identitiy $\{E_\lambda: \lambda \geq 0\}$ by orthognal projections such that

(1.2) $$\overline{P}_t = \int_{[0,\infty)} e^{-\lambda t} dE_\lambda, \quad t > 0.$$

Clearly, the generator of $\{\overline{P}_t: t > 0\}$ is $-\overline{A}$ where $\overline{A} \equiv \int_{[0,\infty)} \lambda dE_\lambda$.

Next define a quadratic form on $L^2(m)$ by

$$(1.3) \qquad \mathcal{E}(f,f) \equiv \int_{[0,\infty)} \lambda d(E_\lambda f,f) \ , \quad f \in L^2(m),$$

(We use (f,g) to denote the inner product of f and g in $L^2(m)$.) The domain $\mathfrak{D}(\mathcal{E})$ of $\mathcal{E}$ is defined to be the subspace of $L^2(m)$ where the integral in (1.3) is finite. Since $\frac{1}{t}(1 - e^{-\lambda t})$ increases to λ as t decreases to 0, another application of the spectral theorem shows that $\mathcal{E}_t(f,f)\uparrow\mathcal{E}(f,f)$ as $t\downarrow 0$, where

$$(1.4) \qquad \mathcal{E}_t(f,f) \equiv \frac{1}{2t}\int (f(y) - f(x))^2 m_t(dx\times dy) = \frac{1}{t}(f - \overline{P}_t f,f) \ ;$$

and that

$$(1.5) \qquad \mathfrak{D}(\mathcal{E}) = \mathfrak{D}(\bar{A}^{1/2}) = \{ f\in L^2(m) \mid \sup_{t>0}\mathcal{E}_t(f,f) < \infty \}$$

(Here $\mathfrak{D}(\bar{A}^{1/2})$ is the domain of the square root of $\bar{A}$.) The bilinear form $\mathcal{E}$ is called the Dirichlet form associated with the symmetric transition function $P(t,x,\cdot)$ on $(E,\mathcal{B},m)$.

It is clear from the (1.4) that $\mathcal{E}_t(|f|,|f|) \le \mathcal{E}_t(f,f)$. Taking the limit as t tends to zero, it is also clear that $\mathcal{E}$ posseses this same property. What is not so clear, and is in fact the key to the beautiful Beurling-Deny theory of symmetric Markov semigroups, is the remarkable fact that this last property of $\mathcal{E}$ essentially characterizes bilinear forms which arise in the way just described. For a complete exposition of the theory of Dirichlet forms, the reader is advised to consult M. Fukushima's monograph [F]. A more cursory treatment of the same subject is given in [L.D.] starting on page 146.

2. Nash-Type Inequalities:

Throughout this section, $P(t,x,\cdot)$ will be a symmetric transition probability function on $(E,\mathcal{B},m)$, and $\{P_t: t > 0\}$, $\{\bar{P}_t: t > 0\}$, $\{E_\lambda: \lambda \geq 0\}$, $\mathcal{E}$, and $\bar{A}$ will denote the associated objects introduced in section 1). Furthermore, we will use $\|f\|_p$ to denote the $L^p(m)$-norm of a function f and $\|K\|_{p\to q}$ to denote $\sup\{\|Kf\|_q: f \in B_o(E) \text{ with } \|f\|_p = 1\}$ for an operator K defined on $B_o(E)$.

As the first step in his famous article on the fundamental solution to heat flow equations, J. Nash proved that if a: $\mathbb{R}^N \longrightarrow \mathbb{R}^N \otimes \mathbb{R}^N$ is a bounded smooth symmetric matrix valued function which is bounded uniformly above and below by positive multiples of the identity, and if $p(t,x,y)$ denotes the non-negative fundamental solution to the heat equation $\partial_t u = \nabla\cdot(a\nabla)u$, then $p(t,x,y) \leq K/t^{N/2}$, $(t,x,y) \in (0,\infty)\times\mathbb{R}^N\times\mathbb{R}^N$, where K can be chosen to depend only on N and the lower bound on $a(\cdot)$.

The proof given below that (2.2) implies (2.3) is taken essentially directly from Nash's argument.

(2.1) Theorem: Let $\nu \in (0,\infty)$ and $\delta \in [0,\infty)$ be given. If

(2.2) $$\|f\|^{2+4/\nu} \leq A\left[\mathcal{E}(f,f) + \delta\|f\|_2^2\right]\|f\|_1^{4/\nu}, \quad f \in L^2(m),$$

for some $A \in (0,\infty)$, then there is a $B \in (0,\infty)$ which depends only on ν and A such that

(2.3) $$\|\bar{P}_t\|_{1\to\infty} \leq Be^{\delta t}/t^{\nu/2}, \quad t > 0.$$

Conversely, if (2.3) holds for some B, then (2.2) holds for an A depending only on B and ν.

Proof: We first note that it suffices to consider $f \in \mathfrak{D}(\bar{A})\cap L^{\infty}(m)\cap L^{1}(m)^{+}$ when proving the equivalence of (2.2) and (2.3). It suffices to consider non-negative functions because $\{\bar{P}_t: t > 0\}$ preserves non-negativity and $\mathcal{E}(|f|,|f|) \le \mathcal{E}(f,f)$. Furthermore, if $f \in L^1(m)^+$ and $f_n \equiv \bar{P}_{1/n}(f\wedge n)$, then $f_n \in \mathfrak{D}(\bar{A})\cap L^{\infty}(m)\cap L^{1}(m)^{+}$, $f_n \longrightarrow f$ in $L^1(m)$, and $\mathcal{E}(f_n,f_n) \le \mathcal{E}(f,f)$.

Assume that (2.2) holds, and let $f \in \mathfrak{D}(\bar{A})\cap L^{1}(m)^{+}$ with $\|f\|_1 = 1$ be given. Set $f_t = \bar{P}_t f$ and $u(t) = e^{-2\delta t}\|f_t\|_2^2$. Then, by (1.2) and (2.2): $-\frac{d}{dt}u(t) = 2e^{-2\delta t}\left[\mathcal{E}(f_t,f_t) + \delta\|f_t\|_2^2\right] \ge \frac{2}{A}u(t)^{1+2/\nu}$, where we have used the fact that $\|f_t\|_1 = \|f\|_1 = 1$. Hence, $\frac{d}{dt}\left[u(t)^{-2/\nu}\right] = -(2/\nu)u(t)^{-1-2/\nu}\frac{d}{dt}u(t) \ge 4/\nu A$; and so, $u(t) \le (4t/\nu A)^{-\nu/2}$. From this and the preceding paragraph, it is clear that $\|\bar{P}_t\|_{1\to 2} \le Ce^{\delta t}/t^{\nu/4}$, where C depends only on ν and A. Next, since $\bar{P}_t$ is symmetric, $\|\bar{P}_t\|_{2\to\infty} = \|\bar{P}_t\|_{1\to 2}$ by duality. Hence, by the semigroup property, $\|\bar{P}_t\|_{1\to\infty} \le \|\bar{P}_{t/2}\|_{1\to 2}^2 \le Be^{\delta t}/t^{\nu/2}$, where again B depends only on ν and A.

To prove the other assertion, assume (2.3). Choose $f \in \mathrm{Dom}(\bar{A})\cap L^1(m)^+$, and set $f_t = e^{-\delta t}\bar{P}_t f$. Then $\|f_t\|_\infty \le B\|f\|_1/t^{\nu/2}$ and $f_t = f - \int_0^t (\delta I + \bar{A})f_s ds$. Hence:

$$\begin{aligned} B\|f_t\|_1^2/t^{\nu/2} &\ge (f,f_t) = \|f\|_2^2 - \int_0^t (f,(\delta I + \bar{A})f_s)ds \\ &\ge \|f\|_2^2 - t\left[\mathcal{E}(f,f) + \delta\|f\|_2^2\right], \end{aligned}$$

where we have used (1.2) to conclude that $(f,(\delta I + \bar{A})f_s) \le \mathcal{E}(f,f) + \delta\|f\|_2^2$ for all $s \ge 0$. After segregating all the t-dependent terms on the right hand side and then minimizing with respect to $t \ge 0$, we conclude that (2.2) holds with an A depending only on ν

and B. Because of the remarks in the first paragraph, the proof is now complete.

Q.E.D.

The estimate (2.3), as it is written, ignores the fact that since $\|\bar{P}_t\|_{1\to 1} \le 1$ for all $t > 0$, $\|\bar{P}_t\|_{1\to\infty}$ is a decreasing function of t. However, it is clear that when $\delta > 0$, (2.3) is equivalent to

(2.3') $$\|\bar{P}_t\|_{1\to\infty} \le B'/(t\wedge 1)^{\nu/2}, \quad t > 0,$$

where $B' = Be^{\delta}$.

(2.4) Remark: The basic example from which the preceding theorem derives is the one treated by Nash. Namely, let $E = \mathbb{R}^N$ and set $P^o(t,x,dy) = (4\pi t)^{-N/2}\exp[-|y - x|^2/4t]dy$. Then it is easy to identify $\mathfrak{D}(\mathcal{E}^o)$ for the associated Dirichlet form $\mathcal{E}^o$ as the Sobolev space $W_2^1(\mathbb{R}^N)$ of $L^2(\mathbb{R}^N)$-functions with first derivatives in $L^2(\mathbb{R}^N)$ and to show that $\mathcal{E}^o(f,f) = \int|\nabla f|^2(x)dx$. In particular, since it is clear from the explicit form of $P^o(t,x,dy)$ that $\|\bar{P}_t\|_{1\to\infty} \le (4\pi t)^{-N/2}$, we can apply the preceding theorem to conclude that

(2.5) $$\|f\|_{L^2(\mathbb{R}^N)}^{2+4/N} \le A_N\left[\int|\nabla f|^2(x)dx\right]\|f\|_{L^1(\mathbb{R}^N)}^{4/N}.$$

On the other hand, and this is the direction in which Nash argued, an easy application of Fourier analysis establishes (2.5) for this example:

$$\begin{aligned}(2\pi)^N\|f\|_{L^2(\mathbb{R}^N)}^2 &\le \int_{|\xi|\le R}|\hat{f}(\xi)|^2d\xi + R^{-2}\int_{|\xi|\ge R}|(\nabla f)^{\wedge}(\xi)|^2d\xi \\ &\le \Omega_N R^N\|f\|_{L^1(\mathbb{R}^N)}^2 + (2\pi)^{-N}R^{-2}\int|\nabla f|^2(x)dx\end{aligned}$$

for all $R > 0$, and therefore (2.5) follows upon minimization with respect to R.

Next, suppose that $a:\mathbb{R}^N \longrightarrow \mathbb{R}^N \otimes \mathbb{R}^N$ is a smooth, symmetric matrix valued function which satisfies $a(\cdot) \geq \alpha I$ for some $\alpha > 0$. Then the fundamental solution $p(t,x,y)$ to $\partial_t u = \nabla\cdot(a\nabla u)$ determines a symmetric transition probability function $P(t,x,dy) = p(t,x,y)dy$ on $(\mathbb{R}^N,dx)$, and the associated Dirichlet form $\mathcal{E}$ is given by $\mathcal{E}(f,f) = \int \nabla f(x)\cdot a(x)\nabla f(x)dx$. While one now has no closed form expression for $P(t,x,dy)$, it is clear that $\mathcal{E}(f,f) \geq \alpha\mathcal{E}^o(f,f)$, and so from (2.5), we see that $\mathcal{E}$ satisfies (2.2) with $A = A_N/\alpha$. Hence, $\|\bar{P}_t\|_{1\longrightarrow\infty} \leq K/t^{N/2}$, where $K \in (0,\infty)$ depends on N and α alone. Obviously, this is the same as saying that $p(t,x,y) \leq K/t^{N/2}$.

The utility of Theorem 2.1 often lies in the fact that it translates a fairly transparent comparison of symmetric Markov semigroups at the infinitessimal level into information relating their kernels; clearly this is the case in Nash's original work.

Our next result is motivated by the following sort of example. Define $p(t,x,y) = \pi_t(y - x)$ on $(0,\infty)\times\mathbb{R}^N\times\mathbb{R}^N$, where $\pi_t(x) \equiv 2/\omega_N \frac{t}{(t^2 + |x|^2)^{(N+1)/2}}$ is the Cauchy (or Poisson) kernel for $\mathbb{R}^N$. Then it is easy to check (cf. the discussion in section 1)) that the associated Dirichlet form $\mathcal{E}$ is given by $\mathcal{E}(f,f) = 1/\omega_N\int dx\int dy\,|y|^{-N+1}(f(x+y) - f(x))^2$. In addition, by either Theorem (2.1) or a Fourier argument like the one given in (2.4), one sees that (2.2) holds with $\delta = 0$ and $\nu = N$. Next, consider the Dirichlet form $\mathcal{E}(f,f) = c\int dx\int dy\,|y|^{-N+1}(f(x+y) - f(x))^2\eta(y)$, where $c > 0$ and $\eta \in B_o(\mathbb{R}^N)^+$ is identically equal to 1 in a neighborhood of the origin and is even. (Note that, by the Levy-Khinchine formula, there is, for each $t > 0$, a unique probability μ_t on $\mathbb{R}^N$ such that $\hat{\mu}_t(\xi) = \exp\left[c' t\int dy\,|y|^{-N+1}(\cos(\xi\cdot y) - 1)\eta(y)\right]$, where $c' = 2c/(2\pi)^N$. Moreover, it is an easy exercise to check that the convolution semigroup $\bar{P}_t f = \mu_t * f$ is symmetric on $L^2(\mathbb{R}^N, dy)$ and has $\mathcal{E}$ as its Dirichlet form.) One can exploit translation invarience by using the Fourier transform to rewrite $\mathcal{E}(f,f)$ as

$$\mathcal{E}(f,f) = c'\int d\xi\left[|\hat{f}(\xi)|^2\int dy\,|y|^{-N+1}(1 - \cos(\xi\cdot y))\eta(y)\right].$$

Note that $\int dy\,|y|^{-N+1}(1 - \cos(\xi\cdot y))\eta(y)$ is asymptoticly proportional to $|\xi|^2$ for ξ small and to $|\xi|$ for ξ large. Then proceeding as in the Fourier analytic derivation of (2.5), one sees that there exists a $C \in (0,\infty)$ (depending only on N, c, $\|\eta\|_\infty$, and the supports of η and $(1 - \eta)$) such that:

$$\|f\|_2^2 \le C\left[(R^{-2} \vee R^{-1})\mathcal{E}(f,f) + R^N\|f\|_1^2\right],\ R > 0. \tag{2.6}$$

From (2.6), we see that if $\mathcal{E}(f,f) \ge \|f\|_1^2$ then $\|f\|^{2+2/N} \le C'\mathcal{E}(f,f)\|f\|_1^{2/N}$, where C' depends only on C and N. At the same time, if $\mathcal{E}(f,f) \le \|f\|_1^2$, then, by taking $R = 1$ in (2.6), we obtain $\|f\|_2^2 \le 2C\|f\|_1^2$ and therefore that $\|f\|_2^{2+2/N} \le (2C)^{1/N}\|f\|_2^2\|f\|_1^{2/N}$. Combining these, we arrive at

$$\|f\|_2^{2+2/N} \le A\left[\mathcal{E}(f,f) + \|f\|_2^2\right]\|f\|_1^{2/N},$$

where A depends only on N and C. Applying Theorem (2.1), we conclude that

$$\|\bar{P}_t\|_{1\to\infty} \le Be^t/t^N,\ t > 0. \tag{2.7}$$

Because the μ_t from which the preceding $\{\bar{P}_t : t > 0\}$ comes is nothing but a truncated Cauchy kernel, one expects that (2.7) is precise for $t \in (0,1]$. However, Central Limit Theorem considerations suggest that it is a very poor estimate for $t \geq 1$. In fact, because the associated stochastic process at any time t and for any $n \in Z^+$ is the sum of n independent random variables having variance approximately proportional to t/n, the Central Limit Theorem leads one to conjecture that the actual decay for large t is $Bt^{-N/2}$. The point is that too much of the information in (2.6) was thrown away when we were considering f's for which $\mathcal{E}(f,f) \leq \|f\|_1^2$. Indeed, from (2.6) we see that

(2.8) $$\|f\|_2^{2+4/N} \leq A\mathcal{E}(f,f)\|f\|_1^{4/N} \text{ when } \mathcal{E}(f,f) \leq \|f\|_1^2.$$

The next theorem addresses the problem of getting decay information from conditional Nash type inequalities like (2.8).

(2.9) Theoerem: Let $\nu \in (0,\infty)$ be given. If

(2.10) $$\|f\|_2^{2+4/\nu} \leq A\mathcal{E}(f,f)\|f\|_1^{4/\nu} \text{ when } \mathcal{E}(f,f) \leq \|f\|_1^2$$

for some $A \in (0,\infty)$ and if $\|\bar{P}_1\|_{1\to\infty} \leq B \in (0,\infty)$, then there is a $C \in (0,\infty)$ depending only on ν, A, and B such that

(2.11) $$\|\bar{P}_t\|_{1\to\infty} \leq C/t^{\nu/2}, \quad t \geq 1.$$

Conversely, (2.11) implies that (2.10) holds for some $A \in (0,\infty)$ depending only on ν and C.

Proof: As in the proof of Theorem (2.1), we restrict our attention to $f \in \mathcal{D}(\bar{A}) \cap L^1(m)^+$ when deriving these relations.

Assume that (2.10) holds and that $\|\bar{P}_1\|_{1\to\infty} \leq B$, and set $T = B/2$. Let $\mathcal{D}(\bar{A}) \cap L^1(m)^+$ with $\|f\|_1 = 1$ be given and define $f_t = \bar{P}_{T+t+1}f$, $t \geq 0$. Then, by (1.2):

$$\mathcal{E}(f_t,f_t) = \int_{[0,\infty)} \lambda e^{-2\lambda(T+t+1)} d(E_\lambda f,f) \le (1/2T)\|\bar{P}_1 f\|_2^2 \le \|f\|_1^2 = \|f_t\|_1^2.$$

Hence, by (2.10), $\|f_t\|_2^{2+4/\nu} \le A\mathcal{E}(f_t,f_t)\|f_t\|_1^{4/\nu} = A\mathcal{E}(f_t,f_t)$, since $\|f_t\|_1 = 1$. Starting from here, the derivation of $\|f_t\|_2^2 \le C'/t^{\nu/2}$ for some C' depending only on N and A is a re-run of the one given in the passage from (2.2) to (2.3). One now completes the proof of (2.11) by first noting that, from the preceding, $\|\bar{P}_{T+1+t}\|_{1\to\infty} \le 2^\nu C'^2/t^{\nu/2}$ and second that $\|\bar{P}_t\|_{1\to\infty} \le \|\bar{P}_1\|_{1\to\infty} \le B$ for $t \ge 1$.

The converse assertion is proved in the same way as we passed from (2.3) back to (2.2).

Q.E.D.

The following statement is an easy corollary of the Theorems (2.1) and (2.9) and the sort of reasoning used in the discussion immediately preceding the statement of (2.9).

(2.12) Corollary: Let $0 < \mu \le \nu < \infty$ be given. If

$$\|f\|_2^2 \le A\left[\left[\frac{\mathcal{E}(f,f)}{\|f\|_1^2}\right]^{\mu/(\mu+2)} + \left[\frac{\mathcal{E}(f,f)}{\|f\|_1^2}\right]^{\nu/(\nu+2)}\right]\|f\|_1^2 \tag{2.13}$$

for some $A \in (0,\infty)$ and all $f \in L^2(m)\setminus\{0\}$, then there is a B, depending only on μ,ν, and A, such that

$$\|\bar{P}_t\|_{1\to\infty} \le \begin{cases} B/t^{\nu/2} & \text{if } t \in (0,1] \\ B/t^{\mu/2} & \text{if } t \in [1,\infty). \end{cases} \tag{2.14}$$

(2.15) Remark: As a consequence of Corollary (2.12), we now have the following result. Let $\{\bar{P}_t : t > 0\}$ have Dirichlet form $\mathcal{E}$ and suppose that $\mathcal{E}(f,f) = \int dx \int (f(x+y) - f(x))^2 M(x,dy)$, where $M: \mathbb{R}^N \times \mathcal{B}_{\mathbb{R}^N\setminus\{0\}} \longrightarrow [0,\infty]$ has the properties that $M(x,\cdot)$ is a locally finite Borel measure on $\mathbb{R}^N\setminus\{0\}$ for each $x \in \mathbb{R}^N$, $M(\cdot,\Gamma)$ is a measurable function for each $\Gamma \in \mathscr{B}_{\mathbb{R}^N\setminus\{0\}}$, $M(x,-\Gamma) = M(x,\Gamma)$, and $\|\int |y|^2/(1 + |y|^2)M(\cdot,dy)\|_\infty \equiv C < \infty$. Next, suppose that $M(x,dy) \geq \eta(y)\frac{dy}{|y|^{N+\alpha}}$ for some $\eta \in B(\mathbb{R}^N)^+$ and $\alpha \in (0,2)$. If $\eta \geq \epsilon$ for some $\epsilon > 0$, then by comparison with the Dirichlet form of the symmetric stable semigroup of order α, we have $\|\bar{P}_t\|_{1\to\infty} \leq B/t^{N/\alpha}$, $t > 0$, where B depends only on N, α, ϵ, $\|\eta\|_\infty$, and C. On the other hand, again by comparison, if $\eta \in B_0(\mathbb{R}^N)$ and if $\eta \geq \epsilon > 0$ on some ball $B(0,r)$, then $\|\bar{P}_t\|_{1\to\infty}$ satisfies (2.14) with $\mu = N/2$, $\nu = 2N/\alpha$, and some B depending only on N, α, ϵ, r, $\|\eta\|_\infty$, and $\text{supp}(\eta)$.

We conclude this section with an explanation of the relationship between Nash inequalities like (2.2) and the more familiar Sobolev inequalies.

(<u>2.16</u>) <u>Theorem</u>: Let $\nu \in (2,\infty)$ be given and define $p \in (2,\infty)$ by the equation $p = 2\nu/(\nu - 2)$ (i.e. $1/p = 1/2 - 1/\nu$). If (2.2) holds for some choice of A and δ, then

$$\|f\|_p^2 \leq A'(\mathcal{E}(f,f) + \delta\|f\|_2^2) \tag{2.17}$$

for some $A' \in (0,\infty)$ which depends only on A and ν. Conversely, (2.17) implies (2.2) for some $A \in (0,\infty)$ depending only on A' and ν.

<u>Proof</u>: At least when $\delta = 0$, Varopoulos proved in [V-2] that (2.3) with $\nu > 2$ is equivalent to (2.17) with $p = 2\nu/(\nu-2)$; and so, since his proof extends easily to the case when $\delta > 0$, Theorem (2.16) follows directly from Varopoulos' theorem and Theorem (2.1).

Q.E.D.

The passage from (2.17) to (2.2) provided above is, however, far from being the most direct. If (2.17) holds, then by Holder's inequality:

$$\|f\|_2 \le \|f\|_p^{p'/2}\|f\|_1^{1-p'/2} \le A'(\mathcal{E}(f,f) + \delta\|f\|_2^2)^{p'/2}\|f\|_1^{1-p'/2},$$

where p' denotes the Holder conjugate of p. The preceding inequality clearly shows that (2.17) yields (2.2) with $A = (A')^{4/p'}$. In view of the crudeness of this argument for going from (2.17) to (2.2), it should come as no suprise that Varopoulos's proof that one can go from (2.3) to (2.17) involves somewhat subtle considerations. In particular, what comes easily from (2.3) is a weak-type version of (2.17); and one applies Marcinkiewicz interpolation to complete the job.

3. Davies's Method for Obtaining Off Diagnal Estimates:

So far we have discussed the derivation of estimates having the form $\|\bar{P}_t\|_{1\to\infty} \le B(t)$. When such an estimate obtains of course, for each t and m-a.e. x, the measure $P(t,x,\cdot)$ must be absolutely continuous with respect to m, and so the semigroup $\{\bar{P}_t: t > 0\}$ posseses a kernel $p(t,x,y)$; that is, for m-a.e. x, we may write $P(t,x,dy) = p(t,x,y)m(dy)$.

In this section we discuss pointwise estimates on the kernel $p(t,x,y)$. To do so conveniently, we will suppose that our semigroup $\{\bar{P}_t: t > 0\}$ is a Feller semigroup; that is, that each $\bar{P}_t$ preserves the space of bounded continuous functions. Under this hypothesis, whenever $\|\bar{P}_t\|_{1\to\infty} \le B(t)$ we have that for every t and x, $P(t,x,dy) = p(t,x,y)m(dy)$, and $p(t,x,\cdot) \le B(t)$ m-a.e. Then in view of the fact that $P(t,x,\cdot)$ is an m-symmetric transition probability function, $p(t,\cdot,*) = p(t,*,\cdot)$ (a.e.,m×m) for all $t > 0$, and $p(s+t,x,\cdot) = \int p(s,x,\xi)p(t,\xi,\cdot)m(d\xi) = \int p(s,x,\xi)p(t,\cdot,\xi)m(d\xi)$ (a.e.,m) for all $(t,x) \in (0,\infty)\times E$. (One may always delete the Feller condition in what follows if one is willing to insert extra a.e. conditions.)

We now enquire after the decay of $p(t,x,y)$ as the distance between x and y increases. The results of section 2) do not address this question. Indeed, under the Feller hypothesis, we have by the Schwarz inequality and the above that $p(t,x,y) \le (p(t,x,x))^{1/2}(p(t,y,y))^{1/2}$ for m×m-a.e. $(x,y) \in E\times E$. Hence, while an estimate on $\|\bar{P}_t\|_{1\to\infty}$ yields a uniform estimate on

$p(t,\cdot,*)$, it is really just an estimate on $p(t,\cdot,*)$ at the diagonal.

In the introduction we briefly sketched an extremely clever method E. B. Davies [D] introduced for obtaining off-diagnal estimates provided the semigroup is generated by a second order elliptic operator. Our primary goal in this section is to show how one can generalize Davies' idea and apply it in a more general non-local setting.

In order to explain what must be done, consider, for a moment, a typical situation handled by Davies. Namely, let $E = \mathbb{R}^N$ and suppose that $\mathcal{E}(f,f) = \int \nabla f \cdot a \nabla f dx$, where $a:\mathbb{R}^N \longrightarrow \mathbb{R}^N \otimes \mathbb{R}^N$ is a smooth, symmetric matrix-valued function, uniformly bounded above and below by positive multiples of the identity; and let $\{\bar{P}_t : t > 0\}$ denote the associated semigroup. Instead of studying the original semigoup $\{P_t: t > 0\}$ directly, Davies proceeded by way of the semigroup $\{P_t^{\psi}: t > 0\}$ where

$$P_t^{\psi} f(x) = e^{\psi(x)}[P_t(e^{-\psi}f)](x) , \tag{3.1}$$

and $\psi \in C_o^{\infty}(\mathbb{R}^N)$. What he showed then is that if $\|P_t\|_{1\longrightarrow\infty} \leq B/t^{\upsilon/2}$, $t > 0$, then, for each $\rho > 0$, there is a $B_\rho \in (0,\infty)$ such that

$$\|P_t^{\psi}\|_{1\longrightarrow\infty} \leq (B_\rho/t^{\upsilon/2})\exp((1 + \rho)\Gamma(\psi)^2 t), \ t > 0,$$

where $\Gamma(\psi)^2 = \| |\sum a^{ij}\partial_i\psi\partial_j\psi| \|_\infty$. As a consequence, he concluded that $p(t,x,y) \leq (B_\rho/t^{\upsilon/2})\exp(\psi(y) - \psi(x) + (1 + \rho)\Gamma(\psi)^2 t)$ for all $\psi \in C_o^{\infty}(\mathbb{R}^N)$ and then got his estimate by varying ψ .

As we will see shortly, the key to carrying out Davies' program is to obtain the inequality

(3.2) $$\mathcal{E}(e^{\psi}f^{2p-1}, e^{-\psi}f) \geq 1/p\mathcal{E}(f^p, f) - p\Gamma(\psi)^2\|f\|_{2p}^{2p}$$

for smooth non-negative f's and any $p \in [1,\infty)$. Although, in the case under consideration, (3.2) is an easy exercise involving nothing more than Leibniz's rule and Schwarz's inequality, it is not immediately clear what replaces (3.2) in the case of more general Dirichlet forms. In particular, we must find a satisfactory version of the Leibnitz rule (cf. (3.8) below) and a suitable quantity to play the role of $\Gamma(\psi)$, and we must then show that a close approximation of (3.2) continues to hold.

(<u>3.3</u>) <u>Warning</u>: Throughout this section we will be assuming that for any Dirichlet form $\mathcal{E}$ under consideration, $C_0(E)\cap\mathcal{D}(\mathcal{E})$ is dense in $C_0(E)$.

In this section we make frequent use of the fact that (cf. section 1)) for $f,g \in \mathcal{D}(\mathcal{E})$,

(3.4) $$\begin{aligned}\mathcal{E}(f,g) &= \lim_{t\downarrow 0}\mathcal{E}_t(f,g)\\ &\equiv \lim_{t\downarrow 0}\frac{1}{2t}\int(f(y) - f(x))(g(y) - g(x))m_t(dx\times dy)\end{aligned}$$

Set $\mathcal{F}_b \equiv \mathcal{D}(\mathcal{E})\cap L^\infty(m)$. We then have the following lemma, which is taken, in part, from [F].

(<u>3.5</u>) <u>Lemma</u>: If φ is a locally Lipshitz coninuous function on $\mathbb{R}^1$ with $\varphi(0) = 0$, then, for all $f \in \mathcal{F}_b$, $\varphi\circ f \in \mathcal{F}_b$. In particular, $\mathcal{F}_b$ is an algebra. Finally, for all $f,g \in \mathcal{F}_b$:

(3.6) $$\lim_{t\downarrow 0}\frac{1}{2t}\int g(x)(f(y) - f(x))^2 m_t(dx\times dy) = \mathcal{E}(gf,f) - 1/2\mathcal{E}(g,f^2).$$

<u>Proof</u>: The proof that $\varphi\circ f \in \mathcal{F}_b$ comes down to checking that

$$\sup_{t>0}\frac{1}{t}\int(\varphi\circ f(y) - \varphi\circ f(x))^2 m_t(dx\times dy) < \infty ;$$

and since $|\varphi\circ f(y) - \varphi\circ f(x)| \le M|f(y) - f(x)|$, where M is the Lipshitz norm of $\varphi\restriction\text{range}(f)$, this is clear. The fact that $\mathcal{F}_b$ is an algebra follows by specialization to $\varphi(\eta) = \eta^2$ and polarization. Finally, to prove (3.6), note that

$$\begin{aligned}(g(x)f(x) - g(y)f(y))(f(x) - f(y))& \\ - 1/2(g(x) - g(y))(f^2(x) - f^2(y))& \\ = 1/2g(x)(f(x) - f(y))^2 + 1/2g(y)(f(y) - f(x))^2;&\end{aligned}$$

and therefore, by the symmetry of m_t, one sees that

$$\begin{aligned}\int g(x)(f(y) - f(x))^2 m_t(dx\times dy) &= \int (g(x)f(x) - g(y)f(y))m_t(dx\times dy) \\ &\quad - 1/2\int (g(x) - g(y))(f^2(x) - f^2(y))m_t(dx\times dy).\end{aligned}$$

After dividing by 2t and letting $t\downarrow 0$, one gets (3.6). Q.E.D.

Given two measures μ and ν on $(E,\mathcal{B})$, recall that $(\mu\nu)^{1/2}$ is the measure which is absolutely continuous with respect to $\mu + \nu$ and has Radon-Nikodym derivative $(fg)^{1/2}$, where f and g denote the Radon-Nikodym derivatives of μ and ν, respectively, with respect to $\mu + \nu$.

(<u>3.7</u>) <u>Theorem</u>: Given $f,g \in \mathcal{F}_b$ and $t > 0$, define the measure $\Gamma(f,g)$ by

$$d\Gamma_t(f,g) = \left[\frac{1}{2t}\int (f(x) - f(y))(g(x) - g(y))P(t,x,dy)\right]m(dx).$$

Then, there is a measure $\Gamma(f,f)$ to which $\Gamma_t(f,f)$ tends weakly as $t\downarrow 0$ (i.e. $\int g(x)d\Gamma_t(f,f) \longrightarrow \int g(x)d\Gamma(f,f)$ for each $g \in C_b(E)$), and $\mathcal{E}(f,f)$ is the total mass of $\Gamma(f,f)$. Furthermore, if $\Gamma(f,g)$ is defined by polarization, then $\Gamma_t(f,g)$ tends weakly to $\Gamma(f,g)$ and $|\Gamma(f,g)| \le (\Gamma(f,f)\Gamma(g,g))^{1/2}$, where $|\sigma|$ denotes the variation measure associated with a signed measure σ.

Finally, if $f,g,h \in \mathcal{F}_0$, then one has the Leibnitz rule:

(3.8) $$\mathcal{E}(fg,h) = \int f d\Gamma(g,h) + \int g d\Gamma(f,h)$$

Proof: Clearly $\Gamma_t(f,f)(E) \longrightarrow \mathcal{E}(f,f)$ as $t \downarrow 0$. Thus we will know that $\Gamma_t(f,f)$ converges weakly as soon as we show that $\lim_{t\downarrow 0}\int g(x)\Gamma_t(f,f)(dx)$ exists for each $g \in C_0(E)$. In turn, since we have assumed that $\mathcal{D}(\mathcal{E}) \cap C_0(E)$ is dense in $C_0(E)$, we need only check this for $g \in \mathcal{D}(\mathcal{E}) \cap C_0(E)$; and for such a g we can apply (3.6).

Clearly both $\Gamma_t(f,g) \longrightarrow \Gamma(f,g)$ and the inequality $|\Gamma(f,g)| \leq (\Gamma(f,f)\Gamma(g,g))^{1/2}$ follow from the definition of $\Gamma(f,g)$ via polarization. Finally, to prove (3.8), observe that

$$\begin{aligned}(f(x)g(x) - f(y)g(y))(h(x) - h(y)) = \\ 1/2(g(x) + g(y))(f(x) - f(y))(h(x) - h(y)) \\ + 1/2(f(x) + f(y))(g(x) - g(y))(h(x) - h(y)).\end{aligned}$$

Hence, by the symmetry of m_t, (3.8) holds with $\mathcal{E}_t$ and Γ_t replacing $\mathcal{E}$ and Γ, respectively; and (3.8) follows upon letting $t \downarrow 0$. Q.E.D.

Clearly we can unambiguously extend the definition of $\mathcal{E}$ and Γ to $f,g \in \hat{\mathcal{F}} \equiv \{h + c\colon h \in \mathcal{F}_b \cap C_b(E) \text{ and } c \in \mathbb{R}^1\}$, and (3.8) will continue to hold even though elements of $\hat{\mathcal{F}}$ need not lie in $L^2(m)$. We now define $\hat{\mathcal{F}}_\infty$ to be the set of $\psi \in \hat{\mathcal{F}}$ such that $e^{-2\psi}\Gamma(e^{\psi},e^{\psi}) \ll m$, $e^{2\psi}\Gamma(e^{-\psi},e^{-\psi}) \ll m$, and

$$\Gamma(\psi) \equiv \left[\left\|\frac{de^{-2\psi}\Gamma(e^{\psi},e^{\psi})}{dm}\right\|_\infty \vee \left\|\frac{de^{2\psi}\Gamma(e^{-\psi},e^{-\psi})}{dm}\right\|_\infty\right]^{1/2} < \infty.$$

(3.9) Theorem: Choose and fix $\psi \in \hat{\mathcal{F}}_\infty$. Then, for all $f \in \hat{\mathcal{F}}^+$:

$$\mathcal{E}(e^{\psi}f, e^{-\psi}f) \geq \mathcal{E}(f,f) - \Gamma(\psi)^2\|f\|_2^2. \tag{3.10}$$

Moreover, all $p \in [2,\infty)$:

$$\mathcal{E}(e^{\psi}f^{2p-1}, e^{-\psi}f) \geq p^{-1}\mathcal{E}(f^p, f^p) - 9p\Gamma(\psi)^2\|f\|_{2p}^{2p}. \tag{3.11}$$

Proof: By polarizing (3.6), we see that:

$$\mathcal{E}(e^{\psi}f^{2p-1}, e^{-\psi}f) = \mathcal{E}(f^{2p-1}, f) + \mathcal{E}(e^{-\psi}f^{2p}, e^{\psi}) - 2\int e^{-\psi}f\,d\Gamma(f^{2p-1}, e^{\psi}).$$

Hence, after applying (3.8) to the second term on the right of the preceding, we obtain:

$$\begin{aligned}\mathcal{E}(e^{\psi}f^{2p-1}, e^{-\psi}f) = \mathcal{E}(f^{2p-1}, f) &+ \int f^{2p-1}d\Gamma(e^{-\psi}f, e^{\psi}) \\ &- \int e^{-\psi}f\,d\Gamma(f^{2p-1}, e^{\psi}).\end{aligned} \tag{3.12}$$

Note that

$$\begin{aligned}&\int f^{2p-1}d\Gamma(e^{-\psi}f, e^{\psi}) - \int e^{-\psi}f\,d\Gamma(f^{2p-1}, e^{\psi}) \\ &= \lim_{t\downarrow 0}\int [e^{-\psi(x)}f(x)f^{2p-1}(y) - e^{-\psi(y)}f(y)f^{2p-1}(x)] \\ &\qquad\times[e^{\psi(x)} - e^{\psi(y)}]m_t(dx\times dy)/2t.\end{aligned} \tag{3.13}$$

In particular, when $p = 1$:

$$\int f\,d\Gamma(e^{-\psi}f, e^{\psi}) - \int e^{-\psi}f\,d\Gamma(f, e^{\psi})$$

$$= \lim_{t\downarrow 0}\int f(x)f(y)[e^{-\psi(x)} - e^{-\psi(y)}][e^{\psi(x)} - e^{\psi(y)}]m_t(dx\times dy)/2t$$

$$= -\lim_{t\downarrow 0}\int f(x)f(y)[e^{-\psi(y)} - e^{-\psi(x)}][e^{\psi(x)} - e^{\psi(y)}]m_t(dx\times dy)/2t$$

$$\begin{aligned}\geq -\lim_{t\downarrow 0}&\left[\int f(x)^2[e^{-\psi(y)} - e^{-\psi(x)}][e^{\psi(x)} - e^{\psi(y)}]m_t(dx\times dy)/2t\right]^{1/2} \\ &\times\left[\int f(y)^2[e^{-\psi(y)} - e^{-\psi(x)}][e^{\psi(x)} - e^{\psi(y)}]m_t(dx\times dy)/2t\right]^{1/2}\end{aligned}$$

$$= \int f^2 d\Gamma(e^{-\psi}, e^{\psi}).$$

At the same time,

$$|\Gamma(e^{-\psi},e^{\psi})| \le \Gamma(\psi)^2 m, \tag{3.14}$$

and so (3.10) now follows from (3.12) with p = 1 and the preceding.

To prove (3.11) when $p \ge 2$, we re-write the right hand side of (3.13) as:

$$\lim_{t\downarrow 0}\int [f^{2p}(y) - f^{2p}(x)]e^{-\psi(x)}[e^{\psi(x)} - e^{\psi(y)}]m_t(dx\times dy)/2t$$

$$+ \lim_{t\downarrow 0}\int f^{2p}(x)[e^{-\psi(x)} - e^{-\psi(y)}][e^{\psi(x)} - e^{\psi(y)}]m_t(dx\times dy)/2t$$

$$+ \lim_{t\downarrow 0}\int f^{2p-1}(y)[f(x) - f(y)]e^{\psi(y)}[e^{\psi(x)} - e^{\psi(y)}]m_t(dx\times dy)/t$$

$$\ge \lim_{t\downarrow 0}\int f^{p}(y)[f^{p}(y) - f^{p}(x)]e^{\psi(y)}[e^{-\psi(y)} - e^{-\psi(x)}]m_t(dx\times dy)/2t$$

$$+ \lim_{t\downarrow 0}\int f^{p}(x)[f^{p}(y) - f^{p}(x)]e^{-\psi(x)}[e^{\psi(x)} - e^{\psi(y)}]m_t(dx\times dy)/2t$$

$$+ \int f^{2p}d\Gamma(e^{-\psi},e^{\psi}) - 2\left[\int f^{2p-2}d\Gamma(f,f)\right]^{1/2}\left[\int f^{2p}e^{-2\psi}d\Gamma(e^{\psi},e^{\psi})\right]^{1/2}$$

$$\ge -\mathcal{E}(f^p,f^p)^{1/2}\left[\left[\int f^{2p}e^{2\psi}d\Gamma(e^{-\psi},e^{-\psi})\right]^{1/2} - \left[\int f^{2p}e^{-2\psi}d\Gamma(e^{\psi},e^{\psi})\right]^{1/2}\right]$$

$$+ \int f^{2p}d\Gamma(e^{-\psi},e^{\psi}) - 2\left[\int f^{2p-2}d\Gamma(f,f)\right]^{1/2}\left[\int f^{2p}e^{-2\psi}d\Gamma(e^{\psi},e^{\psi})\right]^{1/2}.$$

Using (3.14) together with this last expression, we see that:

$$\begin{aligned}\int f^{2p-1}d\Gamma(e^{-\psi}f,e^{\psi}) - \int e^{-\psi}f\,d\Gamma(f^{2p-2},e^{\psi}) &\ge -\Gamma(\psi)^2\|f\|_{2p}^{2p}\\ &\quad - 2\left[\mathcal{E}(f^p,f^p)^{1/2} + \left[\int f^{2p-1}d\Gamma(f,f)\right]^{1/2}\right]\Gamma(\psi)\|f\|_{2p}^{p}.\end{aligned} \tag{3.15}$$

In order to complete the derivation of (3.11), we need two more facts. The first of these is that

$$\mathcal{E}(f^{2p-1},f) \geq \int f^{2p-2}d\Gamma(f,f) \geq \frac{1}{2p-1}\mathcal{E}(f^{2p-1},f) \tag{3.16}$$

and the second is that

$$\mathcal{E}(f^p,f^p) \geq \mathcal{E}(f^{2p-1},f) \geq \frac{2p-1}{p^2}\mathcal{E}(f^p,f^p). \tag{3.17}$$

To prove (3.16), use (3.8) to check that

$$\mathcal{E}(f^{2p-2},f^2) = 2\mathcal{E}(f^{2p-1},f) - 2\int f^{2p-2}d\Gamma(f,f),$$

and use $\mathcal{E}(f^{2p-2},f^2) = \lim_{t\downarrow 0}\mathcal{E}_t(f^{2p-2},f^2) \geq 0$ to conclude that the first part of (3.16) holds. The second part follows from the fact that for all x and y, $\frac{1}{2}(f^{2p-2}(y) + f^{2p-2}(x))(f(x) - f(y))^2 \geq \frac{1}{2p-1}(f^{2p-1}(y) - f^{2p-1}(x))(f(y) - f(x))$, together with (3.4) and Lemma (3.5). The proof of (3.17) is equally easy. Namely, replace $\mathcal{E}$ and by $\mathcal{E}_t$ and note that

$$(f^p(y) - f^p(x))^2 \geq (f^{2p-1}(y) - f^{2p-1}(x))(f(y) - f(x)) \geq \frac{2p-1}{p^2}(f^p(y) - f^p(x))^2.$$

(We do not actually use the second part of (3.16) here, but because it is interesting that there is a two sided bound, we include the short proof here. The second part of (3.17) has appeared already in [L.D.] and [V-2]; only the first part is new.)

Combining (3.12) and (3.15) with (3.16) and (3.17), we now see that

$$\mathcal{E}(e^{\psi}f^{2p-1},e^{-\psi}f) \geq \frac{2p-1}{p^2}\mathcal{E}(f^p,f^p) - 4\mathcal{E}(f^p,f^p)^{1/2}\Gamma(\psi) - \Gamma(\psi)^2\|f\|_{2p}^{2p},$$

of which (3.11) is an easy consequence. Q.E.D.

Now suppose that $\mathcal{E}$ satisfies the Nash inequality

$$(3.18)\qquad \|f\|_2^{2+4/\nu} \le A(\mathcal{E}(f,f) + \delta\|f\|_2^2)\|f\|_1^{4/\nu},\quad f \in L^2(m).$$

Given $\psi \in \hat{\mathcal{F}}_\infty$ and $f \in \hat{\mathcal{F}}^+$, set $f_t = \overline{P}_t^\psi f$. Then, by (3.10) and (3.11), one has that

$$\frac{d}{dt}\|f_t\|_2^2 = -2\mathcal{E}(e^\psi f_t, e^{-\psi} f_t) \le -2\mathcal{E}(f_t,f_t) + \Gamma(\psi)^2\|f_t\|_2^2$$

and

$$\frac{d}{dt}\|f_t\|_{2p}^{2p} = -2p\mathcal{E}(e^\psi f_t^{2p-1}, e^{-\psi} f_t) \le -2\mathcal{E}(f_t^p, f_t^p) + 18p^2\Gamma(\psi)^2\|f_t\|_{2p}^{2p}$$

for $p \in [2,\infty)$. Clearly the first of these implies that

$$(3.19)\qquad \|f_t\|_2 \le \exp(\Gamma(\psi)^2 t)\|f\|_2.$$

At the same time, when combined with (3.18), the second one leads to the differential inequality:

$$(3.20)\qquad \frac{d}{dt}\|f_t\|_{2p} \le -\frac{1}{Ap}\|f_t\|_{2p}^{1+4/\nu}\|f_t\|_p^{-4/\nu} + p(9\Gamma(\psi)^2 + \delta/p^2)\|f_t\|_{2p}$$

for $p \in [2,\infty)$.

The following lemma, which appears in [F-S] and whose proof is repeated here for the sake of completeness, provides the key to exploiting differential inequalities of the sort in (3.20).

(<u>3.21</u>) <u>Lemma</u>: Let $w: [0,\infty)\longrightarrow(0,\infty)$ be a continuous non-decreasing function and suppose that $u \in C^1([0,\infty);(0,\infty))$ satisfies

$$(3.22)\qquad u'(t) \le \frac{\epsilon}{p}\left[\frac{t^{(p-2)/\beta p}}{w(t)}\right]^{\beta p} u^{1+\beta p}(t) + \lambda p u(t),\quad t \in (0,\infty),$$

for some positive ϵ, β, and λ and some $p \in [2,\infty)$. Then, for each $\rho \in (0,1]$, u satisfies

$$(3.23)\qquad u(t) \le -\left[\frac{2p^2}{\rho\epsilon\beta}\right]^{1/\beta p} t^{(1-p)/\beta p} w(t) e^{\rho\lambda t/p},\quad t \in (0,\infty).$$

Proof: Set $v(t) = e^{-\lambda pt}u(t)$ and note that

$$v'(t) \le - \frac{\epsilon t^{(p-2)}}{p^2 w(t)^{\beta p}} e^{\lambda p^2 t} v(t)^{1+\beta p}.$$

Hence,

$$\frac{d}{dt}\left[v(t)^{-\beta p}\right] \ge \epsilon\beta t^{(p-2)} w(t)^{-\beta p} e^{\lambda\beta p^2 t}$$

and so, since w is non-decreasing,

$$e^{-\lambda\beta p^2 t} u(t)^{-\beta p} \ge \epsilon\beta w(t)^{-\beta p} \int_0^t s^{(p-2)} e^{\lambda\beta p^2 s} ds.$$

But, for $\rho \in (0,1]$,

$$\int_0^t s^{(p-2)} e^{\lambda\beta p^2 s} ds \ge \left[t/\lambda\beta p^2\right]^{p-1} \int_{\lambda\beta p^2(1-\rho/p^2)}^{\lambda\beta p^2} s^{(p-2)} s^{ts} ds$$

$$\ge \frac{t^{(p-1)}}{p-1} \exp[\lambda\beta p^2 t - \rho\lambda\beta t] p\left[1 - (1-\rho/p^2)\right].$$

Noting that $p\left[1 - (1-\rho/p^2)\right] \ge \rho/2$ for all $p \in [2,\infty)$, we conclude from the above that u satisfies (3.23).

Q.E.D.

We are now ready to complete our program of estimating $\|\bar{P}_t^{\psi}\|_{1\to\infty}$. To this end, pick an $f \in L^2(m)^+$ with $\|f\|_2 = 1$, set $p_k = 2^k$ for $k \in Z^+$, and define $u_k(t) = \|\bar{P}_t^{\psi} f\|_{p_k}$. Also define $w_k(t) = \max\{s^{(p_k-2)/\beta p_k} u(s): s \in (0,t]\}$. By (3.19), $w_1(t) \le \exp(\Gamma(\psi)^2 t)$. Moreover, by (3.20), u_{k+1} satisfies (3.22) with $\epsilon = 1/A$, $\beta = 4/\nu$, $\lambda = 9\Gamma(\psi)^2 + \delta$, and $w = w_k$. Hence, by (3.23), we see that $w_{k+1}(t)/w_k(t) \le \left[2^{2k+1}/\rho\epsilon\beta\right]^{1/2^k\beta} e^{\rho\lambda t/2^k}$ for any $\rho \in (0,1]$. Putting this together with our estimate on w_1, we arrive at the conclusion $\overline{\lim}_{k\to\infty} w_k(t) \le C(\rho\epsilon)^{-1/\beta} e^{\rho\lambda t}$, where $C = C(\beta) \in (0,\infty)$; and, after replacing ρ by $\rho/9$ and adjusting C accordingly, one easliy passes from here to

$$\|\bar{P}_t^{\psi}\|_{2\to\infty} \le C(A/\rho t)^{\upsilon/4}\exp[(1+\rho)\Gamma(\psi)^2 t + \rho\delta t]$$

for all $\rho \in (0,1]$. Finally, this estimate is obviously unchanged when ψ is replaced by $-\psi$. Thus, since it is clear that $\bar{P}_t^{-\psi}$ is the adjoint of $\bar{P}_t^{\psi}$, we also have that $\|\bar{P}_t^{\psi}\|_{1\to 2} \le C(A/\rho t)^{\upsilon/4}\exp[(1+\rho)\Gamma(\psi)^2 t + \rho\delta t]$ for all $\rho \in (0,1]$. Hence, since $\|\bar{P}_t^{\psi}\|_{1\to\infty} \le \|\bar{P}_{t/2}^{\psi}\|_{1\to 2}\|\bar{P}_{t/2}^{\psi}\|_{2\to\infty}$, we now have

(3.24) $$\|\bar{P}_t^{\psi}\|_{1\to\infty} \le C(A/\rho t)^{\upsilon/2}\exp[(1+\rho)\Gamma(\psi)^2 t + \rho\delta t]$$

for all $\rho \in (0,1]$, where the C in (3.24) is the square of the earlier C.

(<u>3.25</u>) <u>Theorem</u>: Assume that (3.18) holds for some positive υ, A, and δ. Then $P(t,x,dy) = p(t,x,y)m(dy)$ where, for each $\rho \in (0,1]$ and all $(t,x,y) \in (0,\infty)\times E\times E$:

(3.26) $$p(t,x,\cdot) \le C(A/\rho t)^{\upsilon/2} e^{\delta\rho t} e^{-D((1+\rho)t;x,\cdot)} \quad \text{(m-a.e.)}$$

with $C \in (0,\infty)$ depending only on υ and

(3.27) $$D(T;x,y) \equiv \sup\{|\psi(y) - \psi(x)| - T\Gamma(\psi)^2 : \psi \in \hat{\mathcal{F}}_\infty\}.$$

<u>Proof</u>: From (3.24) with $\psi = 0$ we see that $p(t,x,\cdot)$ exists. Moreover, since $\Gamma(\psi) = \Gamma(-\psi)$, (3.24) for general $\psi \in \hat{\mathcal{F}}_\infty$ says that

$$p(t,x,\cdot) \le C(A/\rho t)^{\upsilon/2}\exp\Big[\delta\rho t - |\psi(\cdot) - \psi(x)| + (1+\rho)T\Gamma(\psi)^2\Big],$$

and clearly (3.26) follows from this. Q.E.D.

(<u>3.28</u>) <u>Corollary</u>: Assume that (2.14) holds for some $B \in (0,\infty)$ and $0 < \mu \le \upsilon < \infty$ (or, equivalently, that (2.13) holds for some $A \in (0,\infty)$ and the same μ and υ). Then for all $(t,x,y) \in (0,\infty)\times E\times E$ and each $\rho \in (0,1]$:

$$
(3.29)\quad p(t,x,y) \le \begin{cases} K(\rho t)^{-\nu/2} e^{\delta\rho t} \exp[-D((1+\rho)t;x,y)] & \text{for } t \in (0,1] \\ K(\rho t)^{-\mu/2} e^{\delta\rho t} \exp[-D((1+\rho)t,x,y)] & \text{for } t \in [1,\infty) \end{cases}
$$

where $K \in (0,\infty)$ depends only on B (or A), μ, and ν.

<u>Proof</u>: From (2.14) we have (cf. the proof that (2.3) implies (2.2)) that

$$
(3.30)\quad \|f\|_2^2 \le \begin{cases} Bt^{-\nu/2}\|f\|_1^2 + t\mathcal{E}(f,f), & t \in (0,1] \\ Bt^{-\mu/2}\|f\|_1^2 + t\mathcal{E}(f,f), & t \in [1,\infty). \end{cases}
$$

Hence, if $\delta \in (0,1]$, then $\|f\|_2^2 \le Bt^{-\nu/2}\|f\|_1^2 + t\mathcal{E}(f,f)$ for all $t \in (0,1/\delta]$. In particular, by taking

$$
t = \left[\nu B\delta^{(\mu-\nu)}\|f\|_1^2/2\mathcal{E}(f,f)\right]^{2/(\nu+2)},
$$

we conclude that there is a $B' \in (0,\infty)$, depending only on B, μ, and ν, such that

$$
(3.31)\quad \|f\|_2^{2+4/\nu} \le B'\delta^{\mu/\nu-1}\mathcal{E}(f,f)\|f\|_1^{4/\nu} \text{ if } \|f\|_1^2 \le \frac{2\delta^{-1-\mu/2}}{\nu B}\mathcal{E}(f,f).
$$

On the other hand, by taking $t = 1/\delta$ in (3.30), we see that $\|f\|_2^2 \le B\delta^{\mu/2}\|f\|_1^2 + \delta^{-1}\mathcal{E}(f,f) \le B(1+\nu/2)\delta^{\mu/2}\|f\|_1^2$ and therefore that $\|f\|_2^{4/\nu} \le (B(1+\nu/2)\delta)^{\mu/\nu}\|f\|_2^2\|f\|_1^{4/\nu}$ if $\|f\|_1^2 \ge \frac{2\delta^{-1-\mu/2}}{\nu B}\mathcal{E}(f,f)$. Combining this with (3.30), we conclude that

$$
(3.32)\quad \|f\|_2^{4/\nu} \le A\delta^{\mu/\nu-1}\left[\mathcal{E}(f,f) + \delta\|f\|_2^2\right]\|f\|_1^{4/\nu}, \quad \delta \in (0,1],
$$

where $A \in (0,\infty)$ depends only on B, μ, and ν.

Finally, given $t \in (0,\infty)$, (3.29) follows from (3.32) with $\delta = 1/(1 \vee t)$ and Theorem (3.25).

Q.E.D.

4. The Discrete Time Case:

All our considerations thus far have applied to symmetric Markov semigroups in continuous time for the simple reason that Dirichlet considerations are most natural in that context. However, it is often important to work with a discrete time parameter; and so in the present section we develop the discrete-time analogs of the results in section 2). Unfortunately, we do not know how to extend the results of section 3) to this setting.

Throughout this section $\Pi(x,dy)$ will denote an m-symmetric transition probability on $(E,\mathcal{B})$. Also, we will use $\Pi f(x)$ to denote $\int f(y)\Pi(x,dy)$; and, for $n \geq 1$, the transition function $\Pi^n(x,dy)$ and the operator Π^n are defined inductively by iteration. Note that $\|\Pi\|_{p\to p} = 1$ for all $p \in [1,\infty)$. Finally, set $M(dx\times dy) \equiv \Pi^2(x,dy)m(dx)$ and associate with Π the Dirichlet form $\mathcal{E}(f,f) \equiv 1/2\int(f(y) - f(x))^2 M(dx\times dy)$.

Obviously there is no "small time" in the discrete context and therefore we only seek an analog of Theorem (2.9).

(4.1) Theorem: Let $\nu \in (0,\infty)$ be given. If

(4.2) $$\|f\|_2^{2+4/\nu} \leq A\mathcal{E}(f,f)\|f\|_1^{4/\nu} \text{ when } \mathcal{E}(f,f) \leq \|f\|_1^2$$

for some $A \in (0,\infty)$ and if $\|\Pi\|_{1\to\infty} \leq B \in (0,\infty)$, then there is a $C \in (0,\infty)$ depending only on ν, A, and B such that

(4.3) $$\|\Pi^n\|_{1\to\infty} \leq C/n^{\nu/2},\ n \geq 1.$$

Conversely, (4.3) implies that (4.2) holds for some $A \in (0,\infty)$ depending only on ν and C.

<u>Proof</u>: We begin by observing that

$$(4.4)\qquad \mathcal{E}(f,f) = \|f\|_2^2 - \|\Pi f\|_2^2.$$

In particular,

$$(4.5)\qquad \mathcal{E}(f,f) - \mathcal{E}(\Pi f,\Pi f) = \int (f(x) - \Pi f(x))^2 m(dx) \geq 0,$$

$$\|\Pi f\|_2^2 \geq \sum_1^\infty \left[\|\Pi^n f\|_2^2 - \|\Pi^{n+1} f\|_2^2\right] = \sum_1^\infty \mathcal{E}(\Pi^n f,\Pi^n f),$$

and so

$$(4.6)\qquad \mathcal{E}(\Pi^n f,\Pi^n f) \leq \|\Pi f\|_2^2/n,\ n \geq 1.$$

Now suppose that (4.2) holds and that $\|\Pi\|_{1\to\infty} \leq B$. Then

$$\|\Pi\|_{1\to 2} \leq \|\Pi\|_{1\to 1}^{1/2}\|\Pi\|_{1\to\infty}^{1/2} \leq B^{1/2} ;$$

and so, by (4.6), $\mathcal{E}(\Pi^n f,\Pi^n f) \leq \|\Pi^n f\|_1^2$ for $n \geq N_0 \equiv [B] + 1$. Hence, if $f \in L^1(m)^+$ with $\|f\|_1 = 1$ and $u_n \equiv \|\Pi^n f\|_2^2$, then, by (4.2) and (4.4)

$$(4.7)\qquad u_{n+1} \leq (1 - u_n^{2/\nu}/A)u_n,\ n \geq N_0.$$

Next, choose $N_1 \geq N_0$ so that $(1 - B^{2/\nu}/A(n+1)) \leq (n/(n+1))^{\nu/2}$ for all $n \geq N_1$, and set $C = BN_1^{\nu/2}$. Clearly, $u_n \leq C/n^{\nu/2}$ for $1 \leq n \leq N_1$. Moreover, if $n \geq N_1$ and $u_n \leq C/n^{\nu/2}$, then either $u_n \leq C/(n+1)^{\nu/2}$ or $C/(n+1)^{\nu/2} < u_n \leq C/n^{\nu/2}$. In the first case, since $u_{n+1} \leq u_n$, $u_{n+1} \leq C/(n+1)^{\nu/2}$. On the other hand, in the second case, we apply (4.7) to obtain:

$$u_{n+1} \leq \left[1 - (C/(n+1)^{\nu/2})^{2/\nu}/A\right]u_n \leq (n/(n+1))^{\nu/2}C/n^{\nu/2} \leq C/(n+1)^{\nu/2}.$$

Hence, by induction on $n \geq N_1$, we see that $u_n \leq C/n^{\nu/2}$ for all $n \geq 1$. Obviously, this implies that $\|\Pi^n\|_{1\to 2} \leq C^{1/2}/n^{\nu/4}$; and therefore, by the usual duality argument, (4.3) follows.

To prove that (4.3) implies (4.2), we use (4.4) and (4.5) to conclude that $\|\Pi^n f\|_2^2 - \|f\|_2^2 \geq n\mathcal{E}(f,f)$ and therefore, if (4.3) holds, that $\|f\|_2^2 \leq (C/(2n)^{\nu/2})\|f\|_1^2 + n\mathcal{E}(f,f)$, $n \geq 1$. The passage from here to (4.2) is just the same sort of minimization procedure as was used to get (2.10) from (2.11). Q.E.D.

As a typical application of Theorem (4.1), we present the following. Take $E = \mathbb{R}^N$ and suppose that $\Pi(x,dy) = \pi(x,y)dy$ where π is a symmetric measurable function on $\mathbb{R}^N \times \mathbb{R}^N$ into $[0,B]$ for some $B \in (0,\infty)$. Assume, in addition, that $\pi^2(\cdot,*) \geq \rho(\cdot,*)$ almost everywhere, where ρ is an even function in $L^1(\mathbb{R}^N)^+$ satisfying

(4.8) $$\int (1 - \cos(\xi\cdot y))\rho(y)dy \geq \epsilon|\xi|^\alpha,\ \xi \in \mathbb{R}^N \text{ with } |\xi| \leq 1$$

for some positive α and ϵ.

(<u>4.9</u>) <u>Corollary</u>: Referring to the preceding, there is a $C \in (0,\infty)$, depending only on N, α, ϵ, and B, such that $\pi^n(x,\cdot) \leq C/n^{N/\alpha}$ a.e. for all $x \in \mathbb{R}^N$ and $n \geq 1$.

<u>Proof</u>: Note that

$$\begin{aligned}(2\pi)^N\mathcal{E}(f,f) &\geq (2\pi)^N\int dx\int (f(x+y) - f(x))^2\rho(y)dy \\ &= 2\int\left[\int (1 - \cos(\xi\cdot y))\rho(y)dy\right]|\hat{f}(\xi)|^2 d\xi.\end{aligned}$$

Hence, by (4.8),

$$\begin{aligned}(2\pi)^N\|f\|_2^2 &= \|\hat{f}\|_2^2 = \int_{|\xi|\leq R} |\hat{f}(\xi)|^2 d\xi \\ &+ \int_{|\xi|\geq R} |f(\xi)|^2 d\xi \leq \Omega_N R^N\|f\|_1^2 + \left[(2\pi)^N/2\epsilon R^\alpha\right]\mathcal{E}(f,f)\end{aligned}$$

for all $R \in (0,1]$; and from here it is an easy step to (4.2) with

$\upsilon = 2N/\alpha$ and an $A \in (0,\infty)$ depending only υ, ϵ, and N. Since $\|\Pi\|_{1\to\infty} \leq B$, we can now apply Theorem (4.1) to get the required conclusion.

Q.E.D.

5. Assorted Applications:

We conclude this paper with an assortment of applications of results from previous sections and with some remarks on natural extensions of these results.

Most of these applications, like most of those already discussed, exploit a relatively transparent comparison of Dirichlet forms to yield an interesting comparison of the associated semigroups. By way of counterpoint, the following application of Theorem (2.1) exploits a relatively transparent "multiplicative" property of Markov semigroups to establish an interesting "multiplicative" property of the associated Dirichlet spaces.

Let $E_{(1)}$ and $E_{(2)}$ be two locally compact metric spaces equipped with measures m_1 and m_2, and with symmetric transition probability functions $P^{(1)}(t,x^1,\cdot)$ and $P^{(2)}(t,x^2,\cdot)$, as in the first section. Let $\mathcal{E}^{(1)}$ and $\mathcal{E}^{(2)}$ be the correspnding Dirichlet forms.

Clearly

$$(5.1) \qquad P(t,(x^1,x^2),\cdot) = P^{(1)}(t,x^1,\cdot)\otimes P^{(2)}(t,x^2,\cdot)$$

is a transition probability function on $(E_{(1)}\times E_{(2)},\mathcal{B}_{E_{(1)}\times E_{(2)}})$, which is symmetric with respect to $m \equiv m_1\times m_2$. It is further clear that $P(t,(x^1,x^2),\cdot)$ tends weakly to $\delta_{(x^1,x^2)}$ as t tends to zero, and so (5.1) defines a transition function of the type we have been considering. Let $\mathcal{E}$ be the corresponding Dirichlet form; then it is easy to see that as Hilbert spaces (the inner product on $\mathcal{D}(\mathcal{E})$ being $(\cdot,\cdot) + \mathcal{E}(\cdot,\cdot)$, etc.)

(5.2) $$\mathcal{D}(\mathcal{E}) = \mathcal{D}(\mathcal{E}^{(1)})\otimes\mathcal{D}(\mathcal{E}^{(2)}) .$$

Now suppose that $\mathcal{E}^{(1)}$ and $\mathcal{E}^{(2)}$ each satisfy a Nash type inequality (2.2) for some positive ν_1 and ν_2 . One may naturally ask whether $\mathcal{E}$ then satisfies (2.2) for some ν depending on ν_1 and ν_2 .

It may seem that this question invites an approach using, say, Holder's inequality or Minkowskii's inequality to take apart tensor products directly in (2.2). We know of no such argument. However, the equivalence of (2.2) and (2.3) provides an easy positive answer to the question.

<u>(5.4)</u> <u>Theorem</u>: Let $\mathcal{E}$, $\mathcal{E}^{(1)}$, and $\mathcal{E}^{(2)}$ be related as above, and suppose

(5.5) $$\|f\|_2^{2+4/\nu^i} \le A^{(i)}\left[\mathcal{E}^{(i)}(f,f) + \delta^i\|f\|_2^2\right]\|f\|_1^{4/\nu^i}, \quad f \in L^2(m_i) ,$$

for $i = 1,2$

Then with $\nu = \nu_1 + \nu_2$, $\delta = \delta_1 + \delta_2$ and some $A \in (0,\infty)$, depending only on $A^{(1)}\vee A^{(2)}$:

(5.6) $$\|f\|_2^{2+4/\nu} \le A\left[\mathcal{E}(f,f) + \delta\|f\|_2^2\right]\|f\|^{4/\nu}, \quad f \in L^2(m_1\otimes m_2)$$

Furthermore, provided ν_1 and ν_2 are the smallest values for which (5.5) holds, $\nu_1 + \nu_2$ is the smallest value of ν for which (5.6) holds.

<u>Proof</u>: Let $\{\bar{P}_t^{(1)}: t > 0\}$ and $\{\bar{P}_t^{(2)}: t > 0\}$ be the semigroups corresponding to $\mathcal{E}^{(1)}$ and $\mathcal{E}^{(2)}$. By (5.5) and the

second half of Theorem (2.1), $\|\bar{P}_t^{(i)}\|_{1\to\infty} \le B^{(i)}e^{\delta^i t}/t^{\nu^i/2}$; $i = 1,2$. Then, by Segal's lemma [S],

(5.7) $$\|\bar{P}_t^{(1)}\otimes\bar{P}_t^{(2)}\|_{1\to\infty} \le B^{(1)}B^{(2)}e^{(\delta^1+\delta^2)t}/t^{(\nu^1+\nu^2)/2} ;$$

and so, by the first half of Theorem (2.1), we have (5.6). The optimality of $\nu_1 + \nu_2$ is easily seen by applying $\bar{P}_t^{(1)}\otimes\bar{P}_t^{(2)}$ to the product $f_1\otimes f_2$ where each f_i is chosen with $\|f_i\|_1 = 1$ and $\|\bar{P}_t^{(i)}f_i\|_\infty$ very close to $\|\bar{P}_t^{(i)}\|_{1\to\infty}$. Q.E.D.

A particularly interesting case occurs when $\nu > 2$ in (5.6). Then Theorem (2.17) says that a Sobolev inequality holds for $\mathcal{E}$. This provides an easy way to see that Sobolev inequalities hold for certain Dirichlet forms, and even to find the largest possible p (smallest possible ν) for which the inequality holds.

For the simplest sort of example, take $E_{(1)} = [0,1]$, take m_1 to be xdx, and define $\mathcal{E}^{(1)}$ by

(5.8) $$\mathcal{E}^{(1)}(f,f) = \frac{1}{2}\int_0^1 |f'(x)|^2 xdx$$

for $f \in C_b^\infty([0,1])$ and then closing. Regarding f as a radial function on the unit disk in $\mathbb{R}^2$, one recognizes $\mathcal{E}^{(1)}$ as the restriction to radial functions of the Dirichlet form associated with the Neumann heat kernel on the unit disk in $\mathbb{R}^2$. $\mathcal{E}^{(1)}$ therefore satisfies (5.5) with $\nu_1 = 2$. Next take $E_{(2)}$ to be the unit cube in $\mathbb{R}^{N-1}$, take m_2 to be Lebesgue measure, and take $\mathcal{E}^{(2)}$ to be the Dirichlet form associated with the Neumann heat kernel on $E_{(2)}$. Then with $E = E_{(1)}\times E_{(2)} \subset \mathbb{R}^N$, and with $\mathcal{E}$, $\mathcal{E}^{(1)}$, $\mathcal{E}^{(2)}$ related as above, for any $f \in C_b^\infty(E)$..

(5.9) $$\mathcal{E}(f,f) = \frac{1}{2}\int_0^1 x^1 dx^1 \int_{-1}^1 dx^2 \cdots \int_{-1}^1 dx^N |\nabla f(x)|^2$$

Then clearly Theorem (5.4) applies with $\nu_1 = 2$ and $\nu_2 = N-1$, and so $\mathcal{E}$ satisfies (5.6) with $\nu = N+1$, and does not satisfy (5.6) for any smaller value of ν. Therefore when $N \geq 2$ $\mathcal{E}$ satisfies a Sobolev inequality

(5.10) $$\|f\|_p^2 \leq A'\left[\mathcal{E}(f,f) + \delta\|f\|_2^2\right]$$

with $1/p = 1/2 - 1/N+1$; (5.10) fails for any larger value of p. (The L^p norms are computed with respect to $x^1 dx$.) Of course, if we remove the factor x^1 from the integrals, (5.10) then is satisfied with $1/p = 1/2 - 1/N$. Including the degenerate weight x^1 in our integrals raises the effective dimension ν by one from N to $N+1$.

The same result obtains in less special situations. Let M be a smooth, compact $N-1$ dimensional submanifold of $\mathbb{R}^N$. Let ρ be a weight function on $\mathbb{R}^N$ satisfying, for some $\lambda > 0$, and all x

(5.11) $$\lambda(\mathrm{dist}(x,M)\wedge 1) \leq \rho(x) \leq \lambda^{-1}(\mathrm{dist}(x,M)\wedge 1)$$

By standard results in, for example, Fukushima's book [F]; the closure of

(5.12) $$\mathcal{E}(f,f) = \int_M |\nabla f(x)|^2 \rho(x) dv,$$

defined first for $f \in C_o^\infty(\mathbb{R}^N)$, is a Dirichlet form. Employing a simple partitioning argument, familiar comparison arguments, and otherwise only increasing the complexity of notation; the argument above yields the following result: For some A', $\delta \in (0,\infty)$, $\mathcal{E}$ satisfies the Sobolev inequality (5.10) with $1/p = 1/2 - 1/N+1$.

Before leaving this subject, we briefly look at the limiting case $\nu = 2$. Although $p(\nu) = 2\nu/(\nu-2)$ tends to infinity as ν decreases to 2, it is easy to see that when $\nu = 2$, $\mathcal{E}$ does not in general control the sup norm. There is however a natural definition of the B.M.O. norm in the general Dirichlet form setting. In terms of this B.M.O. norm, one easily obtains a strong limiting case of the Sobolev inequality (5.10) holding whenever $\nu = 2$ holds in (5.6).

Let $\mathcal{E}$ be a fixed Dirichlet form, with $\{\bar{P}_t : t > 0\}$ being the associated semigroup on $L^2(m)$. Using the spectral theorem and the integral $e^{-\lambda} = \pi^{-1/2}\int_0^\infty \frac{dt}{t} e^{-\lambda^2/4t} e^{-t} t^{1/2}$, one sees that with $\bar{Q}_t$ given by $\bar{Q}_t = \pi^{-1/2}\int_0^\infty \frac{ds}{s}\left[\bar{P}_{(s^2/4)}\right] e^{-s} s^{-1/2}$, $\{\bar{Q}_t : t > 0\}$ is a Markov semigroup on $L^2(m)$ generated by $-(\bar{A})^{1/2}$, where $-\bar{A}$ is the generator of $\{\bar{P}_t : t > 0\}$. The B.M.O. norm naturally associated to $\mathcal{E}$ is given by

$$\text{(5.13)} \qquad \|f\|_{B.M.O.} = \sup_{t>0}\left[\|\bar{Q}_t|f - \bar{Q}_t f|\|_\infty\right]$$

(This definition was used by Stroock [St] who established a generalization of the John-Nirenberg inequality; proving that when $m(E) < \infty$ and $\{\bar{Q}_t : t > 0\}$ is a Feller semigroup (so that the corresponding Markov process can be constructed with right continuous paths [W]), there exists an $\alpha > 0$, and a $B < \infty$ so that for all f with $\|f\|_{B.M.O.} < \infty$

$$\text{(5.14)} \qquad \int_E \left[\exp[\alpha f/\|f\|_{B.M.O.}]\right] dm \le B$$

(Note that $\{\bar{Q}_t : t > 0\}$ is a Feller semigroup whenever

$\{\bar{P}_t: t > 0\}$ is a Feller semigroup.) This exponential integrability is what supports the assertion that the B.M.O. norm is a strong substitute for the sup norm. For further discussion of such results, see [D-M].

Now suppose $\mathcal{E}$ satisfies (5.6) with $\nu = 2$. Then $\|\bar{P}_t\|_{1\to\infty} \leq C/t$. The integral representation for $\bar{Q}_t$ shows that then $\|\bar{Q}_t\|_{1\to\infty} \leq C/t^2$, and so by interpolation between this and $\|\bar{Q}_t\|_{\infty\to\infty} = 1$, $\|\bar{Q}_t\|_{2\to\infty} \leq C/t$. (C is of course changing from line to line.) Now suppose that $f \in \mathcal{D}(\mathcal{E})$. Then $t \longrightarrow \bar{Q}_t f$ is strongly differentiable and

$$(5.15) \qquad \bar{Q}_t f - f = -\int_0^t ds \left[\bar{Q}_s \bar{A}^{1/2} f\right] .$$

This gives the estimate $\|\bar{Q}_t f - f\|_2 \leq t\|\bar{A}^{1/2} f\|_2 = t\mathcal{E}(f,f)^{1/2}$ and consequently $\|\bar{Q}_t(\bar{Q}_t f - f)\|_\infty^2 \leq C\mathcal{E}(f,f)$, so that $\|f\|_{B.M.O.}^2 \leq C\mathcal{E}(f,f)$. This discussion is summarized in the following result:

(5.16) Theorem: Let $\mathcal{E}$ be a Dirichlet form such that $\|\bar{P}_t\|_{1\to\infty} \leq C/t$ for all $t \in (0,1)$. Then there is a $C' < \infty$, depending only on C, so that $\|f\|_{B.M.O.}^2 \leq C'\mathcal{E}(f,f)$; and consequently, when $m(E) < \infty$ and $\{\bar{P}_t: t > 0\}$ is a Feller semigroup, there is an $\alpha > 0$, and a $B < \infty$ so that

$$(5.17) \qquad \int_E \left[\exp[\alpha f/\mathcal{E}(f,f)^{1/2}]\right] dm \leq B$$

for all $f \in \mathcal{D}(\mathcal{E})$.

(5.18) Remark: It is not clear to us whether the preceding result has a converse.

We next turn to an application of the results in section 4).

Take $E \subset Z^N$ equipped with the usual metric and a measure m bounded above and below by positve multiples of counting measure. Suppose that E is everywhere connected to infinity, by which we mean that for each $x \in E$, there is an infinite, one sided, loop free chain E_x in E of nearest neighbors starting at x. (One may always erase loops if need be.) Now let $\Pi(x,\cdot)$ be an m-symmetric transition function on E, define $\pi(x,y) = \Pi(x,\{y\})/m(\{y\})$, and assume that

$$(5.19) \qquad 1/\mu \geq \pi(x,y) \geq \mu$$

for some $\mu \in (0,1]$ and all x and y in E which are nearest neighbors. One naturally feels that the associated random walk must spread out at least as fast as a simple random walk on the half line with transition probabilities μ, since starting at x, it can always spread out along E_x. That is, one expects the return probabilities $\Pi^n(x,\{x\})$ to decay like $C/n^{1/2}$. The results of section 4) permit an easy proof of this.

(5.20) Theorem: Let $E \subset Z^N$, Π and m be given as in the preceeding discussion. Then there is a $C < \infty$ depending only on m and μ so that

$$(5.21) \qquad \Pi^n(x,\{x\}) \leq C/n^{1/2} \quad \text{for all } x \in E \text{ and } n \in Z^+.$$

Proof: Let $\mathcal{E}$ denote the Dirichlet form associated with Π^2 as in section 4). Given $x \in E$, let E_x be an infinite, loop free, one sided chain of nearest neighbors in E starting at x. Let $\mathcal{E}_x$ be the Dirichlet form on $L^2(m)$ given by

(5.22) $$\mathcal{E}^{(x)}(f,f) = \sum_{y,z\in E_x} (f(y) - f(z))^2 \Pi^2(y,\{z\})m(z)$$

Clearly $\mathcal{E}^{(x)}(f,f) \leq \mathcal{E}(f,f)$ for all f, so that if $\bar{A}^{(x)}$ denotes the self adjoint operator associated with $\mathcal{E}^{(x)}$, as $\bar{A}$ is with $\mathcal{E}$, then, for any $\lambda > 0$:

(5.23) $$(\bar{A} + \lambda)^{-1} \leq (\bar{A}^{(x)} + \lambda)^{-1}$$

Letting G_λ and $G_\lambda^{(x)}$ denote the kernels of the above operators (with respect to m), (5.23) says in particular that

(5.24) $$G_\lambda(x,x) \leq G_\lambda^{(x)}(x,x)$$

Now identify E_x with the natural numbers $\mathcal{N}$ in the obvious way so that x is identified with 0. By restriction and this identification, we may regard $m_x \equiv m|_{E_x}$ as a measure on $\mathcal{N}$ and $\mathcal{E}^{(x)}$ as a Dirichlet form on the L^2-space over $\mathcal{N}$ relative to this measure. Next, define m_w on $\mathcal{N}$ to be the measure which assigns mass 1 to each element of Z^+ and mass 2 to 0, and define $\mathcal{E}_w$ by

(5.25) $$\mathcal{E}_w(f,f) = \sum_{j,k\in\mathcal{N}} (f(j) - f(k))^2 \Pi_w^2(j,\{k\})m_w(k) ,$$

where $\Pi_w(0,\{1\}) = 1$ and $\Pi_w(n,\{n+1\}) = \Pi_w(n+1,\{n\}) = 1/2$ for all $n \in Z^+$. This is the Dirichlet form of the simple random walk on $\mathcal{N}$ reflected at 0. Since the simple random walk transition function satisfies $\Pi_w^n(k,\{k\}) \leq C/n^{1/2}$, $n \geq 1$, for some $C > 0$ and all $k \in \mathcal{N}$ (this well known fact is also a consequence of Lemma (4.9)), application of Theorem (4.1) yields

(5.26) $$\|f\|^6_{L^2(m_w)} \leq A\mathcal{E}_w(f,f)\|f\|^4_{L^1(m_w)} .$$

But since both m_w and m_x as well as $\mathcal{E}_w$ and $\mathcal{E}^{(x)}$ and are bounded above and below by positive multiples of each other, (5.26) also holds when m_w and $\mathcal{E}_w$ are replaced by m_x and $\mathcal{E}^{(x)}$, respectively. Hence, by Theorem (2.1), $\|e^{-t\bar{A}^{(x)}}\|_{1\to\infty} \le C/t^{1/2}$; and so $\|(\bar{A}^{(x)} + \lambda)^{-1}\|_{1\to\infty} \le C\int_0^\infty \frac{dt}{t}(e^{-\lambda t}t^{1/2}) = \lambda^{-1/2}C$. In particular, $G_\lambda^{(x)}(x,x) \le \lambda^{-1/2}C$, which means, in turn, that $G_\lambda(x,x) \le \lambda^{-1/2}C$. We are now finished with $\mathcal{E}^{(x)}$, and almost with the proof. By the Schwarz inequality, $G_\lambda(x,y) \le (G_\lambda(x,x)G_\lambda(y,y))^{1/2}$, and so $\|(\bar{A} + \lambda)^{-1}\|_{1\to\infty} \le \lambda^{-1/2}C$. Finally

$$\|f\|_2^2 = (f,(\bar{A} + \lambda)(\bar{A} + \lambda)^{-1}f) = \lambda(f,(\bar{A} + \lambda)^{-1}f) + (f,\bar{A}(\bar{A} + \lambda)^{-1}f)$$
$$\le \lambda^{-1/2}C\|f\|_1^2 + \lambda^{-1}\mathcal{E}(f,f) ,$$

and minimizinng in λ leads to $\|f\|^6_{L^2(m)} \le A\mathcal{E}(f,f)\|f\|^4_{L^1(m)}$ for some $A \in (0,\infty)$. Thus Theorem (4.1) gives us (5.21). Q.E.D.

Next we turn to off diagonal bounds and applications of section 3. The trick to applying the results of section 3 is to find, for given x, y, and t , a ψ which maximizes, or nearly maximizes, $\psi(x) - \psi(y) - t\Gamma(\psi)^2$. Hence, in situations where one can guess the correct behavior of the transition function -- and can therefore make a good choice for ψ -- Theorem (3.25) is a good source of pointwise bounds.

In our next example, E is the integers, and m is counting measure. Consider a random walk on the integers given as follows: Let $p:Z\times Z\to\mathbb{R}$ be a non-negative, symmetric function. Suppose that p is dominated by a non-negative even function $\tilde{p}:Z\to\mathbb{R}$ which possesses a moment generating function $M(\lambda)$. That is, suppose that for some $\epsilon > 0$ and some $B < \infty$,

(5.27) $$M(\lambda) = \sum_{n\in Z} e^{\lambda|n|}\tilde{p}(n) \le B < \infty \quad \text{for all } \lambda \in [0,\epsilon) .$$

Then in particular, if we write $\sigma^2 = (\frac{d}{d\lambda})^2 M(\lambda)|_{\lambda=0}$.

(5.28) $$\sum_{n\in Z} p(m,m+n)n^2 \le \sigma^2 \quad \text{for all } m \in Z .$$

It is easy to see that

(5.29) $$\mathcal{E}(f,f) = \frac{1}{2}\sum_{m,n\in Z}(f(m+n) - f(m))^2 p(m,m+n)$$

is the Dirichlet form corresponding to a uniquely determined family $P(t,m,\cdot)$ of probability transition functions with

(5.30) $$P(t,m,\{m+n\}) = p(m,m+n)t + o(t)$$

For this reason, $p(\cdot,\cdot)$ is called a jump rate function.

In general it is very difficult to pass from the infinitessimal description (5.30) of the transition function to a useful closed form formula for it. However, just as in section 2 with the truncated Cauchy processs, Central Limit Theorem considerations suggest that, at least when (5.28) is fairly sharp, and in the Gaussian space-time region where t is much larger than n, $P(t,m,\{m+n\})$ is very nearly $(2\pi\sigma^2 t)^{-1/2}e^{-n^2/2\sigma^2 t}$. We will now prove that there is in fact a pointwise upper bound of this form in the appropriate space time region.

First pick some large N and some $\alpha \neq 0$, and define the even function $\psi_{N,\alpha}$ by $\psi_{N,\alpha}(n) = \alpha N$ for $n \le N$, $\psi_{N,\alpha}(n) = 2\alpha N - \alpha n$ for $n \in [N,2N]$, and $\psi_{N,\alpha}(n) = 0$ for $n \ge 2N$. Clearly, $\psi_{N,\alpha} \in \hat{\mathcal{F}}_\infty$. Next observe that, writing ψ for $\psi_{N,\alpha}$,

$$\Gamma(e^{\psi},e^{\psi})(\{m\}) = \frac{1}{2}\sum_{n\in Z}{}^{\bullet}(e^{\psi(m)} - e^{\psi(m+n)})^2 p(m,m+n)$$
$$= e^{2\psi(m)}\left[\frac{1}{2}\sum_{n\in Z}(1 - e^{\psi(m+n)-\psi(m)})^2 p(m,m+n)\right]$$
$$\leq e^{2\psi(m)}\left[\frac{1}{2}\sum_{n\in Z}(1 - e^{\alpha|n|})^2\tilde{p}(n)\right]$$

Then, by Taylor's theorem, if $K = \partial_\lambda^3 M(\lambda)|_{\lambda=\epsilon/2}$:

(5.31) $$\Gamma(\psi_{N,\alpha})^2 \leq \frac{1}{2}\alpha^2\sigma^2 + \frac{1}{2}M'''(\epsilon/2)|\alpha|^3$$

whenever

(5.32) $$|\alpha| \leq \epsilon/4$$

To use this estimate in Theorem (3.25) we need to know that $\mathcal{E}$ satisfies a Nash type inequality. This will follow easily from a comparison argument if we impose

(5.33) $$p(n,n+1)\wedge p(n,n-1) \geq \mu > 0 \quad \text{for all } n .$$

(5.34) Theorem: Referring to the preceding, there is a $C \in (0,\infty)$, depending only on μ, such that for all ρ and δ from $(0,1)$

(5.35) $$P(t,m,\{n\}) \leq C(\rho t)^{-1/2}\exp[-(1-\delta)|n-m|^2/2(1+\rho)\sigma^2 t]$$

for all $(t,m,n) \in (0,\infty)\times Z\times Z$ satisfying

(5.36) $$t \geq \left[(K/\delta\sigma^4)\vee(4/\epsilon\sigma^2)\right]|n - m|$$

Proof: By the preceding,

$$D(t;m,n) \geq \alpha(m - n) - t(\tfrac{1}{2}\alpha^2\sigma^2 + \tfrac{K}{2}|\alpha|^3)$$

so long as $|\alpha| \leq \epsilon/4$. In particular, if $t \geq \frac{4}{\epsilon^2\sigma^2}|n - m|$, then we can take $\alpha = \frac{m - n}{\sigma^2 t}$ and thereby obtain

$$D(t;m,n) \geq \frac{|n - m|^2}{2\sigma^2 t}\left[1 - \frac{K|n - m|}{\sigma^4 t}\right].$$

Hence, if in addition, $t \geq \frac{K}{\delta\sigma^4}|n - m|$, then we get

$$D(t;m,n) \geq (1 - \delta)|n - m|^2/2(1 + \rho)\sigma^2 t.$$

At the same time, after comparing $\mathcal{E}$ to the Dirichlet form corresponding to the standard random walk on Z, one sees that

$$\|f\|_2^6 \le \frac{7}{\mu}\mathcal{E}(f,f)\|f\|_1^4.$$

Hence, by Theorem (3.25), we arrive at (5.35). Q.E.D.

Note that since ρ and δ are arbitrary elements of (0,1), we get close to what the Central Limit Theorem suggests is the best possible rate of Gaussian decay -- though of course the factors out front diverge as ρ tends to zero.

We give one final example of an interesting situation where we can give a good estimate for the quantity $D(T;x,y)$ defined in (3.26). Namely, consider the case when $E = \mathbb{R}^N$ equipped with Lebesgue measure. Let $\{V_1, \ldots, V_d\} \subseteq C_b^\infty(\mathbb{R}^N;\mathbb{R}^N)$ be a collection of vector fields on $\mathbb{R}^N$ and let $\mathcal{E}$ be the quadratic form on $L^2(\mathbb{R}^N)$ obtained by closing

$$(5.40) \qquad \mathcal{E}(\varphi,\varphi) = \sum_{k=1}^{d} \int_{\mathbb{R}^N} |V_k\varphi|^2 dx\ , \quad \varphi \in C_o^\infty(\mathbb{R}^N)$$

in $L^2(\mathbb{R}^N)$. Again applying standard results from [F], one sees that this closure exists and that the resulting $\mathcal{E}$ is the Dirichlet form associated with the unique transition probability function $P(t,x,\cdot)$ for which the corresponding Markov semigroup $\{P_t: t > 0\}$ satisfies $P_t\varphi = \varphi + \int_0^t P_s L\varphi ds$, $t > 0$, for all $\varphi \in C_o^\infty(\mathbb{R}^N)$, where $L = -\sum_{k=1}^{d} V_k^* V_k$ and we think of V_k as the directional derivative operator $\sum_{i=1}^{N} V_k^i \partial_{x_i}$. (By V_k^* we mean the formal adjoint of the V_k as a differential operator.) Set $a(x) = \sum_{k=1}^{d} V_k(x) \otimes V_k(x)$ and note that an equivalent expression for $L = \nabla \cdot (a\nabla)$. In particular, when $a(\cdot) \geq \epsilon I$ for some $\epsilon > 0$, it is well known that $P(t,x,dy) = p(t,x,y)dy$ where $(t,x,y) \in (0,\infty)\times\mathbb{R}^N\times\mathbb{R}^N \longrightarrow p(t,x,y)$ is a smooth function which is bounded above and below in terms of appropriate heat kernels (cf. [F-S] for a recent treatment of this sort of estimate). Moreover, it is known that, in this non-degenerate situation, $\lim_{t\downarrow 0} t\log(p(t,x,y)) = -d(x,y)^2/4$, where $d(x,y)$ denotes the Riemannian distance between x and y computed with respect to the metric determined by a on $\mathbb{R}^N$ (cf. [V]). These considerations make it clear that we should examine the relation between $d(x,y)$ and the quantity $D(T;x,y)$ introduced in section 3).

In order to make it possible to have our discussion cover cases in which a is allowed to degenerate, we begin by giving an alternate description of $d(x,y)$. Namely, define $\underline{H} = H_d$ to be the Hilbert space of $h \in C([0,\infty);\mathbb{R}^d)$ satisfying $h(0) = 0$ and $\underline{\|h\|_H} \equiv \|\dot{h}\|_{L^2([0,\infty);\mathbb{R}^d)} < \infty$ ($\dot{h} \equiv \partial_t h$). Given $h \in H$, let $\underline{Y^h(\cdot,x)} \in C([0,\infty);\mathbb{R}^N)$ be defined by $Y^h(t,x) = x + \sum_{k=1}^{d} \int_0^t h_k(s) V_k(Y^h(s,x))ds$, $t \geq 0$. Finally, define $\underline{d(x,y)} = \inf\{\|h\|_H : h \in H \text{ and } Y^h(1,x) = y\}$. It is then quite easy to show that, in the non-degenerate case, $d(x,y)$ is the Riemannian distance between x and y determined by the metric a. More generally, one can show that $d(x,y)$ depends on the V_k's only through a.

We next observe that, from (3.6):

$$\Gamma(\psi,\psi)(dy) = \left[\sum_{k=1}^{d} (V_k\psi)^2(y)\right]dy, \quad \psi \in C_0^\infty(\mathbb{R}^N).$$

In particular, $\Gamma(\psi)^2 = \|\sum_{k=1}^{d}(V_k\psi)^2\|_\infty$, $\psi \in C_0^\infty(\mathbb{R}^N)$; and so $D(T;x,y) \geq \sup\{|\psi(y) - \psi(x)| - T\Gamma(\psi)^2 : \psi \in C_0^\infty(\mathbb{R}^N)\}$. Hence,

$$D(T;x,y) \geq D(x,y)^2/4T, \tag{5.41}$$

where $\underline{D(x,y)}^2 \equiv 4\sup\{|\psi(y) - \psi(x)| - \Gamma(\psi)^2 : \psi \in C_0^\infty(\mathbb{R}^N)\} = \sup\{|\psi(y) - \psi(x)|^2 : \psi \in C_0^\infty(\mathbb{R}^N) \text{ and } \Gamma(\psi) \leq 1\}$. On the other hand, since, by Schwarz's inequality, $|\psi(Y^h(1,x) - \psi(x)| \leq \Gamma(\psi)\|h\|_H$, we see that:

$$D(x,y) \leq d(x,y). \tag{5.42}$$

In order to complete our program, we will show that the opposite inequality holds when $d(x,\cdot)$ is continuous at y.

To begin with, suppose that $a(\cdot) \geq \epsilon I$ for some $\epsilon > 0$. It is then easy to see that $d(x,y) \leq (1/\epsilon)|y - x|$. Next, for given $x^o, y^o \in \mathbb{R}^N$ and $\sigma > 0$, define $\psi_\sigma(y) = \eta \circ \left[\int \rho_\sigma(\xi) d(x^o, y-\xi) d\xi\right]$, where $\rho \in C_0^\infty(\mathbb{R}^N)^+$ with $\int \rho(\xi)d\xi = 1$, $\rho_\sigma(\xi) = \sigma^{-N}\rho(\xi/\sigma)$, and $\eta \in C_0^\infty(\mathbb{R}^1)^+$ has the properties that $\|\eta'\|_\infty \leq 1$ and $\eta(u) = u$ for $u \in [0, d(x^o,y^o) + 1]$. Since, for any $\theta \in S^{d-1}$, $|d(x^o, e^{tV_\theta}y) - d(x^o,y)| \leq d(e^{tV_\theta}y, y) \leq t$, where $V_\theta = \sum_{k=1}^{d} \theta_k V_k$, it is easy to see that $\Gamma(\psi_\sigma) \leq 1 + C\sigma$, $\sigma \in (0,1]$, for some $C \in (0,\infty)$. Hence, $D(x^o,y^o) \geq \varliminf_{\sigma\downarrow 0} |\psi_\sigma(y^o) - \eta_\sigma(x^o)| = d(x^o,y^o)$. In other words, when $a(\cdot) \geq \epsilon I$, equality holds in (5.42).

(<u>5.43</u>) <u>Lemma</u>: If $d(x,\cdot)$ is continuous at y, then $d(x,y) = D(x,y)$.

<u>Proof</u>: Given $\epsilon > 0$, define d_ϵ and D_ϵ relative to the vector fields $\{V_1,\ldots,V_d,\epsilon^{1/2}\partial_{x_1},\ldots,\epsilon^{1/2}\partial_{x_N}\}$. Then the corresponding $a_\epsilon = a + \epsilon I$; and so, by the preceding, $d_\epsilon = D_\epsilon$. In addition, it is clear that $D_\epsilon \le D$. Finally, for each $\epsilon > 0$, choose $h_\epsilon = (k_\epsilon,\ell_\epsilon) \in H_{d+N} = H_d \times H_N$ so that $Y^{h_\epsilon}(1,x) = y$ and $\|h_\epsilon\|_{H_{d+N}} = d_\epsilon(x,y)$, and let $y_\epsilon = Y^{k_\epsilon}(1,x)$ where $Y^{k_\epsilon}(\cdot,x) = Y^{(k_\epsilon,0)}(\cdot,x)$. Then, $d(x,y_\epsilon) \le \|k_\epsilon\|_{H_d} \le \|h_\epsilon\|_{H_{d+N}} = D_\epsilon(x,y) \le D(x,y)$. At the same time, since $\|\ell_\epsilon\|_{H_N} \le d(x,y)$, $y_\epsilon \longrightarrow y$ as $\epsilon \downarrow 0$; and so, by continuity, $d(x,y_\epsilon) \longrightarrow d(x,y)$. Q.E.D.

(<u>5.44</u>) <u>Remark</u>: The identification of d with D in the non-degenerate case was known to Davies [D]. In addition, Davies suggested that the two are the same in greater generatlity, but did not provide a proof.

(<u>5.45</u>) <u>Theorem</u>: Suppose that either $\mathcal{E}$ satisfies (1.2) or $\{P_t: t > 0\}$ satifies (1.3) for some $\nu \in (0,\infty)$, $\delta \in [0,1]$, and A or B from $(0,\infty)$. Then, $P(t,x,dy) = p(t,x,y)dy$ where $(t,x,y) \in (0,\infty)\times\mathbb{R}^N\times\mathbb{R}^N \longrightarrow p(t,x,y) \in [0,\infty)$ is measurable and satisfies

(5.46) $$p(t,x,y) \le (C_\rho e^{\delta t}/t^{\nu/2})\exp[-D(x,y)^2/4(1+\rho)t]$$

for all $(t,x) \in (0,\infty)\times\mathbb{R}^N$ and almost every $y \in \mathbb{R}^N$, where $C_\rho \in (0,\infty)$ depends only on ν, ρ, and A or B. In particular, if $d(x,\cdot)$ is continuous, then $D(x,y)$ in (5.46) can be replaced by $d(x,y)$.

(<u>5.47</u>) <u>Remark</u>: Using results of various authors about subelliptic operators, one can show that the preceding theorem applies to a large class of degenerate examples. For instance, if the vector fields $\{V_1,\ldots,V_d\}$ satisfy Hormander's condition in a sufficiently uniform way, then one can check not only that $\mathcal{E}$ satisfies (1.2) but also that the associated $p(t,x,y)$ is smooth and the corresponding $d(x,\cdot)$ is Holder continuous. A closer examination of this situation will be the topic of a forthcoming article [K-S], in which complementary lower bounds on $p(t,x,y)$ will be obtained when $t \in [1,\infty)$.

Bibliography:

[B-E] D. Bakry, M. Emery; Diffusions hypercontractives, pp. 177-207 in Sem. de Probabilites XIX, Lecture Notes in Mathematics #1123, eds. J. Azema and M. Yor, Springer-Verlag, New York, 1985

[D] E. B. Davies; Explicit constants for the Gaussian upper bounds on heat kernels, 1985 preprint, to appear.

[D-M] C. Dellacherie, P. A. Meyer; Probabilities and Potential B, North Holland, Amsterdam, 1982

[F-S] E. B. Fabes, D. W. Stroock; A new proof of Moser's parabolic Harnack inequality via the old ideas of Nash, 1985 preprint, to appear in Arch. for Rat. Mech. and Anal.

[F] M. Fukushima; Dirichlet Forms and Markov Processes, North Holland, Amsterdam, 1980

[S-K III] S. Kusuoka, D. Stroock; Applications of the Malliavin calculus, part III, 1986 preprint, to appear

[N] J. Nash; Continuity of solutions of parabolic and elliptic equations, Amer. J. Math., v. 80, (1958), pp. 931-954

[S] I. E. Segal; Construction of nonlinear local quantum processes: I, Ann. Math., v. 92, (1970), pp. 462-481

[St] D. W. Stroock; Applications of Fefferman-Stein type interpolation to probability and analysis, Comm. Pure Appl. Math., v. XXVI, (1973), pp. 477-496

[L.D.] D. W. Stroock; An Introduction to the Theory of Large Deviations, Springer-Verlag, New York, 1984

[V] S. R. S. Varadhan; On the behavior of the fundamental solution of the heat equation with variable coefficients, CAMS 20#2, pp. 431-455, 1967

[V-1] N. Th. Varopoulos; Isoperimetric inequalities and Markov chains, J. Func. Anal., v. 63, (1985), pp. 215-239

[V-2] N. Th. Varopoulos; Hardy-Littlewood theory for semigroups, J. Func. Anal., v. 63, (1985), pp. 240-260