UPPER BOUNDS OF DERIVATIVES OF THE HEAT KERNEL ON AN ARBITRARY COMPLETE MANIFOLD

Alexander Grigor'yan

0. Introduction

In the present paper, we offer a new method of estimating of the heat kernel derivatives on Riemannian manifolds. Let M denote a smooth connected complete non-compact Riemannian manifold and $p(x,y,t)$ be the heat kernel i.e. the smallest positive fundamental solution to the heat equation

$$\frac{\partial u}{\partial t} - \Delta u = 0$$

Δ being the Laplace operator induced by the Riemannian metric. It is standard now that for a wide variety of Riemannian manifolds the heat kernel can be estimated as follows

$$p(x,y,t) \le \varphi(t) \exp\left(-\frac{r^2}{Ct}\right) \tag{0.1}$$

where $\varphi(t)$ is a increasing function, r is the geodesic distance between x,y and C is a constant.

There are different methods of obtaining inequalities of such a kind starting from some geometric properties of the manifold. We shall be concerned with the following approach offered by the author [6] . Let us consider the weighted integral of the heat kernel

$$E(x,t) = \int_M p^2(x,y,t) \exp\left(\frac{r^2}{Dt}\right) dy$$

where $D > 2$ is given, then as was shown in [6] the function $E(x,t)$ is finite (provided $D > 2$) and decreasing in t . Besides, the semigroup property of the heat equation implies the following universal inequality:

$$p(x,z,t) \le \sqrt{E(x,\frac{t}{2})E(z,\frac{t}{2})} \exp\left(-\frac{r^2}{2Dt}\right) \tag{0.2}$$

where $r = \operatorname{dist}(x,z)$. For example, if one knows that for all x

$$E(x,t) \le \frac{1}{f(t)} \tag{0.3}$$

then (0.2) implies (0.1) with $\varphi(t) = f(t/2)^{-1}$ and $C = 2D$.

Of course, the inequality (0.3) (as well as (0.1)) can be true only under certain geometric conditions. A geometric criterion of (0.3) in terms of an isoperimetric property of a manifold is proved in [6] (see also Section 3 in the present paper).

Grigor'yan A., Upper bounds of derivatives of the heat kernel on an arbitrary complete manifold, *J. Funct. Anal.* , **127** no.2, (1995) 363-389.

In this paper, we start from the point that $E(x,t)$ is a known quantity and will show how to prove analogous estimates of the heat kernel derivatives. Let us set

$$E_1(x,t) = \int_M |\nabla p|^2 (x,y,t) \exp\left(\frac{r^2}{Dt} \right) dy$$

and

$$E_2(x,t) = \int_M |\Delta p|^2 (x,y,t) \exp\left(\frac{r^2}{Dt} \right) dy$$

where the both operators $\nabla,\ \Delta$ relate to y .

Our main result states, in particular, that whenever (0.3) is known to be true for a point x and for all $t > 0$ then the following inequalities hold as well:

$$E_1(x,t) \le \frac{\text{const}}{f_1(t)} \tag{0.4}$$

and

$$E_2(x,t) \le \frac{\text{const}}{f_2(t)} \tag{0.5}$$

where

$$f_1(t) \equiv \int_0^t f(\tau)d\tau \ , \quad f_2(t) \equiv \int_0^t f_1(\tau)d\tau \ .$$

Moreover, similar to (0.2) one can estimate pointwise the time derivative of the heat kernel via the quantities E_i :

$$\left| \frac{\partial p}{\partial t} \right| (x,z,t) \le \sqrt{E_2(x,\frac{t}{2})E(z,\frac{t}{2})} \exp\left(-\frac{r^2}{2Dt} \right) \ . \tag{0.6}$$

The quantity E_2 enters this estimate because of the fact that $\frac{\partial p}{\partial t} = \Delta p$. Corresponding inequalities for the higher derivatives are found in Section 1 below where the basic result are proved.

In the Section 2, we get some consequences for heat kernel estimations without the Gaussian term. The main result there states that the inequality

$$p(x,x,t) \le \frac{1}{f(t)} \tag{0.7}$$

implies a similar estimate of the derivatives $\frac{\partial^n p}{\partial t^n}(x,y,t)$ which takes the following simplest form provided the function $f(t)$ is regular in some sense and grows at least polynomially:

$$\left| \frac{\partial^n p}{\partial t^n} \right| (x,y,t) \le \frac{\text{const}}{f_n(t)} \ . \tag{0.8}$$

Moreover, in this case $\left| \frac{d}{dt} \frac{1}{f_n} \right|$ is finitely proportional to $\frac{1}{f_{n+1}}$ which means that the dependence of t in the upper bound (0.8) is sharp.

If the function $f(t)$ satisfies some more regularity assumptions then one can deduce from (0.7) the Gaussian upper bound of $\frac{\partial^n p}{\partial t^n}(x,y,t)$ similar to (0.7) which is done in Section

3. In that Section, we give also some examples of linking the previous results of [6] on estimation of the heat kernel itself with that of the present paper to make clear a geometric background of all these considerations. For instance, we present upper bounds of the time derivatives on some classes of manifolds.

In this connection, we would like to emphasize the following advantages of our approach. First, it gives a sharp dependence in t of the time derivatives as was mentioned above. Second, it gives a sharp Gaussian exponent $2D$ in (0.6) which can be taken arbitrarily close to 4 or even equal to 4 if one pays for it by appearing of a polynomial correction term. Third, it can be localized in the sense that in order to get an upper bound of $\frac{\partial^n p}{\partial t^n}(x,z,t)$ at given points x, z, we need only to know $E(x,t)$ and $E(z,t)$ at the same points - we can further estimate $E_2(x,t)$ by (0.5) and get a desired pointwise estimate by (0.6) .

There are other approaches of estimation of derivatives of the heat kernel which can be found in the following works:

- for parabolic equations in $\mathbb{R}^n$: Porper [10] (see also [11]) and Kovalenko, Semenov [8] ;
- in context of Lie groups: Varopoulos [13] as well as the books by Varopoulos, Saloff-Coste, Coulhon [14] and by Robinson [12] ;
- on Riemannian manifolds: Cheng, Li, Yau [2] , Varopoulos [13] , Davies [3] .

The sharpest method was developed by Davies [3] who exploited crucially the fact that the heat semigroup is a holomorphic one and, hence, the heat kernel can be considered for complex values of the time. This enables one to use the Cauchy integral formula to represent the time derivative directly via the function in question. This method uses essentially an analytic form of the function $f(t)$ which governs a decay of the heat kernel and although it works well for a polynomial function $f(t) = t^\nu$ it is not clear whether it goes through, say, for the function $f(t) = \exp t^\nu$. Neither it is evident that it is possible to localize this machinery in the above sense.

This is a common drawback of all known before approaches that they yield the upper bound

$$\left|\frac{\partial p}{\partial t}\right| \leq \frac{\text{const}}{tf((1-\epsilon)t)}$$

whereas our estimate (0.8) gives

$$\left|\frac{\partial p}{\partial t}\right| \leq \frac{\text{const}}{\int_0^t f(t)dt}$$

which is essentially sharper whenever $f(t)$ increases superpolynomially.

As far as the space derivative ∇p is concerned a pointwise upper bound was obtained by B.Davies [3] from that of the time derivative by means of the Harnack inequality of P.Li and S.T.Yau [9] which needs in turn boundedness from below of the Ricci curvature. See also [2] concerning the higher order space derivatives.

Here we hardly touch pointwise estimates of space derivatives because it seems not to be possible to get such estimates at the same level of generality as that of time derivatives. Indeed, it is well-known that the magnitude of spatial derivatives of a solution of a parabolic or elliptic equation can not be estimated only through the *sup* -norm of the coefficients (which are expressed in our case through the Riemannian metric) - some extra information concerning the modulus of continuity of the coefficients must be involved. At the same time, the upper bounds of the heat kernel itself such as (0.1) or (0.7) do not provide such information.

Nevertheless, if one knows a priori that two possible gradients $\nabla_x p(x,y,t)$ and $\nabla_y p(x,y,t)$ have finite proportional moduli then it becomes possible to deduce a pointwise upper bound from the integral estimate (0.4) .

1. Basic results

Here and in what follows we shall keep to the following convention: ∇^n denotes $\Delta^{n/2}$ if n is even (for example, $\Delta = \nabla^2$) and $\nabla\Delta^{\frac{n-1}{2}}$ if n is odd. Therefore, in the former case ∇^n is a scalar operator whereas in the latter case ∇^n is a vector. The operator ∇ relates always to a variable y unless otherwise specified.

Let us fix an arbitrary number $D > 2$ and introduce a notation

$$E_n(x,t) = \int_M |\nabla^n p|^2 (x,y,t) \exp\left(\frac{r^2}{Dt}\right) dy$$

where $n \geq 0$ is an integer, $x \in M,\ t > 0$ are arbitrary.

Let us consider also a pre-compact region $\Omega \subset M$ with a smooth boundary and denote by $p_\Omega(x,y,t)$ the heat kernel in Ω with the vanishing Dirichlet boundary values. It is known (see [4] , [1]) that p_Ω increases on expansion of Ω and approaches to p as $\Omega \to M$ where $\Omega \to M$ is an exhaustion of M by a sequence of regions Ω . Moreover, the convergence $p_\Omega(x,y,t) \to p(x,y,t)$ is uniform on any compact in $M \times M \times (0,+\infty)$ and the derivatives of any order converge locally uniformly, too.

We shall refer to p_Ω as a local heat kernel in contrast to a global one p and introduce the corresponding quantities E_n^Ω for a local heat kernel p_Ω too:

$$E_n^\Omega(x,t) = \int_\Omega |\nabla^n p_\Omega|^2 (x,y,t) \exp\left(\frac{r^2}{Dt}\right) dy\ .$$

Normally the value of $D > 2$ will be fixed but if we need to consider the quantities E_n for different values of D we shall use extended notations $E_{n,D}(x,t)$ and, respectively, $E_{n,D}^\Omega(x,t)$.

Let us fix also some $T \in (0,+\infty]$ which is allowed to be equal to $+\infty$ and relate to any positive integrable function $f(t)$ on $(0,T)$ a sequence of its integrals:

$$f_0 = f, \qquad f_n(t) = \int_0^t f_{n-1}(\tau)d\tau, \quad n \geq 1$$

or, equivalently, for any $n \geq 1$ we have

$$f_n(t) = \int_0^t \frac{(t-\tau)^{n-1}}{(n-1)!} f(\tau)d\tau\ . \tag{1.1}$$

Our main result states the following.

Theorem 1.1 *Suppose that for some $x \in M$, for some constant C and for all $t \in (0,T)$*

$$E_0(x,t) \leq C/f(t) \ , \tag{1.2}$$

$f(t)$ being a positive integrable function in $(0,T)$ then it follows that for any $n \geq 1$ and for all $t \in (0,T)$

$$E_n(x,t) \leq \frac{Cc^{-n}}{f_n(t)} \tag{1.3}$$

where

$$c = \frac{D-2}{D/2+8} \ . \tag{1.4}$$

Proof. It suffices to treat the case $C = 1$ which will be assumed to hold in the course of the proof. Let us consider a pre-compact region $\Omega \subset M$ with a smooth boundary and note that $p_\Omega \leq p$ implies that the hypothesis (1.2) holds also if we replace E_0 by E_0^Ω . The main part of the proof will consist of obtaining the inequality (1.3) for E_n^Ω . The rest will follow upon passing to a limit from the observation that

$$E_n(x,t) \leq \liminf_{\Omega \to M} E_n^\Omega(x,t) \ . \tag{1.5}$$

Indeed, for any compact $K \subset M$ containing the point x and for any Ω covering K we have

$$\int_K |\nabla^n p_\Omega|^2 (x,y,t) \exp\left(\frac{r^2}{Dt}\right) \, dy \leq E_n^\Omega(x,t)$$

whence letting $\Omega \to M$ we get

$$\int_K |\nabla^n p|^2 (x,y,t) \exp\left(\frac{r^2}{Dt}\right) \, dy \leq \liminf_{\Omega \to M} E_n^\Omega(x,t)$$

which implies (1.5) due to the arbitrariness of the compact K .

From now on we shall be found inside a given region Ω and for the sake of simplicity we are going to modify the notations by suppressing the subscript Ω by p_Ω and E_n^Ω . In other words, in the course of the proof we shall denote the local heat kernel as p and the quantities E_n are supposed to relate to the local heat kernel as well.

We shall be considering expressions of the kind $\nabla^n u \nabla^m v$ which is understood either as the inner product of two vectors or as the product of a vector and a scalar or as the product of two scalars depending on whether the integers n, m are even or odd. The following lemma is nothing other than integration by parts in our notations.

Lemma 1.1 *If u, v, w are smooth enough functions in a pre-compact region Ω with a smooth boundary and $n > 0, m \geq 0$ are integers being simultaneously even or odd then*

$$\begin{aligned} &\int_\Omega \nabla^n u \nabla^m v w \\ &= -\int_\Omega \nabla^{n-1} u \nabla^{m+1} v w - \int_\Omega \nabla^{n-1} u \nabla^m v \nabla w + \int_{\partial\Omega} \nabla^{n-1} u \nu \nabla^m v w \end{aligned} \tag{1.6}$$

where ν is the unit outer normal vector field on $\partial\Omega$.

Proof. Let first n, m be both even, then denoting for simplicity $\hat{u} = \Delta^{\frac{n}{2}-1}u$ and $\hat{v} = \Delta^{m/2}v$ we have

$$\int_\Omega \nabla^n u \nabla^m v w = \int_\Omega \Delta\hat{u}\hat{v}w = -\int_\Omega \nabla\hat{u}\nabla(\hat{v}w) + \int_{\partial\Omega} \nabla\hat{u}\nu\hat{v}w$$
$$= -\int_\Omega \nabla\hat{u}\nabla\hat{v}w - \int_\Omega \nabla\hat{u}\hat{v}\nabla w + \int_{\partial\Omega} \nabla\hat{u}\nu\hat{v}w$$

which coincides with the right-hand side of (1.6) .

If n, m are odd then we put $\hat{u} = \Delta^{\frac{n-1}{2}}u$ and $\hat{v} = \Delta^{\frac{m-1}{2}}v$ and obtain similarly

$$\int_\Omega \nabla^n u \nabla^m v w = \int_\Omega \nabla\hat{u}\nabla\hat{v}w = -\int_\Omega \hat{u}\operatorname{div}(\nabla\hat{v}w) + \int_{\partial\Omega} \hat{u}\nu\nabla\hat{v}w$$
$$= -\int_\Omega \hat{u}\Delta\hat{v}w - \int_\Omega \hat{u}\nabla\hat{v}\nabla w + \int_{\partial\Omega} \hat{u}\nu\nabla\hat{v}w$$

which was to be proved.

To proceed with the proof of the theorem let us put

$$\xi(y,t) = \frac{r^2}{Dt}$$

where $r = \operatorname{dist}(x,y)$ as above. In fact, we shall use only the the following property of the function ξ :

$$\xi_t + \frac{D}{4}\left|\nabla\xi\right|^2 \leq 0 \tag{1.7}$$

which follows from $|\nabla r| \leq 1$.

Let us start with integration by parts of the integral E_n using the above lemma:

$$\begin{aligned} E_n &= \int_\Omega \left|\nabla^n p\right|^2(x,y,t)e^\xi dy \\ &= -\int_\Omega \nabla^{n-1}p\nabla^{n+1}pe^\xi - \int_\Omega \nabla^{n-1}p\nabla^n p\nabla\xi e^\xi + \int_{\partial\Omega} \nu\nabla^{n-1}p\nabla^n pe^\xi \end{aligned} \tag{1.8}$$

Since either $n-1$ or n is even then one of the factors $\nabla^{n-1}p,\ \nabla^n p$ is equal to an integer power of Laplacian $\Delta^k p,\ k = [n/2]$ which vanishes on $\partial\Omega$ because $\Delta^k p = \frac{\partial^k p}{\partial t^k} = 0$ as $p|_{\partial\Omega} = 0$. Therefore, the third term on the right-hand side of (1.8) is equal to 0 . The first term is estimated from above by the product

$$E_{n-1}^{\frac{1}{2}}E_{n+1}^{\frac{1}{2}}$$

the second is less than or equal to

$$E_{n-1}^{\frac{1}{2}}\hat{E}_n^{\frac{1}{2}}$$

where we have introduced the notation

$$\hat{E}_n(x,t) = \int_\Omega \left|\nabla^n p\right|^2\left|\nabla\xi\right|^2 e^\xi \ .$$

Hence, we have

$$E_n \leq E_{n-1}^{\frac{1}{2}}\left(E_{n+1}^{\frac{1}{2}} + \hat{E}_n^{\frac{1}{2}}\right) \ . \tag{1.9}$$

Next we consider the time derivative $E_n^{'}$:

$$E_n^{'} = 2\int_\Omega \nabla^n p_t \nabla^n p e^\xi + \int_\Omega |\nabla^n p|^2 \xi_t e^\xi \ .$$

Replacing p_t by $\Delta p = \nabla^2 p$ and ξ_t according to (1.7) by a larger value $-\frac{D}{4}|\nabla\xi|^2$ we get

$$E_n^{'} \leq 2\int_\Omega \nabla^{n+2} p \nabla^n p e^\xi - \frac{D}{4}\int_\Omega |\nabla^n p|^2 |\nabla\xi|^2 e^\xi \ .$$

Integration by parts yields

$$E_n^{'} \leq -2\int_\Omega \left|\nabla^{n+1} p\right|^2 e^\xi - 2\int_\Omega \nabla^{n+1} p \nabla^n p \nabla\xi e^\xi - \frac{D}{4}\hat{E}_n \ ,$$

$$E_n^{'} \leq -2E_{n+1} + 2E_{n+1}^{\frac{1}{2}}\hat{E}_n^{\frac{1}{2}} - \frac{D}{4}\hat{E}_n \ . \tag{1.10}$$

Now we combine the inequalities (1.9) and (1.10) : for any positive number c we obtain

$$E_n^{'} + c\frac{E_n^2}{E_{n-1}} \leq -2E_{n+1} + 2E_{n+1}^{\frac{1}{2}}\hat{E}_n^{\frac{1}{2}} - \frac{D}{4}\hat{E}_n + c\Big(E_{n+1}^{\frac{1}{2}} + \hat{E}_n^{\frac{1}{2}}\Big)^2$$

whence it follows after simplifying that

$$E_n^{'} + c\frac{E_n^2}{E_{n-1}} \leq -\Big((2-c)E_{n+1} - 2(1+c)E_{n+1}^{\frac{1}{2}}\hat{E}_n^{\frac{1}{2}} + (\frac{D}{4}-c)\hat{E}_n\Big) \ . \tag{1.11}$$

Let us note that the expression in brackets appearing on the right-hand side of (1.11) is a quadratic polynomial with respect to $E_{n+1}^{\frac{1}{2}}$ and $\hat{E}_n^{\frac{1}{2}}$ which means that it is always non-negative provided its discriminant is non-positive i.e.

$$(1+c)^2 - (2-c)\Big(\frac{D}{4} - c\Big) \leq 0$$

and

$$2 - c > 0, \qquad \frac{D}{4} - c > 0$$

which is equivalent to

$$c \leq \frac{D-2}{D/2+8} \ .$$

We take for the further considerations the largest possible value of c which coincides with one given by (1.3) . It follows from (1.11) that

$$E_n^{'} \leq -c\frac{E_n^2}{E_{n-1}} \ . \tag{1.12}$$

We are left to integrate by induction this ordinary differential inequality. Indeed, for $n = 0$ (1.3) is true by the hypothesis of the theorem. Suppose, that (1.3) has been proved for the order $n-1$ i.e.

$$E_{n-1} \leq \frac{c^{1-n}}{f_{n-1}(t)} \tag{1.13}$$

Dividing (1.12) by E_n^2 and integrating over the interval $(0,t)$ we have

$$E_n(x,t) \le c^{-1}\left(\int_0^t \frac{d\tau}{E_{n-1}(x,\tau)}\right)^{-1} . \tag{1.14}$$

Applying (1.13) we get

$$\int_0^t \frac{d\tau}{E_{n-1}(x,\tau)} \ge c^{n-1}\int_0^t f_{n-1}(\tau)d\tau = c^{n-1}f_n(t) .$$

Substituting this into (1.14) we complete the proof. □

Let us observe that we may treat the infinite value of D as well. In this case $\xi(x,t) = 0$ and all computations carried out in the course of the proof of Theorem 1.1 become much easier. For instance, we have $\hat{E}_n = 0$ and the expression on the right-hand side of (1.11) consists only of the first summand which enables us to take $c = 2$. Therefore, the conclusion of Theorem 1.1 holds with $c = 2$ which formally can be seen also from (1.4) upon letting $D \to \infty$ there.

Corollary 1.1 *We have on any manifold M*

$$\lim_{\Omega\to M} E_n^\Omega(x,t) = E_n(x,t) .$$

Indeed, let us estimate for a given x the difference

$$\left|E_{n,D}^\Omega(x,t) - E_{n,D}(x,t)\right| . \tag{1.15}$$

We use here the extended notation emphasizing the dependence on D because we are going to consider simultaneously two different values of this parameter. The difference (1.15) as a difference between two integrals can be split into a sum of the differences of the corresponding integrals over a ball $B(x,R)$ and over its exterior. The first summand goes to 0 as $\Omega \to M$ because the derivatives of the local heat kernel p_Ω converge to that of the global heat kernel p uniformly in any compact, in particular, in the ball $B(x,R)$.

To estimate the second summand it suffices to show that either integral over the exterior of the ball tends to 0 uniformly in Ω as $R \to \infty$. To this end we apply the estimate of Theorem 1.1 to $E_{n,D-\varepsilon}$ where ε is a small positive number such that $D - \varepsilon > 2$. Let us put

$$f(t) \equiv \frac{1}{E_{0,D-\varepsilon}(x,t)} . \tag{1.16}$$

Since $p_\Omega \le p$ it follows that $E_0^\Omega \le E_0$ and, thereby

$$E_{0,D-\varepsilon}^\Omega(x,t) \le \frac{1}{f(t)} .$$

By Theorem 1.1 we have the inequalities

$$E_{n,D-\varepsilon}(x,t) \le \frac{\text{const}}{f_n(t)} , \quad E_{n,D-\varepsilon}^\Omega(x,t) \le \frac{\text{const}}{f_n(t)} \tag{1.17}$$

where the latter is formally not stated by this theorem but was shown to be valid in the course of its proof. Taking into account that

$$\frac{r^2}{(D-\varepsilon)t} = \frac{r^2}{Dt} + \frac{\varepsilon r^2}{D(D-\varepsilon)t}$$

we deduce from (1.17)

$$\int_{M\setminus B(x,R)} |\nabla^n p|^2 (x,y,t) \exp\left(\frac{r^2}{Dt}\right) dy \le \frac{\text{const}}{f_n(t)} \exp\left(-\frac{\varepsilon R^2}{D(D-\varepsilon)t}\right)$$

and the similar estimate for p_Ω which imply that the either integral outside $B(x,R)$ goes to 0 as $R \to \infty$ uniformly with respect to Ω . □

In the course of the proof we used a regularity of the function $E_0(x,t)$ with respect to t because the function $f(t)$ defined by (1.16) must be positive and integrable. Finiteness of E_0 provided $D > 2$ is a general fact established in [6] and it implies non-vanishing of $f(t)$. To show integrability of $f(t)$ we can prove first the statement of Corollary 1.1 for $n = 0$ (which needs only finiteness of E_0 and does not require Theorem 1.1) and obtain as a consequence that $E_0(x,t)$ is a decreasing function of t because E_0^Ω is such a function (see [7]). Hence, the function $f(t)$ is increasing and thereby integrable. Therefore, we may repeat again the proof of Corollary 1.1 and on this occasion for any $n \ge 0$.

The proof yields some additional information concerning E_n , for example, this function is always finite provided $D > 2$. Moreover, tracing back the proof it is easy to see that the convergence $E_n^\Omega(x,t) \to E_n(x,t)$ is uniform as long as t lies in a finite interval bounded away from 0 which implies that the function $E_n(x,t)$ is continuous in t , indeed. Finally, let us observe that the inequality (1.12) implies that the function $E_n^\Omega(x,t)$ decreases in t whence the same property of $E_n(x,t)$ follows.

We summarize this discussion of general properties of E_n as a function of time in the following statement.

Corollary 1.2 *The function $E_{n,D}(x,t)$ is a finite continuous decreasing function of $t \in (0,+\infty)$ for any $n \ge 0$ and any $D > 2$.*

Corollary 1.3 *We have for all $x \in M, t > 0$*

$$E_n(x,t) \le \frac{n!2^n c^{-n}}{t^n} E_0(x, \frac{t}{2}) \ . \tag{1.18}$$

An analogous estimate for the integrals of the heat kernel and its derivatives without the weight $\exp\left(\frac{r^2}{Dt}\right)$ (or, in other words, if $D = \infty$) was proved in [2] (Lemma 7).

To prove (1.18) let us put $f(t) = \frac{1}{E_0(x,t)}$ and apply Theorem 1.1 estimating $f_n(t)$ in the following manner:

$$f_n(t) \ge \frac{t^n}{n!2^n} f(\frac{t}{2})$$

which follows immediately from (1.1) upon reducing the interval of integration to $(t/2, t)$.

Let us proceed to pointwise estimates of the derivatives of the heat kernel.

Theorem 1.2 *For any two points $x, z \in M$ and for all $t > 0$*

$$\left|\frac{\partial^n p}{\partial t^n}\right|(x,z,t) \le \sqrt{E_{2n}(x,\frac{t}{2})E_0(z,\frac{t}{2})}\exp\left(-\frac{r^2}{2Dt}\right) \quad . \tag{1.19}$$

where $n \ge 0$ is an integer and $r = \mathrm{dist}(x,z)$.

Remark. Since the variables x, z enter the statement of Theorem 1.2 likewise a symmetric estimate holds as well

$$\left|\frac{\partial^n p}{\partial t^n}\right|(x,z,t) \le \sqrt{E_0(x,\frac{t}{2})E_{2n}(z,\frac{t}{2})}\exp\left(-\frac{r^2}{2Dt}\right) \tag{1.20}$$

Proof. By the semigroup property we have for any $s \in (0,t)$

$$p(x,z,t) = \int_M p(x,y,t-s)p(y,z,s)dy \tag{1.21}$$

which implies upon differentiation in t

$$\begin{aligned}\frac{\partial^n p}{\partial t^n}(x,z,t) &= \int_M \frac{\partial^n p}{\partial t^n}(x,y,t-s)p(y,z,s)dy \\ &= \int_M \Delta_y^n p(x,y,t-s)p(y,z,s)dy \ .\end{aligned}$$

or, taking here $s = t/2$,

$$\frac{\partial^n p}{\partial t^n}(x,z,t) = \int_M \Delta_y^n p(x,y,\frac{t}{2})p(y,z,\frac{t}{2})dy \ . \tag{1.22}$$

Let us put $r_1 = \mathrm{dist}(x,y),\ r_2 = \mathrm{dist}(z,y)$ and note that by a triangle inequality

$$r_1^2 + r_2^2 \ge \frac{1}{2}(r_1+r_2)^2 \ge \frac{1}{2}r^2$$

and, thereby,

$$1 \le \exp\left(-\frac{r^2}{2Dt}\right)\exp\left(\frac{r_1^2}{Dt}\right)\exp\left(\frac{r_2^2}{Dt}\right)$$

which together with (1.22) yields

$$\left|\frac{\partial^n p}{\partial t^n}\right|(x,z,t) \le \exp\left(-\frac{r^2}{2Dt}\right)\int_M \left|\Delta_y^n p\right|(x,y,\frac{t}{2})\exp\left(\frac{r_1^2}{Dt}\right) p(y,z,\frac{t}{2})\exp\left(\frac{r_2^2}{Dt}\right) dy$$

whence (1.19) follows upon application of Cauchy-Schwarz inequality. □

Similar arguments go through for estimation of the spatial derivative ∇p but with a very essential distinction: if we apply the gradient ∇_x to the semigroup identity (1.21) and use Cauchy-Schwarz inequality we are faced with a necessity to do with the integral of $\nabla_x p(x,y,t)$ instead of $\nabla_y p(x,y,t)$ which enters $E_1(x,t)$.

This obstacle can be overcome if it is known a priori that $\nabla_x p(x,y,t)$ and $\nabla_y p(x,y,t)$ are finitely proportional but generally speaking this is not so. More precisely, the following statement is true.

Theorem 1.3 *Let for some point $x \in M$ and for all $y \in M,\ t > 0$ the following inequality hold:*

$$\left|\nabla_x p\right|(x,y,t) \leq \mathrm{const}\left|\nabla_y p\right|(x,y,t)\ ,$$

then for any $z \in M$ and all $t > 0$ we have

$$\left|\nabla_x p\right|(x,z,t) \leq \mathrm{const}\sqrt{E_1(x,\frac{t}{2})E_0(z,\frac{t}{2})}\exp\left(-\frac{r^2}{2Dt}\right)\ .$$

2. Pointwise estimates without the Gaussian factor

In this Section we obtain some consequences of the foregoing theorems for the case $D = \infty$ which means that

$$E_n(x,t) = \int_M \left|\nabla^n p\right|^2(x,y,t)dy$$

and corresponds to the heat kernel estimates without the Gaussian correction term. As was explained above in this case Theorem 1.1 holds with $c = 2$.

Corollary 2.1 *Let us assume that for some points $x, z \in M$ and all $t \in (0,T)$*

$$p(x,x,t) \leq \frac{1}{f(t)},\quad p(z,z,t) \leq \frac{1}{g(t)}$$

where $f(t), g(t)$ are positive continuous functions on $(0,T)$ then it follows that for any integer $n \geq 0$ and for all $t \in (0,T)$

$$\left|\frac{\partial^n p}{\partial t^n}\right|(x,z,t) \leq \frac{1}{\sqrt{f_{2n}(t)g(t)}}\ . \tag{2.1}$$

In particular,

$$\left|\frac{\partial^n p}{\partial t^n}\right|(x,x,t) \leq \frac{1}{\sqrt{f_{2n}(t)f(t)}}\ . \tag{2.2}$$

Remark. A strengthened version of this statement is presented in the Section 3 below and includes the Gaussian correction term albeit under some additional restriction on the functions f, g .

Proof. Indeed, by the semigroup property of the heat kernel we have

$$E_0(x,t) = \int_M p^2(x,y,t)dy = p(x,x,2t)\ .$$

Therefore,

$$E_0(x,t) \leq \frac{1}{f(2t)}$$

and by Theorem 1.1 for any integer k

$$E_k(x,t) \leq \frac{c^{-k}}{2^{-k}f_k(2t)}$$

where 2^{-k} comes from integration k times of the function $f(2t)$. Applying Theorem 1.2 and observing that $c = 2$ we get

$$\left|\frac{\partial^n p}{\partial t^n}\right|(x,z,t) \leq \sqrt{E_{2n}(x,\frac{t}{2})E_0(z,\frac{t}{2})} \leq \frac{1}{\sqrt{f_{2n}(t)g(t)}}$$

which was to be proved. □

Corollary 2.2 *The functions $E_n(x,t)$ and $\log E_n(x,t)$ are convex in t for any integer $n \geq 0$ and for any $x \in M$.*

In view of Corollary 1.1 we may replace in the present statement the global heat kernel by a local one in a region Ω as was done in the proof of Theorem 1.1. Differentiating E_n with respect to t we have

$$\begin{aligned} E_n^{'} &= 2\int_\Omega \nabla^n p_t \nabla^n p = 2\int_\Omega \nabla^{n+2} p \nabla^n p \\ &= -2\int_\Omega \left|\nabla^{n+1} p\right|^2 = -2E_{n+1} \end{aligned} \tag{2.3}$$

Next we apply the inequality (1.12) proved in the course of the proof of Theorem 1.1 with $c = 2$:

$$E_{n+1}^{'} \leq -2\frac{E_{n+1}^2}{E_n}$$

and replace here E_{n+1} by $-\frac{1}{2}E_n^{'}$ according to the identity (2.3) . We get

$$E_n^{''} \geq \frac{E_n^{'\,2}}{E_n}$$

which is equivalent to $(\log E_n)^{''} \geq 0$. □

Corollary 2.3 *The function $\log p(x,x,t)$ is convex in t .*

Indeed, it follows immediately from the foregoing statement because

$$p(x,x,t) = \int_M p^2(x,y,t/2)dy = E_0(x,t/2) \ .$$

Returning to Corollary 2.1 we are going to show to what extent the statement is sharp. We would like to have a right rate of decay of the derivatives of the heat kernel in a time variable. Of course, if one is given the inequality

$$p(x,x,t) \leq \frac{1}{f(t)}$$

then one may conjecture that

$$\left|\frac{\partial p}{\partial t}\right|(x,x,t) \leq \mathrm{const}\left|\frac{d}{dt}\frac{1}{f(t)}\right| \tag{2.4}$$

at least if $f(t)$ is regular enough. By Corollary 2.1 we have on the right-hand side of the inequality (2.2) the second integral $f_2(t)$ of $f(t)$ instead of expected derivative. It turned out that for a wide class of functions $f(t)$ it yields basically the same estimate as (2.4) as will be established below. But first let us note that, for example, if $f(t) = \log t$ (for values of t bounded away from 0) then (2.2) acquires for $n = 1$ the form

$$\left|\frac{\partial p}{\partial t}\right|(x,x,t) \leq \frac{\mathrm{const}}{t\log t}$$

whereas the conjecture (2.4) would give a better bound:

$$\left|\frac{\partial p}{\partial t}\right|(x,x,t) \leq \frac{\mathrm{const}}{t\log^2 t} \ .$$

It is not clear whether the latter one is actually true. A gap between Corollary 2.1 and the conjecture (2.4) occurs whenever the function $f(t)$ increases as $t \to \infty$ slower as polynomially. On the contrary, if the function $f(t)$ grows as $t \to \infty$ at least as fast as a polynomial then the Corollary 2.1 gives a sharp bound which is shown by the following statement.

Corollary 2.4 *Let a function $f(t) \in C^1(0,T)$ satisfy the following assumption:*

(*) for some $\kappa \in (0,1)$ we have $\kappa \leq \frac{f' f_1}{f^2} \leq 1$ for all $t \in (0,T)$

Suppose, that for some points $x, z \in M$ and all $t \in (0,T)$

$$p(x,x,t) \leq \frac{1}{f(t)}, \quad p(z,z,t) \leq \frac{1}{f(t)} \tag{2.5}$$

then for any integer $n \geq 0$

$$\left|\frac{\partial^n p}{\partial t^n}\right|(x,z,t) \leq \frac{\mathrm{const}_{n,\kappa}}{f_n(t)} \ . \tag{2.6}$$

Moreover, we have in this case

$$\left|\frac{d}{dt}\frac{1}{f_n}\right| \asymp \frac{1}{f_{n+1}} \tag{2.7}$$

where the sign "$\asymp$" means that the ratio of the left- and the right-hand sides is bounded from above and from below by positive finite constants depending on n and κ .

Remark. The upper bound $\frac{f' f_1}{f^2} \leq 1$ in the hypothesis (*) simply means that the function $\log \frac{1}{f_1}$ is convex. One can show that convexity of $\log \frac{1}{f_1}$ follows from that of $\log \frac{1}{f}$. In the view of the preceding corollary the assumption of convexity of $\log \frac{1}{f}$ would not be a restrictive one. On the contrary, the lower bound in (*) puts an essential restriction on f , for example, it holds if $f(t) = t^\nu$, $\nu > 0$ or $f(t) = \exp t^\nu, 0 < \nu \leq 1$ but does not if $f(t) = \log^\nu t$.

Let us also note that the relation (2.7) clearly means that the rate $1/f_n(t)$ of decay of the derivative $\frac{\partial^n p}{\partial t^n}$ on increasing of time t is a sharp one up to a constant multiple provided the given inequality (2.5) is sharp. Moreover, by the lemma below the hypothesis (*) implies that

$$f_n(t) \asymp f\left(\frac{f}{f'}\right)^n \ .$$

Proof. The following elementary lemma plays the main role in the proof.

Lemma 2.1 *Suppose that a function $f(t) \in C^1(0,T)$ is such that $f'(t) > 0$ for all $t \in (0,T)$ and for some $a, b \in (0,1]$*

$$b\frac{f^2}{f'} \le f_1 \le a\frac{f^2}{f'} \tag{2.8}$$

then for all $n \ge 1$ and $t \in (0,T)$ we have, first,

$$\beta_n \frac{f_n^2}{f_{n-1}} \le f_{n+1} \le \alpha_n \frac{f_n^2}{f_{n-1}} \tag{2.9}$$

where $\alpha_n, \beta_n \in (0,1]$, and, second,

$$b_n f\left(\frac{f}{f'}\right)^n \le f_n \le a_n f\left(\frac{f}{f'}\right)^n \tag{2.10}$$

with positive constants a_n, b_n .

Proof of lemma. The given inequality (2.8) can be considered as the (2.9) for $n = 0$ if f_{-1} is understood as f' which forms the inductive basis. Let us take (2.9) as the inductive hypothesis and prove that for the next value of n . Indeed, (2.9) implies that $\frac{f_{n-1}f_{n+1}}{f_n^2} \ge \beta_n$ whence it follows

$$f_{n+1} \ge \frac{f_{n+1}}{2-\beta_n}\left(2 - \frac{f_{n-1}f_{n+1}}{f_n^2}\right) = \frac{1}{2-\beta_n}\frac{d}{dt}\left(\frac{f_{n+1}^2}{f_n}\right) .$$

Integrating this inequality from 0 to t we get

$$f_{n+2} \ge \beta_{n+1}\frac{f_{n+1}^2}{f_n}$$

with $\beta_{n+1} = \frac{1}{2-\beta_n}$ (we have to check also that $\frac{f_{n+1}^2}{f_n}|_{t=0} = 0$ which follows from $f_{n+1}(t) \le t f_n(t)$). The right-hand inequality of (2.9) is proved in the same way.

Let us turn to (2.10) which is going to be proved also by induction. The inductive basis consists of two cases: $n = 0$ - evidently, and $n = 1$ - given by (2.8) so that $a_0 = b_0 = 1,\ a_1 = a, b_1 = b$. Suppose that we have proved (2.10) for some $n \ge 1$ and the same for the preceding value $n-1$ and let us prove (2.10) for the next value $n+1$. Indeed, by the inductive hypothesis we can, in particular, estimate f_n^2 from above and f_{n-1} from below as follows

$$f_n^2 \le a_n^2\frac{f^{2n+2}}{f'^{2n}} , \quad f_{n-1} \ge b_{n-1}\frac{f^n}{f'^{n-1}}$$

which yields together with (2.9) that

$$f_{n+1} \le \alpha_n\frac{f_n^2}{f_{n-1}} \le a_{n+1} f\left(\frac{f}{f'}\right)^{n+1}$$

where $a_{n+1} = \alpha_n\frac{a_n^2}{b_{n-1}}$. The inequality in the opposite direction is proved similarly.

The proof of Corollary 2.4 is completed as follows. We apply the upper bound (2.1) of the Corollary 2.1 with $g = f$. The hypothesis (*) enables us to apply Lemma 2.1 with $a = 1$, $b = \kappa$ and we get by (2.10)

$$\sqrt{b_{2n}} f \left(\frac{f}{f'} \right)^n \leq \sqrt{f f_{2n}} \leq \sqrt{a_{2n}} f \left(\frac{f}{f'} \right)^n .$$

Comparing with the similar estimate of f_n we see that $f_n \asymp \sqrt{f f_{2n}}$. Substituting this into (2.1) we obtain finally (2.6) .

The second statement (2.7) of Corollary 2.4 follows directly from the inequality (2.9) . □

3. Gaussian estimates

We return now to the case of a finite D and consider some applications of Theorems 1.1 and 1.2 based upon the results of [6] where the function $E_0(x,t)$ was estimated provided some isoperimetric properties of the manifold in question were known. Following [6] let us say that a Λ-isoperimetric inequality holds in a region $\Omega \subset M$ if for any sub-region D Ω

$$\lambda_1(D) \geq \Lambda(\mu D)$$

where $\lambda_1(D)$ denotes the first Dirichlet eigenvalue of the Laplacian in D and $\Lambda(v)$ is a monotonically decreasing positive continuous function in $(0, +\infty)$.

To formulate the next theorem we introduce the following notation. In any ball $B(x,R)$ on a complete manifold the Λ-isoperimetric inequality holds with the function $\Lambda = \Lambda_{x,R}$:

$$\Lambda_{x,R}(v) = a v^{-\nu} \tag{3.1}$$

where $\nu = 2/\dim M$ and the coefficient $a = a(x,R) > 0$ depends on the geometry of the ball. Indeed, if a region D $B(x,R)$ is small enough so that it lies in a chart being "almost" a Euclidean ball then D satisfies the Euclidean Λ-isoperimetric inequality :

$$\lambda_1(D) \geq \mathrm{const}(\mu D)^{-\nu} .$$

For an arbitrary D $B(x,R)$ the same inequality (with a smaller constant const > 0) follows from compactness arguments exploiting the fact that a geodesic ball is a pre-compact set due to the completeness of the manifold.

Theorem 3.1 *On an arbitrary complete manifold M the following estimate of the n -th time derivative of the heat kernel holds for any $R > 0$ and for all $x, z \in M,\ t > 0$*

$$\left| \frac{\partial^n p}{\partial t^n} \right| (x,z,t) \leq \frac{\mathrm{const}_{m,n} \left(1 + \frac{r^2}{t} \right)^N}{(a(x,R)a(z,R))^{m/2}} \frac{\exp\left(-\frac{r^2}{4t} \right)}{t^n \min(t, R^2)^m} . \tag{3.2}$$

$n \geq 0$ being an arbitrary integer, $m = \frac{1}{\nu},\ N = m + n + 1$ and $r = \mathrm{dist}(x,y)$.

Besides, if $\lambda > 0$ is the spectral gap of the Laplace operator in $L^2(M)$ then

$$\left| \frac{\partial^n p}{\partial t^n} \right| (x,z,t) \leq \frac{\mathrm{const}_{m,n,\lambda,R} \left(1 + \frac{r^2}{t} \right)^N}{(a(x,R)a(z,R))^{m/2}} \frac{\exp\left(-\frac{r^2}{4t} - \lambda t \right)}{\min(t^{n+m}, R^{2m})} . \tag{3.3}$$

Remark. We would like to put the attention of the reader to the factor t^n in the denominator on the right-hand side of (3.2) which means that the derivative $\frac{\partial^n p}{\partial t^n}$ has a priori decay as $t \to \infty$ at least as fast as t^{-n} , and this is true for an arbitrary complete manifold irrespective of its geometry. Of course, if the spectral gap λ is strictly positive then (3.3) give the better decay.

Let us note also that in the statement of the theorem ν is not necessarily equal to $2/\dim M$ although in all known application of this theorem this is so and, consequently, $m = \frac{1}{2}\dim M$. Basically, we need only that ν is positive.

Proof. As was shown in [6] (see the proof of Theorem 5.2 where the inequality (3.2) was obtained for $n = 0$) the following estimate is valid:

$$E_0(x,t) \le \frac{\text{const}_m}{\delta^{1+m} a(x,R)^m \min(t,R^2)^m} \tag{3.4}$$

where

$$\delta = \min(D-2,\ \bar{c}) \tag{3.5}$$

and $\bar{c}$ is a small absolute constant.

To apply Theorems 1.1, 1.2 we have to evaluate the n -th integral of the function

$$f(t) = \min(t,R^2)^m \ .$$

Of course one can compute $f_n(t)$ explicitly but we prefer to deal with a simpler estimate of f_n which is stated in the following lemma.

Lemma 3.1 *Suppose that $f(t)$ is an increasing function in $(0,+\infty)$ growing at most polynomially i.e. for some $\alpha > 0$ and for all $t > 0$*

$$f(\frac{1}{2}t) \ge \alpha f(t) \ , \tag{3.6}$$

then for all $t > 0$ and any integer $n > 0$

$$\frac{\alpha}{2^n n!} f(t) t^n \le f_n(t) \le \frac{1}{n!} f(t) t^n \ .$$

Proof. Indeed, we have by (1.1) and (3.6) :

$$f_n(t) \ge \int_{t/2}^t \frac{(t-\tau)^{n-1}}{(n-1)!} f(\tau) d\tau \ge f(t/2) \int_{t/2}^t \frac{(t-\tau)^{n-1}}{(n-1)!} d\tau \ge \alpha f(t) \frac{(t/2)^n}{n!} \ .$$

The upper bound of f is proved similarly. □

Therefore, by Theorems 1.1, 1.2 and Lemma 3.1 we have for any integer $n \ge 0$ the following estimate of the n -th derivative of the heat kernel:

$$\left| \frac{\partial^n p}{\partial t^n} \right| (x,z,t) \le \frac{\text{const}_{m,n} c^{-n}}{\delta^{1+m} (a(x,R)a(z,R))^{m/2}} \frac{\exp\left(-\frac{r^2}{2Dt}\right)}{t^n \min(t,R^2)^m} \ . \tag{3.7}$$

Next we are going to choose the optimal value of D for any pair r, t exploiting the explicit form of c given by Theorem 1.1. Indeed, let us put

$$D = 2 + \min\left(\bar{c}, \frac{t}{r^2}\right)$$

then $D - 2 \le \bar{c}$ and $\delta = D - 2 = \min\left(\bar{c}, \frac{t}{r^2}\right)$. On the other hand

$$c = \frac{D-2}{D/2+8} \ge \frac{\delta}{\frac{1}{2}(2+\bar{c})+8} = \mathrm{const}\delta$$

Substituting this into (3.7) and observing that

$$\frac{r^2}{2t} - \frac{r^2}{Dt} = \frac{\delta r^2}{2Dt} \le \frac{t}{r^2}\frac{r^2}{2Dt} \le \frac{1}{4}$$

and that

$$\frac{1}{\delta} \le \mathrm{const}\left(1 + \frac{r^2}{t}\right)$$

we get finally (3.2) .

Tracing back the computations one can find that the constant $\mathrm{const}_{n,m}$ in the numerator of (3.2) depends on n is the following manner: $\mathrm{const}_{n,m} = \mathrm{const}_m 2^n \sqrt{(2n)!}$.

In order to prove (3.3) we replace the basic inequality (3.4) by another. As was proved in [6] the function $E_0(x,t)e^{\lambda t}$ is monotonically decreasing in t which implies that for $t > R^2$

$$E_0(x,t) \le E_0(x,R^2)e^{-\lambda(t-R^2)}$$

or, together with (3.4) , we have that for all $t > 0$

$$E_0(x,t) \le \frac{\mathrm{const}_m}{\delta^{1+m}a(x,R)^m \min(t,R^2)^m} e^{-\lambda(t-R^2)_+} .$$

Next we shall get an upper bound of E_n upon application of Theorem 1.1 but first we state the following elementary lemma whose proof is plain.

Lemma 3.2 *Let*

$$f(t) = \min(t,T)^m e^{\lambda(t-T)_+}$$

then

$$f_n(t) \ge C \min(t,T)^{m+n} e^{\lambda(t-T)_+}$$

where $C = \max(m+1, \lambda T)^{-n}$ *and all constants* m, λ, T *are positive.*

Applying this statement and Theorem 1.1 we obtain

$$E_n(x,t) \le \frac{\mathrm{const}_{m,n,\lambda,R}c^{-n}}{\delta^{1+m}a(x,R)^m \min(t,R^2)^{m+n}} e^{-\lambda(t-R^2)_+}$$

whence (3.3) follows in the same manner as above. □

Of course, if one is given some more information about the coefficients $a(x,R)$ then one may hope to obtain from (3.2) extra information about behaviour of $\frac{\partial^n p}{\partial t^n}$. For example, if the manifold is more or less homogeneous in the sense that for a small R all quantities $a(x,R)$ are bounded away from 0 uniformly in x then the decay of $\frac{\partial^n p}{\partial t^n}$ in the distance r is given by the Gaussian factor.

On the other hand if $a(x,R)$ is known for all values of R one may try to optimize (3.2) with respect to R and to obtain thereby a sharper dependence of t . An example of this kind is presented below.

As was shown in [5] for a manifold of a non-negative Ricci curvature one may take

$$a(x,R) = \frac{b}{R^2}\Big(\mu B(x,R)\Big)^{\nu} \tag{3.8}$$

where $b > 0$ is determined by the dimension of the manifold. Application of Theorem 3.1 in this case yields the following.

Corollary 3.1 *Suppose that in any ball the Λ-isoperimetric inequality holds with the function (3.1) , the coefficient $a(x,R)$ being defined by (3.8) where b is a positive constant, then*

$$\left|\frac{\partial^n p}{\partial t^n}\right|(x,z,t) \le \mathrm{const}_{m,n,b}\left(1+\frac{r^2}{t}\right)^{N}\frac{\exp\left(-\frac{r^2}{4t}\right)}{t^n \mu B(x,\sqrt{t})} \tag{3.9}$$

where $m = 1/\nu$ as above and $N = 1 + \frac{3}{2}m + n$.

Indeed, if we simply substitute into (3.2) the values of $a(x,R)$, $a(z,R)$ and set $R = \sqrt{t}$ we get

$$\left|\frac{\partial^n p}{\partial t^n}\right|(x,z,t) \le \frac{\mathrm{const}_{m,n}\left(1+\frac{r^2}{t}\right)^{1+m+n}}{t^n\Big(\mu B(x,\sqrt{t})\mu B(z,\sqrt{t})\Big)^{\frac{1}{2}}}\exp\left(-\frac{r^2}{4t}\right) . \tag{3.10}$$

The idea of a further reduction of the expression on the right-hand side is well-known in the case of a non-negatively curved manifold : one should estimate the ratio of the volumes $\mu B(x,\sqrt{t})$ and $\mu B(z,\sqrt{t})$ via the distance r between points x, z using the volume comparison theorems for such a manifold (see [9]). Under the assumptions of Corollary 3.1 it is a little bit more tricky. As was shown in [6] (see Proposition 5.2 there) the Λ-isoperimetric inequality in question with the coefficient (3.8) implies that for any two balls $B(x,R)$ and $B(z,\rho)$ such that $B(z,\rho) \subset B(x,R)$ the ratio of their volumes is bounded as follows:

$$\frac{\mu B(x,R)}{\mu B(z,\rho)} \le \mathrm{const}_m b^{-m}\left(\frac{R}{\rho}\right)^{2m} . \tag{3.11}$$

Therefore, we have by (3.11) :

$$\begin{aligned}\mu B(x,\sqrt{t}) \le \mu B(z, r+\sqrt{t}) &\le \mathrm{const}_{m,b}\left(\frac{r+\sqrt{t}}{\sqrt{t}}\right)^{2m}\mu B(z,\sqrt{t})\\ &\le \mathrm{const}_{m,b}\left(1+\frac{r^2}{t}\right)^{m}\mu B(z,\sqrt{t})\end{aligned}$$

which enables us to replace in (3.10) the volume of the ball $B(z,\sqrt{t})$ by its lower bound through $\mu B(x,\sqrt{t})$ and obtain finally (3.9) . □

Let us note that Corollary 3.1 is applicable not only to a manifold of a non-negative Ricci curvature but also to a manifold which is quasi-isometric to that of a non-negative Ricci curvature because its hypothesis is stable under a quasi-isometric transformation.

In what follows we assume that the heat kernel p satisfies for all $t \in (0, +\infty)$ and for any $x \in M$ the inequality

$$p(x, x, t) \leq \frac{1}{f(t)} \tag{3.12}$$

where $f(t)$ is a regular enough function. We are going to deduce from (3.12) the Gaussian pointwise estimates of the heat kernel time derivatives. For this purpose we have to impose some regularity conditions on $f(t)$ so that the results of [6] concerning estimation of $E_0(x, t)$ are applicable. Let us suppose that $f(t) \in C^1(0, +\infty)$, $f'(t) > 0$, $f(0+) = 0$, $f(+\infty) = \infty$ and denote $l(t) = \frac{f'(t)}{f(t)}$. We assume the function $l(t)$ to satisfy the following three conditions:

(i) the function $l(t)$ is decreasing

(ii) for some $\alpha > 0$ and any $k \in [1, 2]$ and for all $t > 0$ the following is valid:

$$l(kt) \geq \alpha l(t) \tag{3.13}$$

(iii) for some $K \leq +\infty$ and a positive N

$$\begin{cases} tl(t) \leq N for\ t \leq 2K \\ tl(t)\ is\ increasing\ for\ t > K \end{cases} .$$

The hypothesis (i) implies that the function $f(t)$ grows at least exponentially which is natural because the heat kernel can not decay faster than exponentially. The hypothesis (ii) means that the function $l(t)$ does not decrease very fast - to imagine to what extent it restricts applications let us take $l(t) = t^{-\gamma}$ which satisfies (3.13) for all γ but if $\gamma > 1$ then by integrating the function $l(t)$ we get that $f(+\infty) < \infty$ which is out of our considerations. Finally, the hypothesis (iii) clearly means that the function $tl(t)$ either remains bounded (and in this case $K = \infty$) or is increasing in a neighbourhood of ∞ (which corresponds to a finite K) which excludes only big jumps of the graph of this function.

For example, the function $f(t) = \log^A t \cdot t^B \exp t^C$, $t > 2$ satisfies all the hypotheses (i)-(iii) provided $A \geq 0, B \geq 0, 0 \leq C \leq 1$.

Corollary 3.2 *Suppose that the heat kernel satisfies the inequalities (3.12) with the function $f(t)$ subject to (i)-(iii), then for any integer $n \geq 0$ and any $D > 2$ the following estimate holds for all $t > 0$ and all $x, z \in M$*

$$\left| \frac{\partial^n p}{\partial t^n} \right| (x, z, t) \leq \frac{\mathrm{const}_{N,K} C^n \exp\left(-\frac{r^2}{2Dt} \right)}{\delta \sqrt{f(\widetilde{c}t) f_{2n}(\widetilde{c}t)}} . \tag{3.14}$$

where $r = \mathrm{dist}(x, z)$, the constant $\delta = \delta(D)$ is defined by (3.5) , $\widetilde{c} = \mathrm{const}_\alpha \delta$ and $C = C(\alpha)$.

To prove this we shall apply two theorems from [6] : the first is Theorem 2.2 there which states that the heat kernel on-diagonal estimate (3.12) implies some Λ-isoperimetric inequality , and the second is the Theorem 4.2 which estimates $E_0(x,t)$ provided a Λ-isoperimetric inequality is given (we used a particular case of this theorem when proving Theorem 3.1). Afterwards we can apply Theorem 1.2.

The Theorem 2.2 from [6] is applicable under the conditions (i) and (ii) and states that (3.12) implies that for any bounded region $\Omega \subset M$

$$\lambda_1(\Omega) \geq \text{const}_\alpha \Lambda(\mu\Omega) \tag{3.15}$$

where the function Λ is defined by

$$\Lambda(f(t)) = \frac{f'(t)}{f(t)}.$$

The Theorem 4.2 from [6] states that under the hypothesis (3.15) (and if the function $f(t)$ satisfies (iii)) the following inequality holds

$$E_0(x,t) \leq \frac{\text{const}_{N,K}}{\delta f(\hat{c}\delta t)}$$

where $\hat{c}$ is proportional to the constant const_α from (3.15) . Observing that

$$f(\hat{c}\delta t)_n = \delta^{-n}\hat{c}^{-n} f_n(\hat{c}\delta t)$$

we get upon application of Theorems 1.1, 1.2 that

$$\left|\frac{\partial^n p}{\partial t^n}\right|(x,z,t) \leq \frac{\text{const}_{N,K} c^{-n}\delta^n \hat{c}^n \exp\left(-\frac{r^2}{2Dt}\right)}{\delta\sqrt{f(\frac{\hat{c}}{2}\delta t) f_{2n}(\frac{\hat{c}}{2}\delta t)}} \tag{3.16}$$

where c is defined as above by (1.4) . Let us note that

$$c^{-1}\delta = \frac{D/2+8}{D-2}\min(D-2,\bar{c}) \leq 9 + \frac{\bar{c}}{2}$$

and the right-hand side here is an absolute constant. Hence, $c^{-1}\delta\hat{c} \leq C = C(\alpha)$ and substituting it into (3.16) we obtain finally (3.14) . □

The explicit dependence of δ in the inequality (3.14) can be essential when optimizing the right-hand side with respect to D . For example, if the function $f(t)$ grows at most polynomially in the sense of Lemma 3.1 then it follows that for some positive constant m

$$f(\tilde{c}t) \geq \text{const}\tilde{c}^m f(t) = \text{const}_{\alpha,m}\delta^m f(t)$$

which enables us to make clear dependence of δ . Applying Lemma 3.1 to estimate f_{2n} and choosing the value D as in Theorem 3.1 we obtain

$$\left|\frac{\partial^n p}{\partial t^n}\right|(x,z,t) \leq \text{const}\left(1+\frac{r^2}{t}\right)^{1+m+n}\frac{\exp\left(-\frac{r^2}{4t}\right)}{t^n f(t)} . \tag{3.17}$$

In the particular case of the function $f(t) = \text{const}\, t^m$ an estimate of $\frac{\partial p}{\partial t}$ similar to (3.17) was proved by B.Davies [3] . His result is slightly sharper because it contains the factor $\left(1+\frac{r^2}{t}\right)$ to the power $1+m$ instead of our $2+m$.

REFERENCES

[1] **Chavel I.**, “Eigenvalues in Riemannian geometry” Academic Press, New York, 1984.

[2] **Cheng S.Y., Li P., Yau S.-T.**, On the upper estimate of the heat kernel of a complete Riemannian manifold, *Amer. J. Math.* , **103** (1981) no.5, 1021-1063.

[3] **Davies E.B.**, Pointwise bounds on the space and time derivatives of the heat kernel, *J. Operator Theory* , **21** (1989) 367-378.

[4] **Dodziuk J.**, Maximum principle for parabolic inequalities and the heat flow on open manifolds, *Indiana Univ. Math. J.* , **32** (1983) no.5, 703-716.

[5] **Grigor'yan A.A.**, The heat equation on non-compact Riemannian manifolds, (in Russian) *Matem. Sbornik* , **182** (1991) no.1, 55-87. Engl. transl. *Math. USSR Sb.* , **72** (1992) no.1, 47-77.

[6] **Grigor'yan A.**, Heat kernel upper bounds on a complete non-compact manifold, *Revista Mathemática Iberoamericana* , **10** no.2, (1994) 395-452.

[7] **Grigor'yan A.**, Integral maximum principle and its applications, *Proc. Roy. Soc. Edinburgh* , **124A** (1994) 353-362.

[8] **Kovalenko V.P., Semenov U.A.**, Semigroups generated by an elliptic operator of second order, (in Russian) *in:* “Methods of Functional Analysis in Problems of Mathematical Physics” , Kiev Inst. of Math., Ukranian Acad. of Sciences, 1987. 17-36.

[9] **Li P., Yau S.-T.**, On the parabolic kernel of the Schrödinger operator, *Acta Math.* , **156** (1986) no.3-4, 153-201.

[10] **Porper F.O.**, Estimates of the derivatives of the fundamental solution of a stationary parabolic divergence equation with constants that do not depend on the smoothness of the coefficients, (in Russian) *Dokl. AN SSSR* , **235** (1977) 1022-1025. Engl. transl. *Soviet Math. Dokl.* , **18** (1977) 1092-1096.

[11] **Porper F.O., Eidel'man S.D.**, Two-side estimates of fundamental solutions of second-order parabolic equations and some applications, (in Russian) *Uspechi Mat. Nauk* , **39** (1984) no.3, 101-156. Engl. transl. *Russian Math. Surveys* , **39** (1984) no.3, 119-178.

[12] **Robinson D.W.**, “Elliptic operators and Lie groups” Oxford New York Tokyo, Clarenton Press, Oxford Math. Mono., 1991.

[13] **Varopoulos N.Th.**, Semi-groupes d'opérateurs sur les espaces L^p, *C. R. Acad. Sci. Paris Sér. I Math.* , **301** (1985) 865-868.

[14] **Varopoulos N.Th., Saloff-Coste L., Coulhon T.**, “Analysis and geometry on groups” Cambridge University Press, Cambridge U.K., 1992.