# GAUSSIAN UPPER BOUNDS FOR THE HEAT KERNEL AND FOR ITS DERIVATIVES ON A RIEMANNIAN MANIFOLD

Alexander Grigor'yan

## 1. Main results

Given a Riemannian manifold $M$ we consider the heat kernel $p(x, y, t)$ being by definition the smallest positive fundamental solution to the heat equation $u_t - \Delta u = 0$ on $M \times (0, +\infty)$ where $\Delta$ is the Laplace operator associated with the Riemannian metric. The question to be discussed here is how to get Gaussian upper estimates of $p(x, y, t)$ and of its time derivatives $\frac{\partial^m p}{\partial t^m}(x, y, t)$ provided we know a priori for all $t > 0$ an on-diagonal upper bound

$$p(x, x, t) \leq \frac{1}{f(t)} \tag{1.1}$$

where $f(t)$ is an increasing function, $x$ is a given point on $M$.

In the simplest case when the manifolds $M$ is a Euclidean space $\mathbb{R}^n$ we have

$$p(x, y, t) = \frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{r^2}{4t}\right) \tag{1.2}$$

where $r = \operatorname{dist}(x, y)$, and (1.1) holds with $f(t) = \frac{\text{const}}{t^{n/2}}$. A plain computation yields that the $m-$ th time derivative $\frac{\partial^m p}{\partial t^m}$ has in this case the sign $(-1)^m$ and its absolute value is estimated as follows

$$\left|\frac{\partial^m p}{\partial t^m}\right|(x, y, t) \asymp \frac{1}{t^{n/2+m}} \left(1 + \frac{r^2}{t}\right)^m \exp\left(-\frac{r^2}{4t}\right) \tag{1.3}$$

where the sign $\asymp$ means that the ratio of the left- and right-hand sides in (1.3) is bounded from above and below by constants depending only on $n$ and $m$.

Similar inequalities can be obtained in a more general situation as will be shown below. From now on we assume that the manifold in question is non-compact and complete. The most interesting aspects of what follows are connected to behaviour of the heat kernel $p(x, y, t)$ and its derivatives as $t \to \infty$ and as $r \to \infty$ where $r$ denotes a geodesic distance between the points $x$ and $y$.

In order to formulate the main result let us introduce the following notation. Let us fix some constant $D > 2$ and put

$$E^D(x, t) = \int_M p^2(x, y, t) \exp\left(\frac{r^2}{Dt}\right) dy \tag{1.4}$$

where $r = \operatorname{dist}(x, y)$ (dist stands for a geodesic distance and $dy$ is the Riemannian volume element).

---

**Grigor'yan A.**, Gaussian upper bounds for the heat kernel and for its derivatives on a Riemannian manifold, *in:* "Proceeding of the ARW on Potential Theory, Chateau de Bonas, July 1993" *ed. K.GowriSankaran* , Kluwer Academic Publisher, 1994. 237-252.

**Theorem 1.1** *Let $f(t)$ be an increasing function on $(0,\infty)$ of at most a polynomial growth in the sense that*

$$f(2t) \leq Af(t) \tag{1.5}$$

*for all $t > 0$ and for some constant $A$. Suppose that the heat kernel on a manifold $M$ satisfies the on-diagonal estimate (1.1) for some (fixed) point $x \in M$ and for all $t > 0$. Then for any $D > 2$ and for all $t > 0$ we have*

$$E^D(x,t) \leq \frac{\mathrm{const}_A}{f(\delta t)} \tag{1.6}$$

*where $\delta = \min(D-2, 1)$.*

Of course, applying (1.5) one sees that $f(\delta t) \geq \mathrm{const}_\delta f(t)$ and the factor $\delta$ can be absorbed into the constant in the numerator of (1.6) but sometimes the dependence on $D$ in (1.6) is essential and the inequality (1.6) enables one to catch it. Note that the hypothesis $D > 2$ can not be relaxed: if $D = 2$ then in the Euclidean space we have $E^D = \infty$.

Let us also observe that the assumption that (1.1) is true for *all* $t > 0$ does not restrict applicability of the theorem to the case when (1.1) is known only for $t < T$. Indeed, the function $p(x,x,t)$ is decreasing in $t$, therefore, if (1.1) holds for $t < T$ then the function $f(t)$ can be extended for $t \geq T$ simply as the constant $f(T)$ and (1.1) will be valid for all $t > 0$.

Theorem 1.1 will be proved in Section 3 below. In Section 2, we shall outline other related results and methods. Now, let us explain how to get pointwise upper Gaussian estimations similar to (1.3) applying this theorem. As was proved in [9] the following inequality holds always irrespective of geometry of a manifold

$$p(x,y,t) \leq \sqrt{E^D(x,t/2)E^D(y,t/2)}\exp\left(-\frac{r^2}{2Dt}\right) . \tag{1.7}$$

Therefore, if we are given that the on-diagonal estimate (1.1) holds for a point $x$ as well as for a point $y$ then by (1.6) and (1.7)

$$p(x,y,t) \leq \frac{\mathrm{const}}{f(\frac{\delta}{2}t)}\exp\left(-\frac{r^2}{2Dt}\right) . \tag{1.8}$$

In order to get inequalities involving derivatives of the heat kernel let us define the powers of the gradient $\nabla$ as follows: $\nabla^m$ means $\Delta^m$ if $m$ is even and $\nabla\Delta^{\frac{m-1}{2}}$ if $m$ is odd. Let us fix some number $D > 2$ and introduce the following series of functions $E_m(x,t)$ :

$$E_m(x,t) = \int_M |\nabla^m p|^2(x,y,t)\exp\left(\frac{r^2}{Dt}\right) dy, \ m = 0,1,2,... \tag{1.9}$$

where the operator $\nabla$ relates to the variable $y$ and $r = \mathrm{dist}(x,y)$. Obviously, $E^D$ is the same as $E_0$ in the new notation. Of course, it would be correct to write $E_m^D$ in place of $E_m$ but we skip the superscript $D$ in order to simplify notations when the constant $D$ is fixed.

The following two theorem were proved in [11] .

**Theorem 1.2** *If $D > 2$ then for any $x \in M$ and any integer $m \geq 0$ the function $E_m(x,t)$ is finite and decreasing in $t$. Besides, for all $x, y \in M$ and any $t > 0$*

$$\left|\frac{\partial^m p}{\partial t^m}\right|(x,y,t) \leq \sqrt{E_{2m}(x,t/2)E_0(y,t/2)}\exp\left(-\frac{r^2}{2Dt}\right) \tag{1.10}$$

In fact, the inequality (1.10) as well as (1.7) above are derived from the semigroup identity

$$p(x,y,t) = \int_M p(x,z,t-s)p(z,y,s)dy \tag{1.11}$$

upon differentiation in $t$ and a proper application of Cauchy-Schwarz inequality using the fact that $\frac{\partial^m p}{\partial t^m} = \Delta^m p = \nabla^{2m} p$.

Theorem 1.2 reduces the question of finding upper bounds of $\left|\frac{\partial^m p}{\partial t^m}\right|$ to that of $E_m$ . It turns out that $E_m$ for $m > 0$ can be estimated directly via $E_0$ as follows.

**Theorem 1.3** *Suppose that for some point $x \in M$ and all $t > 0$*

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

*where $\varphi(t)$ is a positive increasing function on $(0,\infty)$ then for any integer $m \geq 1$ and for all $t > 0$*

$$E_m(x,t) \leq \frac{C^m}{\varphi_m(t)} \tag{1.12}$$

*where $C = \frac{D/2+8}{D-2}$ and $\varphi_m(t)$ denotes the $m-$ th integral of the function $\varphi(t)$, that is to say, $\varphi_0 = \varphi$ and for $m > 0$*

$$\varphi_m(t) = \int_0^t \varphi_{m-1}(\tau)d\tau \ .$$

Let us observe that the inequality (1.12) can be rewritten in the following form

$$E_m(x,t) \leq C^m\left(\int_0^t \frac{(t-\tau)^{m-1}}{(m-1)!}\frac{d\tau}{E_0(x,\tau)}\right)^{-1}$$

if we put in Theorem 1.3 $\varphi(t) = 1/E_0(x,t)$.

The following statement is a straightforward consequence of the theorems 1.1-1.3.

**Corollary 1.1** *Let $f(t)$ and $g(t)$ be increasing in $t$ functions either satisfying the hypothesis (1.5) of at most a polynomial growth. Assume that for two points $x, y \in M$ and for all $t > 0$*

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

*then for any $D > 2$, any integer $m \geq 0$ and for all $t > 0$*

$$\left|\frac{\partial^m p}{\partial t^m}\right|(x,y,t) \leq \frac{\mathrm{const}_{A,D,m}}{t^m\sqrt{f(\frac{\delta}{2}t)g(\frac{\delta}{2}t)}}\exp\left(-\frac{r^2}{2Dt}\right) \tag{1.13}$$

*where*

$$\mathrm{const}_{A,D,m} = \mathrm{const}_A \delta^{-m} 20^m \sqrt{(2m)!} \tag{1.14}$$

*and* $\delta = \min(D-2, 1)$ *is the same as in Theorem 1.1.*

*Moreover, for some* $\nu = \nu(A)$

$$\left|\frac{\partial^m p}{\partial t^m}\right|(x,y,t) \le \frac{\mathrm{const}_{A,m}}{t^m\sqrt{f(t)g(t)}}\left(1+\frac{r^2}{t}\right)^{\nu+m} \exp\left(-\frac{r^2}{4t}\right) \quad . \tag{1.15}$$

In fact, the exponent $\nu$ is exactly the number for which the inequality is valid:

$$\frac{f(t_2)}{f(t_1)} \le \mathrm{const}\left(\frac{t_2}{t_1}\right)^{\nu} \quad \forall t_2 > t_1 > 0 \tag{1.16}$$

and the same for the function $g$. It is plain that (1.16) follows from (1.5) but sometimes it is useful to postulate (1.16) separately in order to have a better value of $\nu$.

## 2. An outline of previous results and methods

Let us compare the announced Theorem 1.1 and Corollary 1.1 with the previous results of this kind. A numerous works are devoted to heat kernel's estimations - the size of this note is not enough to mention even a small part of them. We consider below several types of the known theorems from our standpoint - what Gaussian bounds of the heat kernel can be derived from an on-diagonal estimate. Let us note that the basic results in this direction are due to Davies [4] , [5] and Varopoulos [16] .

We discuss more detailed approaches to estimation of the time derivatives of the heat kernel which are due to Porper [13] , Cheng, Li, Yau [3] , Varopoulos [16] , [18] , Davies [8] and and Kovalenko, Semenov [12] based on various ideas. The common achievement of these works is that upper bounds of the time derivatives follows from upper bounds of the heat kernel itself without any additional geometric assumptions. Of course, the same is stated also by Corollary 1.1 in the most flexible and sharp form.

*1. Bounds which are uniform in x*

These are theorems which yield the off-diagonal upper bound (1.8) under the hypothesis that the on-diagonal bound (1.1) holds for *all* points $x$. Such a statement for the case of a polynomial function

$$f(t) = \mathrm{const}\begin{cases} t^{\alpha} & , \ t < 1 \\ t^{\beta} & , \ t \ge 1 \end{cases}$$

was first proved by Davies [4] using a log-Sobolev inequality as an intermediate step between on-diagonal and off-diagonal upper bounds of the heat kernel. Another approach was offered earlier by Ushakov [15] for the setting of parabolic equations in unbounded domains in $\mathbb{R}^n$ but without sharp exponent $2D$ in (1.8) . We apply an improved version of the latter method in the proof of Theorem 1.1. Let us emphasize is this connection that when applying Corollary 1.1 the on-diagonal upper bound need be checked as a hypothesis only at *two* points $x, y$ rather than for all points.

*2. Bounds which are non-uniform in x*

These are non-homogeneous estimations when one assumes that for any $x$

$$p(x, x, t) \leq \varphi(x, t)$$

where behaviour in $t$ might be different for different $x$. A theorem of Davies [5] states that

$$p(x, y, t) \leq \mathrm{const}\sqrt{\varphi(x, t)\varphi(y, t)}\exp\left(-\frac{r^2}{2Dt}\right) \tag{2.1}$$

provided the function $\psi(x, t) = \sqrt{\varphi(x, t)}$ satisfies the conditions

$$|\psi_t| \leq \mathrm{const}\frac{\psi}{t} \tag{2.2}$$

and

$$\Delta\psi \leq \mathrm{const}\frac{\psi}{t}\ . \tag{2.3}$$

In the view of Corollary 1.1, the most restrictive second condition (2.3) is actually superfluous: indeed, (2.2) implies that the function $f_x(t) = \frac{1}{\varphi(x,t)}$ satisfies the condition (1.5) of a polynomial growth in $t$ for any point $x$. Therefore, by Corollary 1.1 (case $m = 0$ ) we get (2.1) .

*3. Gaussian estimates of $p(x, y, t)$ with a polynomial correction term*

The first results are due to Varopoulos [16] , [17] and in the sharpest form to Davies, Pang [6] . Following [16] , such estimations as (1.15) containing the factor $1 + \frac{r^2}{t}$ to some power are derived from (1.13) upon optimization with respect to $D$ provided one knows an explicit dependence on $D$ of other constants. Our estimate (1.15) (case m=0) gives the same power of this factor as in [6] provided $f(t) = g(t) = \mathrm{const}t^\nu$. Again, the advantage of Corollary 1.1 is that it needs the initial on-diagonal estimate only at two points $x, y$ in contrast to all the previous results.

*4. Superpolynomial decay of the heat kernel*

By the hypotheses of Theorem 1.1 and Corollary 1.1, the function $f(t)$ can not increase faster than polynomially. There exists, in fact, only one related result which catches the opposite situation. This is a consequence of theorems 2.2 and 4.2 from [9] , and it states that if the inequality (1.1) is true for any $x \in M$ and for all $t > 0$ then, again, for any $x \in M$ and $t > 0$

$$E_0(x, t) \leq \frac{\mathrm{const}}{f(\delta t)}$$

where $\delta = \delta(D)$ and $D > 2$ is arbitrary, provided the function $f(t)$ satisfies certain regularity conditions. Without going into details of these conditions, let us only mention that they admit also a superpolynomial function $f(t)$, for example, $f(t) = \exp t^\nu$, for large $t$ where $0 < \nu \leq 1$. On the other hand, in order to run this theorem, one must have the hypothesis (1.1) be true at once for *all* $x$ whereas in Theorem 1.1, one needs the same at a *single* point $x$. Hence, in this sense, Theorem 1.1 is more flexible.

*4. Estimations of the time derivatives by the method of Porper*

This method was developed by Porper [13] , [14] for the setting of a uniform parabolic equation in $\mathbb{R}^n$, and it goes through on a manifold as well. The starting point is the assumption that for some points $x, y$ we are given that for all $t > 0$

$$p(x,x,t) \leq \frac{\text{const}}{t^\nu}, \quad p(y,y,t) \leq \frac{\text{const}}{t^\nu} \tag{2.4}$$

and

$$p(x,y,t) \leq \frac{\text{const}}{t^\nu} \exp\left(-\frac{r^2}{\sigma t}\right) \tag{2.5}$$

and the objective is to obtain similar bounds for the time derivatives of the heat kernel.

Let us introduce the notation $W_k(x,t)$ similar to $E_k(x,t)$ but without the Gaussian weight:

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

in particular, we have by (1.11)

$$W_0(x,t) = p(x,x,2t) \; . \tag{2.6}$$

By differentiation in $t$ one can show that

$$\frac{\partial^k W_0}{\partial t^k}(x,t) = (-2)^k W_k(x,t)$$

which implies, in particular, that the function $W_0(x,t)$ is convex and decreasing in $t$ whence it follows that

$$W_1(x,t) = -\frac{1}{2}\frac{\partial W_0}{\partial t}(x,t) \leq \frac{1}{2}\frac{W_0(x,t/2) - W_0(x,t)}{t/2} \leq \frac{1}{t} W_0(x,t/2) \; . \tag{2.7}$$

Similarly, for any integer $k$

$$W_k(x,t) \leq \frac{1}{t} W_{k-1}(x,t/2) \leq \frac{1}{t^k} W_0(x,t/2^k) \; . \tag{2.8}$$

Besides, the semigroup identity (1.11) enables one to get the following initial estimate

$$\left|\frac{\partial^k p}{\partial t^k}\right|(x,y,t) \leq \sqrt{W_{2k}(x,t/2)W_0(y,t)} \; . \tag{2.9}$$

Let now $x, y$ be the points for which (2.4) and (2.5) hold, then by substituting into (2.9) successively (2.8) , (2.6) and (2.4) we get that

$$\left|\frac{\partial^k p}{\partial t^k}\right|(x,y,t) \leq \frac{\text{const}}{t^{\nu+k}} \; . \tag{2.10}$$

In order to involve the Gaussian factor one estimates first $\frac{\partial p}{\partial t}(x,y,\theta)$ at a mean point $\theta \in (t, t+\delta)$ (for some $\delta > 0$ ) by the mean-value theorem

$$\left|\frac{\partial p}{\partial t}\right|(x,y,\theta) \leq \frac{1}{\delta}\left|p(x,y,t+\delta) - p(x,y,t)\right| \leq \frac{p(x,y,t+\delta) + p(x,y,t)}{\delta} \; . \tag{2.11}$$

Applying again the mean-value theorem to the function $\frac{\partial p}{\partial t}$ we have

$$\left|\frac{\partial p}{\partial t}\right|(x,y,t) \leq \left|\frac{\partial p}{\partial t}\right|(x,y,\theta) + \delta \sup_{\tau \in (t, t+\delta)} \left|\frac{\partial^2 p}{\partial t^2}\right|(x,y,\tau) \ . \tag{2.12}$$

Next we substitute into (2.12) the upper bound of $\left|\frac{\partial p}{\partial t}\right|(x,y,\theta)$ from (2.11) , the upper bound of $\left|\frac{\partial^2 p}{\partial t^2}\right|(x,y,\tau)$ obtained by (2.10) and the upper bounds of $p(x,y,t+\delta), p(x,y,t)$ according to (2.5) which yields

$$\left|\frac{\partial p}{\partial t}\right|(x,y,t) \leq \frac{1}{\delta}\frac{\text{const}}{t^\nu} \exp\left(-\frac{r^2}{\sigma(t+\delta)}\right) + \delta \frac{\text{const}}{t^{\nu+2}}$$

whence by choosing an optimal value of $\delta$ which is to be $\delta = \varepsilon t \exp\left(-\frac{r^2}{2\sigma t}\right) \leq \varepsilon t$ for a small $\varepsilon > 0$ it follows finally

$$\left|\frac{\partial p}{\partial t}\right|(x,y,t) \leq \frac{\text{const}}{t^{\nu+1}} \exp\left(-\frac{r^2}{\sigma_1 t}\right) \tag{2.13}$$

with $\sigma_1 = 2\frac{1+\varepsilon}{1-\varepsilon}\sigma$.

The advantage of this method is that in order to get an upper bound of $\left|\frac{\partial p}{\partial t}\right|(x,y,t)$ one need only be given the upper bound of the heat kernel itself at the same points $x, y$. On the other hand, it does not yield the sharp Gaussian exponent - we have in (2.13) under the exponential the coefficient $\sigma_1 > 2\sigma$ instead of expected $\sigma$.

5. *Integral estimations of derivatives according to Cheng-Li-Yau*

The original purpose of the method to be outlined below was to get pointwise upper bounds of the space derivatives of the heat kernel under the assumption that the curvature of the manifold is bounded. We have extracted a part of arguments of Cheng, Li, Yau [3] which involve no curvature and which enable one to get crucial integral estimations of $\nabla^k p$.

Let us consider side by side with the quantities $W_k(x,t)$ the integrals of $\nabla^k p$ over an exterior of a ball

$$W_k^R(x,t) = \int_{r \geq R} \left|\nabla^k p\right|^2 (x,y,t) dy \ .$$

where $r = \text{dist}(x,y)$. We shall concentrate on obtaining a Gaussian estimation of $W_k$ of the following kind:

$$W_k^R(x,t) \leq \frac{\text{const}}{t^{\nu+k}} \exp\left(-\frac{r^2}{D_k t}\right) \tag{2.14}$$

(which implies in turn pointwise Gaussian upper bounds of the time derivatives of the heat kernel upon application of the semigroup identity) under assumption that (2.14) is known to hold for $k = 0$ :

$$W_0^R(x,t) \leq \frac{\text{const}}{t^\nu} \exp\left(-\frac{r^2}{Dt}\right) \tag{2.15}$$

for a given point $x$ and for all $t > 0$.

First, one proves (2.8) by means of the eigenvalue expansion. We skip this part of the arguments and pass directly to Gaussian bounds. Obviously, we have by (2.15) and (2.8)

$$W_k^R(x,t) \le W_k(x,t) \le \frac{1}{t^k} W_0(x, t/2^k) \le \frac{\mathrm{const}}{t^{\nu+k}} \ . \tag{2.16}$$

The following inequality is true for any smooth functions $u(y)$, $\varphi(y)$ provided $\varphi(y)$ is finitely supported:

$$\int_M \left|\nabla^{k+1} u\right|^2 \varphi^2 \le 4\int_M \left|\nabla^k u\right|^2 \left|\nabla \varphi\right|^2 + 2\Big(\int_M \left|\nabla^k u\right|^2 \varphi^2 \int_M \left|\nabla^{k+2} u\right|^2 \varphi^2\Big)^{\frac{1}{2}}$$

which is proved by a standard technique of integration by parts. By an appropriate choice of $\varphi$ and by putting $u(y) = p(x,y,t)$ one deduces from it for any $\gamma < 1$

$$W_{k+1}^R \le \frac{\mathrm{const}}{R^2} W_k^{\gamma R} + 2\Big(W_k^{\gamma R} W_{k+2}^{\gamma R}\Big)^{\frac{1}{2}} \le \mathrm{const}\sqrt{W_k^{\gamma R}}\Big(\frac{1}{R^2}\sqrt{W_k} + \sqrt{W_{k+2}}\Big) \ .$$

all functions $W$ being taken at the point $(x,t)$.

Let us estimate $W_k$ and $W_{k+2}$ by means of (2.16) and suppose that (2.14) is true by the inductive hypothesis, then the formula above yields

$$W_{k+1}^R(x,t) \le \frac{\mathrm{const}}{t^{\nu+k+1}} \exp\Big(-\frac{\gamma^2 R^2}{2D_k t}\Big)\Big(\frac{t}{R^2} + 1\Big) \ .$$

If $R^2/t > 1$ then the last factor on the right-hand side can be absorbed into the Gaussian term and we get

$$W_{k+1}^R(x,t) \le \frac{\mathrm{const}}{t^{\nu+k+1}} \exp\Big(-\frac{r^2}{D_{k+1} t}\Big) \tag{2.17}$$

where $D_{k+1} > 2D_k$ as in the above method of Porper. If $R^2/t \le 1$ then (2.17) follows directly from (2.16) .

We see that this methods gives no sharp Gaussian term either. On the other side, an important advantage of the method is that one need not know that $p$ is a solution to the heat equation in order to deduce (2.14) from (2.15) and (2.16) . Also, the hypotheses (2.15) and (2.16) need hold only at a single point $x$.

*6. A method based upon the semigroup theory*

The following approach arose independently in the papers of Varopoulos [16] and Kovalenko, Semenov [12] and is based upon the fact the heat semigroup $e^{t\Delta}$ in $L^2(M)$ is holomorphic. Since this semigroup is also submarkovian then by general properties of holomorphic semigroups we have

$$\left\|\Delta e^{t\Delta}\right\|_{2\to 2} \le \frac{\mathrm{const}}{t} \tag{2.18}$$

where $p \to q$ means that one considers an operator acting from $L^p(M)$ to $L^q(M)$. By a standard interpolation we have

$$\left\|\Delta e^{3t\Delta}\right\|_{1\to\infty} \le \left\|e^{t\Delta}\right\|_{1\to 2}\left\|\Delta e^{t\Delta}\right\|_{2\to 2}\left\|e^{t\Delta}\right\|_{2\to\infty} \ . \tag{2.19}$$

Since either norm $\left\|e^{t\Delta}\right\|_{1\to 2}$ and $\left\|e^{t\Delta}\right\|_{2\to\infty}$ is equal to $\sup_x p(x,x,t)$ and the operator $\Delta e^{t\Delta}$ has the kernel $\frac{\partial p}{\partial t}(x,y,t)$ then (2.19) implies that

$$\sup_{x,y}\left|\frac{\partial p}{\partial t}\right|(x,y,3t) \le \frac{\text{const}}{t}\sup_x p(x,x,t)\ . \tag{2.20}$$

The higher order derivatives are estimated similarly. The Gaussian bounds are obtained using a perturbed semigroup in the spirit of Davies [7] . We do not go into further details because these arguments are presented very well in the literature. We shall only mention that the sharpest Gaussian term in estimations of the time derivative of the heat kernel is obtained by Davies [8] by a modification of this approach - in place of using (2.18) he considered the heat kernel for complex values of the time and applied the Cauchy integral formula in order to estimate $\frac{\partial p}{\partial t}$.

Although this method gives a sharp Gaussian factor and a sharp polynomial correction term it has also a drawback - its applicability depends very heavily on the initial on-diagonal upper bound of the heat kernel which must be valid uniformly for *all* $x$ as it is seen from (2.20) .

Comparing the methods above with one based upon Theorems 1.1-1.3 we should like to mention that our approach possesses all their advantages - it yields the sharp Gaussian factor as well as it is not tied to a uniform in $x$ behaviour of $p(x,x,t)$.

## 3. Proof of Theorem 1.1 and Corollary 1.1

*Proof of Theorem 1.1.* In the course of the proof, we would have to deal with different unproper integrals of the heat kernel which could a priori be equal to $\infty$. In order to avoid such difficulties we replace the heat kernel $p$ by that of a bounded region. Indeed, by definition the heat kernel $p$ is constructed as a limit of the local heat kernels $p_\Omega$ :

$$p(x,y,t) = \lim_{\Omega\to M} p_\Omega(x,y,t)$$

where $\Omega$ is a bounded region with a smooth boundary, $p_\Omega$ is a heat kernel in $\Omega$ with the vanishing Dirichlet boundary values and $\Omega \to M$ means an exhaustion of $M$ by an expanding sequence of such regions (see [2] for details of this definition).

Since by the maximum principle $p_\Omega \le p$ the hypothesis (1.1) holds for the local heat kernel $p_\Omega$ too. We fix from the very beginning a region $\Omega$ and shall perform all the proof for the heat kernel in this region obtaining the inequality (1.6) for $E^D_\Omega$ (with an obvious definition of $E^D_\Omega$ ) instead of $E^D$.

If we have proved already that $E^D_\Omega(x,t) \le \mathcal{R}$ where $\mathcal{R}$ stands for the right-hand side of (1.6) then for any compact $K \subset M$ and for any $\Omega$ containing $K$ we have

$$\int_K p^2_\Omega(x,y,t)\exp\!\left(\frac{r^2}{Dt}\right)\, dy \le \mathcal{R}$$

whence it follows upon passage to the limit as $\Omega \to M$ that the same inequality is true for $p$ in place of $p_\Omega$. Since $K$ is arbitrary it follows that in this inequality $K$ can be replaced by $M$ which yields $E^D \le \mathcal{R}$.

Thus, from now on we shall deal with the local heat kernel $p_\Omega$ and in order to simplify notations we shall suppress the subscript $\Omega$. We suppose that $\Omega$ is wide enough so that the point $x$ for which (1.1) holds belongs to $\Omega$. Also, we extend the function $p(x,y,t) \equiv p_\Omega(x,y,t)$ to all $y \in M$ so that it vanishes if $y$ lands outside $\Omega$.

The main ingredient of the proof is the following lemma which is proved by a version of the method invented by Ushakov [15] for parabolic equations in $\mathbb{R}^n$. Let us denote by $B(x,R)$ a geodesic ball of radius $R$ centred at the point $x$.

**Lemma 3.1** *Suppose that $f(t)$ is an increasing function which is at most polynomial in the sense of (1.5) and let $D \geq D_0$ where $D_0$ is a large absolute constant (say, $D_0 = 200$ ). Then the on-diagonal bound (1.1) implies that for any $R > 0$*

$$\int_{M\setminus B(x,R)} p^2(x,y,t)dy \leq \frac{\mathrm{const}_A}{f(t)} \exp\left(-\frac{R^2}{Dt}\right) \tag{3.1}$$

*Proof.* Let us introduce the function:

$$d(y) = \begin{cases} R-r & , \ r \leq R \\ 0 & , \ r > R \end{cases}$$

where $r = \mathrm{dist}(x,y)$ and for some $T > 0$ consider the function

$$\xi(y,t) = \frac{d(y)^2}{2(t-T)}, \ 0 < t < T,$$

then we have obviously

$$\frac{\partial \xi}{\partial t} + \frac{1}{2}|\nabla \xi|^2 \leq 0 \ . \tag{3.2}$$

We shall apply the integral maximum principle which is stated in the following lemma.

**Lemma 3.2** *If a function $u(y,t)$ satisfies the heat equation $u_t - \Delta u = 0$ in $\Omega \times (t_1,t_2)$ with Dirichlet boundary values $u|_{\partial\Omega} = 0$ then the integral*

$$\int_\Omega u^2(y,t)e^{\xi(y,t)}dy$$

*is a decreasing function of $t \in (t_1,t_2)$ provided the function $\xi$ satisfies (3.2) in $\Omega \times (t_1,t_2)$.*

This property of solutions to parabolic equations was discovered by Aronson [1] . The proof for the setting of manifolds is found in [3] and in [10] . Let us note also that the proof of the first part of Theorem 1.2 is based on this lemma as well.

Applying Lemma 3.2 to the function $u(y,t) = p(x,y,t)$ we see for any $\tau < t < T$

$$\int_M p^2(x,y,t)e^{\xi(y,t)}dy \leq \int_M p^2(x,y,\tau)e^{\xi(y,\tau)}dy \tag{3.3}$$

whence, using the specific form of the function $d$, it follows that

$$\int_{M\setminus B(x,R)} p^2(x,y,t)dy \leq \int_{B(x,\rho)} p^2(x,y,\tau)\exp\left(-\frac{d(y)^2}{2(T-\tau)}\right) dy + \int_{M\setminus B(x,\rho)} p^2(x,y,\tau)dy$$

where $\rho < R$. Observing here that $d(y)$ under the exponential is at least as much as $R-\rho$ and letting $T \to t+$ we obtain

$$\int_{M\setminus B(x,R)} p^2(x,y,t)dy \le \exp\left(-\frac{(R-\rho)^2}{2(t-\tau)}\right)\int_M p^2(x,y,\tau)dy + \int_{M\setminus B(x,\rho)} p^2(x,y,\tau)dy$$

Since by the semigroup property (1.11) and (1.1)

$$\int_M p^2(x,y,\tau)dy \equiv p(x,x,2\tau) \le \frac{1}{f(2\tau)} \le \frac{1}{f(\tau)} \tag{3.4}$$

it follows that

$$\int_{M\setminus B(x,R)} p^2(x,y,t)dy \le \frac{1}{f(\tau)}\exp\left(-\frac{(R-\rho)^2}{2(t-\tau)}\right) + \int_{M\setminus B(x,\rho)} p^2(x,y,\tau)dy\ . \tag{3.5}$$

Let us arrange now two decreasing sequences $t_k = t2^{-k}$ and $R_k = (\frac{1}{2}+\frac{1}{k+2})R$ where $k = 0,1,2,\dots$ . Obviously, $t_k \to 0$ as $k \to \infty$ and $\frac{1}{2}R < R_k \le R$ for any $k$. We apply the inequality (3.5) for pairs $t_k, t_{k-1}$ and $R_k, R_{k-1}$ in place of $\tau, t$ and, respectively, $\rho, R$ and sum up all such inequalities. Since

$$\int_{M\setminus B(x,R_k)} p^2(x,y,t_k)dy \le \int_{M\setminus B(x,R/2)} p^2(x,y,t_k)dy$$

and the right-hand side integral approaches to 0 as $t_k \to 0$ (which follows from the fact that $p(x,y,t) \to 0$ as $t \to 0$ uniformly in $y \in \Omega \setminus B(x,R/2)$ ) we obtain from (3.5)

$$\int_{M\setminus B(x,R)} p^2(x,y,t)dy \le \sum_{k=0}^{\infty}\frac{1}{f(t_k)}\exp\left(-\frac{(R_{k-1}-R_k)^2}{2(t_{k-1}-t_k)}\right)$$

or, applying (1.16) in the form $f(t) \le \mathrm{const}_A 2^{k\nu} f(t_k)$ in order to estimate $f(t_k)$ via $f(t)$

$$\int_{M\setminus B(x,R)} p^2(x,y,t)dy \le \frac{\mathrm{const}_A}{f(t)}\sum_{k=0}^{\infty}\exp\left(ck - \frac{2^{k-1}}{(k+2)^4}\frac{R^2}{t}\right) \tag{3.6}$$

where $c = \nu\log 2$.

Let us note that for some positive absolute constants $c_1, c_2$ the following inequality holds for all $k \ge 0$

$$\frac{2^{k-1}}{(k+2)^4} > c_1 k + c_2$$

(if $c_1$ is small enough then $c_2$ can be taken 0.006 ). Therefore, putting $X = R^2/t$ we have

$$ck - \frac{2^{k-1}}{(k+2)^4}X < (c - c_1X)k - c_2X \le -ck - c_2X$$

provided

$$X \ge 2\frac{c}{c_1}\ . \tag{3.7}$$

Hence, under assumption (3.7) the sum on the right-hand side in (3.6) does not exceed the geometric series

$$\sum_{k=0}^{\infty} \exp(-ck - c_2 X) = \frac{e^{-c_2 X}}{1 - e^{-c}}$$

and (3.6) acquires the from

$$\int_{M \setminus B(x,R)} p^2(x,y,t)dy \leq \frac{\mathrm{const}_A}{f(t)} \exp\left( -c_2 \frac{R^2}{t} \right)$$

which was to be proved.

Finally, let $R^2/t < 2\frac{c}{c_1}$ , then

$$\int_{M \setminus B(x,R)} p^2(x,y,t)dy \leq \int_M p^2(x,y,t)dy \leq \frac{1}{f(t)} \leq \frac{\mathrm{const}_A}{f(t)} \exp\left( -c_2 \frac{R^2}{t} \right)$$

where we have applied (3.4) and boundedness of $R^2/t$.

Thus, we have proved (3.1) with $D = D_0 \equiv 1/c_2$, and the more so is valid for $D > D_0$. □

Next, we proceed with the proof of the Theorem 1.1. There are two points on which we shall focus our attention. First, how to pass from the integral (3.1) to $E^D$ and second, how to diminish the constant $D$ so that it is arbitrarily close to 2.

First we prove (1.6) for a large $D$. To that end, let us observe that the integral $E^{5D}$ is decomposed into the sum of the integrals

$$\int_{\{2^k R \leq r \leq 2^{k+1} R\}} p^2(x,y,t) \exp\left( \frac{r^2}{5Dt} \right) dy \tag{3.8}$$

for $k = 0, 1, 2, ...$ and of the integral

$$\int_{B(x,R)} p^2(x,y,t) \exp\left( \frac{r^2}{5Dt} \right) dy \tag{3.9}$$

where $R$ is an arbitrary positive number. We estimate the integral (3.9) from above by

$$\frac{1}{f(t)} \exp\left( \frac{R^2}{5Dt} \right)$$

which follows from (3.4) . The integral (3.8) does not exceed the following

$$\begin{aligned} \exp\left( \frac{4^{k+1} R^2}{5Dt} \right) \int_{\{r \geq 2^k R\}} p^2(x,y,t)dy &\leq \frac{\mathrm{const}_A}{f(t)} \exp\left( -\frac{4^k R^2}{Dt} + \frac{4^{k+1} R^2}{5Dt} \right) \\ &= \frac{\mathrm{const}_A}{f(t)} \exp\left( -\frac{4^k R^2}{5Dt} \right) \end{aligned}$$

where we have applied Lemma 3.1 assuming that $D \geq D_0$ . Adding all these inequalities we obtain

$$E^{5D}(x,t) \leq \frac{1}{f(t)} \exp\left( \frac{R^2}{5Dt} \right) + \frac{\mathrm{const}_A}{f(t)} \sum_{k=0}^{\infty} \exp\left( -\frac{4^k R^2}{5Dt} \right)$$

Taking here $R=\sqrt{Dt}$ and denoting $5D$ by $D$ we get that for any $D\geq D_1\equiv 5D_0$

$$E^D\leq \frac{\mathrm{const}_A}{f(t)} \tag{3.10}$$

which coincides with (1.6) .

Finally, let us prove (1.6) for any $D$ such that $2<D<D_1$ . We apply again Lemma 3.2 and following from it the inequality (3.3) , on this occasion with the weight function $\xi(y,t)=\frac{r^2}{2(t+T)}$ for a positive $T$. Given $t$, we choose consecutively $T$ and $\tau$ so that $2(t+T)=Dt$ i.e. $T=(D-2)t/2$ and $2(\tau+T)=D_1\tau$ i.e. $\tau=2T/(D_1-2)=\frac{D-2}{D_1-2}t<t$ . Hence, (3.3) takes the form

$$E^D(x,t)\leq E^{D_1}(x,\tau)$$

whence, applying (3.10) to estimate the right-hand side we get

$$E^D(x,t)\leq \frac{\mathrm{const}_A}{f(\tau)}\ .$$

Since we have by (1.16) $f(\tau)\geq \mathrm{const}_A f((D-2)t)$ the desired estimate (1.6) follows from the inequality above. □

*Proof of Corollary 1.1.* Theorem 1.1 applies under the conditions of this corollary and gives (1.6) , or, using the notation $E_0$ instead of $E^D$

$$E_0(x,t)\leq \frac{\mathrm{const}_A}{f(\delta t)}\ .$$

By Theorem 1.2 we obtain that for any $m\geq 0$

$$E_m(x,t)\leq \frac{\mathrm{const}_A C^m\delta^m}{f_m(\delta t)}$$

where $\delta^m$ in the numerator comes from integration of $f(\delta t)$ in $t$. Let us note that for $m\geq 1$

$$f_m(t)=\int_0^t\frac{(t-\tau)^{m-1}}{(m-1)!}f(\tau)d\tau>\int_{t/2}^t\frac{(t-\tau)^{m-1}}{(m-1)!}f(\tau)\geq f(t/2)\frac{(t/2)^m}{m!}$$

whence it follows that

$$f_m(t)\geq \frac{t^m}{A2^m m!}f(t)$$

and

$$E_m(x,t)\leq \frac{\mathrm{const}_A(2C)^m m!}{t^m f(\delta t)}\ .$$

The analogous inequality holds for the point $y$. Applying finally Theorem 1.3 we obtain (1.13) . The form of the coefficient (1.14) comes from a remark that $2C\delta=\frac{D+16}{D-2}\min(D-2,1)<20$ for any $D>2$ and thereby $2C<20\delta^{-1}$ .

In order to prove (1.15) we apply (1.16) which implies $f(\delta t/2)\geq \mathrm{const}_A\delta^\nu f(t)$ and (1.13) is transformed to

$$\left|\frac{\partial^m p}{\partial t^m}\right|(x,y,t)\leq \frac{\mathrm{const}_{A,m}}{\delta^{\nu+m}t^m\sqrt{f(t)g(t)}}\exp\left(-\frac{r^2}{2Dt}\right)\ .$$

Let us put here $D = 2 + \min(1, \frac{t}{r^2})$. Since $D - 2 \le 1$ it follows that $\delta = D - 2 \le \frac{t}{r^2}$ and we have evidently

$$\frac{r^2}{4t} - \frac{r^2}{2Dt} = \frac{\delta}{4D}\frac{r^2}{t} \le \frac{t}{r^2}\frac{r^2}{4Dt} = \frac{1}{4D} \le \frac{1}{8}$$

and

$$\frac{1}{\delta} = \max(1, \frac{r^2}{t}) < 1 + \frac{r^2}{t}$$

whence (1.15) follows.

## REFERENCES


[1] **Aronson D.G.**, Bounds for the fundamental solution of a parabolic equation, *Bull. of AMS* , **73** (1967) 890-896.

[2] **Chavel I.**, "Eigenvalues in Riemannian geometry" Academic Press, New York, 1984.

[3] **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.

[4] **Davies E.B.**, Explicit constants for Gaussian upper bounds on heat kernels, *Amer. J. Math.* , **109** (1987) 319-334.

[5] **Davies E.B.**, Gaussian upper bounds for the heat kernel of some second-order operators on Riemannian manifolds, *J. Funct. Anal.* , **80** (1988) 16-32.

[6] **Davies E.B., Pang M.M.H.**, Sharp heat kernel bounds for some Laplace operators, *Qurt. J. Math.* , **40** (1989) 281-290.

[7] **Davies E.B.**, "Heat kernels and spectral theory" Cambridge University Press, Cambridge, 1989.

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

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

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

[11] **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.

[12] **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.

[13] **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.

[14] **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.

[15] **Ushakov V.I.**, Stabilization of solutions of the third mixed problem for a second order parabolic equation in a non-cylindric domain, (in Russian) *Matem. Sbornik* , **111** (1980) 95-115. Engl. transl. *Math. USSR Sb.* , **39** (1981) 87-105.

[16] **Varopoulos N.Th.**, Semi-groupes d'opérateurs sur les espaces $L^p$, *C. R. Acad. Sci. Paris Sér. I Math.* , **301** (1985) 865-868.
[17] **Varopoulos N.Th.**, Analysis on Lie groups, *J. Funct. Anal.* , **76** (1988) 346-410.
[18] **Varopoulos N.Th.**, Small time Gaussian estimates of heat diffusion kernel.I. The semigroup technique, *Bull. Sci. Math.(2)* , **113** (1989) no.3, 253-277.