SCHAUDER ESTIMATES FOR DEGENERATE ELLIPTIC
AND PARABOLIC PROBLEMS WITH UNBOUNDED
COEFFICIENTS IN $\mathbb{R}^N$

LUCA LORENZI
Dipartimento di Matematica, Università di Parma
Via M. D’Azeglio 85/A, 43100 Parma, Italy

(Submitted by: Reza Aftabizadeh)

Abstract. We consider a class of second-order degenerate elliptic operators. Continuing the study started in [3], we prove Schauder estimates for the distributional solutions of the nonhomogeneous elliptic equation $\lambda u - Au = f$ and the Cauchy problem $Du = Au + g, u(0, \cdot) = f$.

1. Introduction

In this paper we continue the study of the degenerate elliptic operator

$$\mathcal{A}u(x) = \sum_{i,j=1}^{r} q_{ij}(x) D_{ij}u(x) + \sum_{i,j=1}^{N} b_{ij} x_j D_i u(x), \quad x \in \mathbb{R}^N,$$

started in [3]. The operator $\mathcal{A}$ is a generalization of the well known degenerate Ornstein-Uhlenbeck operator, which has constant coefficients $q_{ij}$. The degenerate Ornstein-Uhlenbeck operator has been extensively studied in [6]. In that paper the author proved uniform estimates for the space derivatives of the associated semigroup that she then used to prove some regularity results for the distributional solutions to elliptic equations and to parabolic problems associated with the operator $\mathcal{A}$. In [3] we have shown that under suitable growth assumptions on the diffusion matrix $Q = (q_{ij})$ and on the rank of the matrix $B$, the Cauchy problem

$$(DCP) \quad \begin{cases} 
D_t u(t, x) = \mathcal{A}u(t, x), & t > 0, \quad x \in \mathbb{R}^N, \\
u(0, x) = f(x), & x \in \mathbb{R}^N,
\end{cases}$$

Accepted for publication: October 2004.

AMS Subject Classifications: 35K65, 35J70, 35B65, 35K15.

Work partially supported by the research projects “Equazioni di evoluzione deterministiche e stocastiche” of the Ministero dell’Istruzione, dell’Università e della Ricerca (M.I.U.R.) and “Equazioni di Kolmogorov” of the Istituto Nazionale di Alta Matematica (INdAM).
with \( f \in C_b(\mathbb{R}^N) \), admits a unique bounded classical solution; i.e., there exists a unique bounded and continuous function \( u : [0, +\infty) \times \mathbb{R}^N \to \mathbb{R} \) which is once continuously differentiable with respect to time and twice continuously differentiable with respect to the space variables in \((0, +\infty) \times \mathbb{R}^N\), and such that \( u(0, \cdot) = f \). This allowed us to define a semigroup of bounded operators in \( C_b(\mathbb{R}^N) \) by setting \( T(t)f = u(t, \cdot) \) for any \( t > 0 \). Moreover, we exploited the behavior near \( t = 0 \) of the sup norm of \( T(t)f \) and of its derivatives up to the third order, when \( f \in C^k_b(\mathbb{R}^N) (k \leq 3) \). The assumptions on \( r, Q \) and \( B \) that we made in \([3]\), and that we still assume in this paper, are the following ones:

1. \( \frac{N}{2} \leq r < N \) and
   \[ \sum_{i,j=1}^{r} q_{ij}(x)\xi_i \xi_j \geq \nu(x)|\xi|^2, \quad \forall \xi \in \mathbb{R}^r, \quad x \in \mathbb{R}^N, \]
for some function \( \nu : \mathbb{R}^N \to \mathbb{R}_+ \) such that \( \nu_0 := \inf_{x \in \mathbb{R}^N} \nu(x) > 0 \); 

2. \( q_{ij} \in C^{3+\delta}_{\text{loc}}(\mathbb{R}^N) \) for any \( i, j = 1, \ldots, r \) and some \( \delta \in (0, 1) \), and there exists a positive constant \( C \) such that
   \[ |D^\alpha q_{ij}(x)| \leq C|x|^{(1-|\alpha|)+} \sqrt{\nu(x)}, \quad \forall x \in \mathbb{R}^N, \quad i, j = 1, \ldots, r, \quad |\alpha| \leq 3; \]

3. the matrix \( B \) can be split into blocks as follows
   \[ B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}, \]
   with \( B_1 \in L(\mathbb{R}^r) \) invertible, \( B_2, B_3 \in L(\mathbb{R}^{N-r}, \mathbb{R}^r) \), \( B_4 \in L(\mathbb{R}^{N-r}) \), and \( \text{rank}(B_3) = N - r \).

Our assumptions allow us to recover all the cases when the diffusion coefficients are bounded, as well as some cases in which they are unbounded. We stress that our results are new also in the case of bounded diffusion coefficients. Indeed, to the author’s knowledge, there are results similar to ours only in the particular case when the matrix \( Q(x) \) converges to a strictly positive definite matrix as \( |x| \) tends to +\( \infty \) (see \([6]\)), whereas we do not need to assume any convergence hypothesis. In this paper, first we prove some useful continuity properties of the semigroup \( \{T(t)\}_{t\geq 0} \). In particular, we show that the semigroup is strong Feller. This means that, for any \( t > 0 \), \( T(t) \) maps the space of all bounded Borel measurable functions \( f : \mathbb{R}^N \to \mathbb{R} \) into \( C_b(\mathbb{R}^N) \). Then, we characterize the domain of its “generator.” It is worth noticing that the semigroup \( \{T(t)\}_{t\geq 0} \) is not strongly continuous in \( C_b(\mathbb{R}^N) \) and, in general, it is neither strongly continuous nor analytic in \( BUC(\mathbb{R}^N) \).
This prevents us from defining the infinitesimal generator of \( \{T(t)\}_{t \geq 0} \) in the usual classical way. Nevertheless, we can still associate a closed operator \( A \) (usually called the weak generator) to the semigroup, which is a generalization of the classical notion of generator of a \( C_0 \)-semigroup. Next, we prove Schauder estimates for the distributional solution to the nonhomogeneous equation

\[
\lambda u - Au = f, \quad \lambda > 0, \tag{1.1}
\]

and for the distributional solution to the parabolic nonhomogeneous Cauchy problem

\[
\begin{aligned}
D_t u(t, x) &= Au(t, x) + g(t, x), \quad t > 0, \ x \in \mathbb{R}^N, \\
u(0, x) &= f(x), \quad x \in \mathbb{R}^N.
\end{aligned} \tag{1.2}
\]

More precisely, we prove the following theorems.

**Theorem A.** Let \( 0 < \theta < 1, \lambda > 0 \). Then, for any \( f \in C^0_{b, \theta} / 3 (\mathbb{R}^N) \) (see Definition 2.3) there exists a function \( u \in C^2_{b, (2+\theta) / 3} (\mathbb{R}^N) \) solving equation (1.1) in the sense of distributions. Moreover, there exists a positive constant \( C \), independent of \( u \) and \( f \), such that

\[
\| u \|_{C^2_{b, (2+\theta) / 3} (\mathbb{R}^N)} \leq C \| f \|_{C^0_{b, \theta} / 3 (\mathbb{R}^N)}; \tag{1.3}
\]

such a function \( u \) is the unique distributional solution to the equation (1.1) which is bounded and continuous in \( \mathbb{R}^N \) and is twice continuously differentiable in \( \mathbb{R}^N \) with respect to the first \( r \) variables, with bounded derivatives.

**Theorem B.** Let \( \theta \in (0, 1), T > 0 \). Further, let \( f \in C^2_{b, (2+\theta) / 3} (\mathbb{R}^N) \) (see Definition 2.3) and let \( g \in C^0_b ([0, T] \times \mathbb{R}^N) \) be such that \( g(t, \cdot) \in C^0_{b, \theta} / 3 (\mathbb{R}^N) \) for any \( t \in [0, T] \) and \( \sup_{t \in [0, T]} \| g(t, \cdot) \|_{C^0_{b, \theta} / 3 (\mathbb{R}^N)} < +\infty \). Then, there exists a function \( u \in C^0_b ([0, T] \times \mathbb{R}^N) \), solving problem (1.2) in the sense of distributions, such that \( u(t, \cdot) \in C^2_{b, (2+\theta) / 3} (\mathbb{R}^N) \) for any \( t \in [0, T] \) and

\[
\sup_{t \in [0, T]} \| u(t, \cdot) \|_{C^2_{b, (2+\theta) / 3} (\mathbb{R}^N)} \leq C \left( \| f \|_{C^2_{b, (2+\theta) / 3} (\mathbb{R}^N)} + \sup_{t \in [0, T]} \| g(t, \cdot) \|_{C^0_{b, \theta} / 3 (\mathbb{R}^N)} \right), \tag{1.4}
\]

for some positive constant \( C \), independent of \( u, f, g \). Moreover, \( u \) is the unique distributional solution to problem (1.2) which is bounded and continuous in \( [0, T] \times \mathbb{R}^N \) and is twice continuously differentiable with respect to the first \( r \) space variables in \( [0, T] \times \mathbb{R}^N \), with bounded derivatives.

Finally, we show that, if the function \( f \) in Theorem A belongs to \( C^0_{b} (\mathbb{R}^N) \) for some \( \theta > 1 / 3 \), then the function \( u \) provided by Theorem A is once continuously differentiable in \( \mathbb{R}^N \) with bounded derivatives. Similarly, we
show that, if in addition to all the assumptions of Theorem B, \( \theta > 1/3 \) and 
\( g(t, \cdot) \in C^\theta_b(\mathbb{R}^N) \) for any \( t \in [0, T] \) with \( \sup_{t \in [0,T]} \| g(t, \cdot) \|_{C^\theta_b(\mathbb{R}^N)} < +\infty \), then 
the function \( u \) provided by Theorem B is once continuously differentiable 
with respect to the time and space variables in \([0, T] \times \mathbb{R}^N\) and all the first-
order space derivatives are bounded in \([0, T] \times \mathbb{R}^N\).

The paper is structured as follows. First, in Section 2 we introduce the 
function spaces we deal with throughout the paper. In Section 3, we recall 
some results from \([3]\) that we need in the sequel. Then, in Section 4, we 
prove some properties of the semigroup, which provide us some fundamental 
tools to prove the Schauder estimates of the forthcoming section. Moreover, 
we prove that the semigroup is strong Feller, and we characterize the domain 
of the weak generator. Finally in Section 5 we prove Theorems A and B. For 
this purpose we provide sharp estimates for the behavior (with respect to \( t \)) 
of the semigroup in the anisotropic spaces \( C^{3\theta, \theta}_b(\mathbb{R}^N) \). The keystone is the 
case where \( \theta = 1 \). To estimate the behavior of the semigroup in \( C^{3,1}_b(\mathbb{R}^N) \), 
we adapt to our situation the classical Bernstein’s method (see \([1]\)). The 
general case is then obtained by an interpolation argument.

Notation. Throughout the paper, for any \( u : \mathbb{R}_+ \times \mathbb{R}^N \to \mathbb{R} \) we indifferently 
write \( u(t, \cdot) \) and \( u(t) \) when we want to stress the dependence of \( u \) on the 
time variable \( t \). Moreover, for any smooth real-valued function \( v \) defined 
on a domain of \( \mathbb{R}^N \), we denote by \( Dv \) the gradient of \( v \) and by \( |Dv(x)| \) 
its Euclidean norm at \( x \). Similarly, by \( D^k v \) (\( k \in \mathbb{N} \)) we denote the vector 
consisting of all the \( k \)-th order derivatives of \( v \), and by \( |D^k v(x)| \) its Euclidean 
norm at \( x \). By \( \mathbb{1} \) we denote the function which is identically equal to 1.

By \( I_k \) we denote the identity matrix in \( \mathbb{R}^k \). For any matrix \( A \), we denote by 
\( A^* \) its transpose matrix. When \( a \) is a vector we denote by \( a^T \) its transpose. 
For any matrix \( A \) we denote by \( \| A \| \) its Euclidean norm. For any symmetric 
matrix \( A \) we denote by \( \lambda_{\text{max}}(A) \) and by \( \lambda_{\text{min}}(A) \), respectively, the maximum 
and the minimum eigenvalue of \( A \). When \( A \) is a square matrix, we denote by 
\( \text{Tr}(A) \) its trace; i.e., the sum of elements on the main diagonal. By \( L(\mathbb{R}^m, \mathbb{R}^n) \) 
we denote the set of all linear operators from \( \mathbb{R}^m \) to \( \mathbb{R}^n \) (or, equivalently, the 
set of all \( n \times m \) matrices). When \( m = n \) we simply write \( L(\mathbb{R}^m) \). Finally, by 
\( a \lor b \) (respectively \( a \land b \)) we denote the maximum (respectively the minimum) 
between \( a \) and \( b \), and we set \( a^+ = a \lor 0 \).

2. Function spaces

In this section we introduce the function spaces we deal with throughout 
this paper. We begin with the following definitions.
Definition 2.1. We denote by $B_b(\mathbb{R}^N)$ the space of all bounded Borel measurable functions $f : \mathbb{R}^N \to \mathbb{R}$. Moreover, for any $k \geq 0$ we denote by $C^k_b(\mathbb{R}^N)$ the space of all functions $f : \mathbb{R}^N \to \mathbb{R}$ which are continuously differentiable up to the $[k]$-th order in $\mathbb{R}^N$ with bounded derivatives and $D^\alpha f$ is Hölder continuous of order $k - [k]$. Here $[k]$ denotes the integer part of $k$. We endow $C^k_b(\mathbb{R}^N)$ with the norm

$$
\|u\|_{C^k_b(\mathbb{R}^N)} = \sum_{|\alpha| \leq [k]} \|D^\alpha f\|_\infty + \sum_{|\alpha| = [k]} [D^\alpha f]_{C_b^{k-[k]}(\mathbb{R}^N)},
$$

where $\| \cdot \|_\infty$ and $[\cdot]_{C_b^{k-[k]}(\mathbb{R}^N)}$ denote, respectively, the sup norm and the Hölder seminorm of order $k - [k]$. Finally, we denote by $C^\infty_b(\mathbb{R}^N)$ the space of all functions $f$ which belong to $C^k_b(\mathbb{R}^N)$ for any $k \in \mathbb{N}$.

Definition 2.2. We denote by $C^{1,2}_{loc}((0, +\infty) \times \mathbb{R}^N)$ the space of all $u : (0, +\infty) \times \mathbb{R}^N \to \mathbb{R}$ which are once continuously differentiable with respect to time and twice continuously differentiable with respect to the space variables in $(0, +\infty) \times \mathbb{R}^N$. For any $\alpha \in (0, 1)$, $C^{1+\alpha/2,2+\alpha}_{loc}((0, +\infty) \times \mathbb{R}^N)$ is the subset of $C^{1,2}_{loc}((0, +\infty) \times \mathbb{R}^N)$ of all functions $u$ such that for any compact set $F \subset (0, +\infty) \times \mathbb{R}^N$, $D_t u, D_x^\beta u (|\beta| \leq 2)$ are $\alpha$ Hölder continuous in $F$ with respect to the parabolic distance $d((t,x),(s,y)) = ((t-s) + |x-y|^2)^{1/2}$.

We now define the anisotropic spaces $C_b^{3\theta,\theta}(\mathbb{R}^N) (\theta \in \mathbb{R}_+).$ To simplify the notation, we identify $\mathbb{R}^N$ with $\mathbb{R}^r \times \mathbb{R}^{N-r}$. Moreover, we denote by $C_b^\theta(\mathbb{R}^N) (\theta \in (0, +\infty))$ the usual Zygmund spaces.

Definition 2.3. For any $\theta \in (0, 1]$ we define the space $C_b^{3\theta,\theta}(\mathbb{R}^N)$ by setting

$$
C_b^{3\theta,\theta}(\mathbb{R}^N) = \left\{ f : \mathbb{R}^r \times \mathbb{R}^{N-r} \to \mathbb{R} : f(\cdot, y) \in C_b^{3\theta}(\mathbb{R}^r) \; \forall y \in \mathbb{R}^{N-r}, \right. \left. \sup_{y \in \mathbb{R}^{N-r}} \|f(\cdot, y)\|_{C_b^{3\theta}(\mathbb{R}^r)} < +\infty, \right. \left. f(x, \cdot) \in C_b^\theta(\mathbb{R}^{N-r}) \; \forall x \in \mathbb{R}^r, \sup_{x \in \mathbb{R}^r} \|f(x, \cdot)\|_{C_b^\theta(\mathbb{R}^{N-r})} < +\infty \right\}
$$

and we norm it by

$$
\|f\|_{C_b^{3\theta,\theta}(\mathbb{R}^N)} = \sup_{y \in \mathbb{R}^{N-r}} \|f(\cdot, y)\|_{C_b^{3\theta}(\mathbb{R}^r)} + \sup_{x \in \mathbb{R}^r} \|f(x, \cdot)\|_{C_b^\theta(\mathbb{R}^{N-r})}.
$$

Similarly, we define and norm the space $C_b^{3\theta,\theta}(\mathbb{R}^N)$, replacing everywhere the Zygmund space $C_b^\alpha$ with the usual Hölder spaces $C_b^\alpha$. When $\theta = 1$ we assume
that the derivatives $D_{ijh}f \ (1 \leq i,j,h \leq r)$ of $f \in C^3_b(R^N)$ are continuous in $R^N$.

Remark 2.4. From the characterization of the Zygmund spaces, it follows that $C^{3\theta,\theta}_b(R^N) = C^{3\theta,\theta}_b(R^N)$ with equivalence of the corresponding norms, if $\theta \in R_+ \setminus N$ is such that $3\theta \notin N$, while $C^{3\theta,\theta}_b(R^N)$ is properly continuously embedded in $C^{3\theta,\theta}_b(R^N)$ otherwise.

Remark 2.5. By interpolation, it is easy to check that if $u \in C^{3\theta,\theta}_b(R^N)$ and $3\theta < 1$, then all the derivatives of $u$, with respect to the first $r$ variables and up to the $[3\theta]$-th order are (bounded and) continuous in $R^N$. Indeed, for any $x \in R^N$ split $x = (y, z)$ where $y = (x_1, \ldots, x_r)$ and $z = (x_{r+1}, \ldots, x_N)$, and let $k \in N$ be such that $k < 3\theta$. Since there exists a positive constant $C$ such that $\|f\|_{C^k_b(R^r)} \leq C\|f\|_{C^{1-k/(3\theta)}_b(R^r)} \|f\|_{C^{k/(3\theta)}_b(R^r)}$, $f \in C^{3\theta}_b(R^N)$, (see [4, Proposition 1.1.3]), we easily deduce that

$$\|u(\cdot, z_2) - u(\cdot, z_1)\|_{C^k_b(R^r)} \leq C\|u(\cdot, z_2) - u(\cdot, z_1)\|_{C^{1-k/(3\theta)}_b(R^r)} \|u(\cdot, z_2) - u(\cdot, z_1)\|_{C^{k/(3\theta)}_b(R^r)}$$

$$\leq 2C\|u\|_{C^{3\theta,k,\theta-k/3}_b(R^N)} \|y_2 - y_1\|^{\theta-k/3}, \quad (2.1)$$

for any $z_1, z_2 \in R^{N-r}$. From (2.1) we deduce that the $k$-th order derivatives of $u$ with respect to the first $r$ variables belong to $C^{3\theta-k,\theta-k/3}_b(R^N)$, so, in particular, they are continuous in $R^N$.

3. SOME RESULTS FROM [3]

In this section we recall all the results from [3] that we need in what follows. For this purpose, for any $\varepsilon > 0$ we introduce the (uniformly) elliptic operator $A_\varepsilon$, defined on smooth functions by

$$A_\varepsilon u = \sum_{i,j=1}^r q_{ij}D_{ij}u + \varepsilon \sum_{j=r+1}^N D_{jj}u + \sum_{i,j=1}^N b_{ij}x_jD_iu. \quad (3.1)$$

The following proposition provides us a very useful maximum principle for both the operators $A_\varepsilon$ and $A$.

Proposition 3.1. Suppose that assumptions H1–H3 hold true and let $u : [0, T] \times R^N \rightarrow R \ (T > 0)$ be a bounded classical solution of the Cauchy
Schauder estimates for degenerate elliptic and parabolic problems

\[
\begin{aligned}
D_t u(t, x) &= A_\varepsilon u(t, x) + g(t, x), \quad t \in (0, T), \ x \in \mathbb{R}^N, \\
u(0, x) &= f(x), \quad x \in \mathbb{R}^N,
\end{aligned}
\]

where \( f \in C_b(\mathbb{R}^N) \), \( g \in C((0, T) \times \mathbb{R}^N) \) and \( \varepsilon \geq 0 \). If \( g(t, x) \leq 0 \) for any \((t, x) \in (0, T) \times \mathbb{R}^N\), then

\[
\sup_{x \in \mathbb{R}^N} u(t, x) \leq \sup_{x \in \mathbb{R}^N} f(x), \quad t \in [0, T].
\]

(3.2)

In particular, if \( g \equiv 0 \), then

\[
\|u(t, \cdot)\|_{\infty} \leq \|f\|_{\infty}, \quad t \in [0, T].
\]

(3.3)

Proof. See [3, Proposition 2.7 and Remark 3.3].

Let us now introduce the semigroup \( \{T_\varepsilon(t)\}_{t \geq 0} \) associated with the operator \( A_\varepsilon \), defined as follows: for any \( f \in C_b(\mathbb{R}^N) \) and any \( t > 0 \), \( T_\varepsilon(t)f = u(t, \cdot) \), where \( u \) is the unique classical solution to the Cauchy problem

\[
\begin{aligned}
D_t u(t, x) &= A_\varepsilon u(t, x), \quad t > 0, \ x \in \mathbb{R}^N, \\
u(0, x) &= f(x), \quad x \in \mathbb{R}^N.
\end{aligned}
\]

(3.4)

The existence of a classical solution to (3.4) follows from [7, Theorems 4.2 and 4.5], while its uniqueness follows from Proposition 3.1. The next proposition is concerned with the behavior of \( T_\varepsilon(t)f \) and its space derivatives as \( t \) tends to 0, when \( f \in C^k_b(\mathbb{R}^N) \) (\( k \in \mathbb{N}, k \leq 3 \)).

Proposition 3.2. For any \( \varepsilon > 0 \), any \( f \in C^k_b(\mathbb{R}^N) \) (\( k = 0, \ldots, 3 \)) and any \( j = 0, \ldots, 3 \), the function \((t, x) \mapsto t^{(j-k)+/2}(D^j T_\varepsilon(t)f)(x)\) is continuous in \([0, +\infty) \times \mathbb{R}^N\). Moreover,

\[
\lim_{t \to 0^+} t^{(j-k)+/2}(D^j T_\varepsilon(t)f)(x) = 0, \ x \in \mathbb{R}^N, \ j = 0, \ldots, 3.
\]

Further, there exists a positive constant \( C = C(T, k, \varepsilon) \) such that

\[
\sum_{j=0}^{3} t^{(j-k)+/2} \|D^j T_\varepsilon(t)f\|_{\infty} \leq C \|f\|_{C^k(\mathbb{R}^N)}, \quad t \in (0, T].
\]

Proof. See [3, Theorems 2.8 and 2.13].

In the following proposition we recall the main properties of the semigroup \( \{T(t)\}_{t \geq 0} \) that we need in what follows.
Proposition 3.3. For any $f \in C_b(\mathbb{R}^N)$ there exists a unique solution $u$ to the Cauchy problem \((DCP)\) belonging to
\[ C_b([0, +\infty) \times \mathbb{R}^N) \cap C^{1+\alpha/2,2+\alpha}_{\text{loc}}((0, +\infty) \times \mathbb{R}^N) \]
for any $\alpha \in (0, 1)$. The family of bounded operators $\{T(t)\}_{t \geq 0}$ defined by $T(t)f = u(t, \cdot)$ for any $t > 0$, where $u$ is as above, is a contraction semigroup of linear operators in $C_b(\mathbb{R}^N)$. Moreover, $T_\varepsilon(\cdot)f$ converges to $T(\cdot)f$ in $C^{1,2}(F)$ as $\varepsilon$ tends to 0 for any compact set $F \subset (0, +\infty) \times \mathbb{R}^N$. Further, $T(t)f \in C^2_b(\mathbb{R}^N)$ for any $t > 0$ and any $f \in C_b(\mathbb{R}^N)$, and for any $\omega > 0$ there exists a positive constant $C = C(\omega)$ such that
\[ \|T(t)f\|_{C^k_b(\mathbb{R}^N)} \leq C e^{-\frac{\omega t}{t^{3/2}}} \|f\|_{C^k_b(\mathbb{R}^N)}, \quad t > 0, \]  
and
\[ \|D^\alpha T(t)f\|_{\infty} \leq C e^{-\frac{\omega t}{t^{k/2}}} \|f\|_{\infty}, \quad t > 0, \]
for any multindex $\alpha = (\alpha_1, \ldots, \alpha_N)$ with length $k = 1, 2, 3$ and such that $\alpha_j = 0$ for any $j > r$.

Proof. See [3, Theorem 3.4].

Remark 3.4. The maximum principles in Proposition 3.1 implies that for any $\varepsilon \geq 0$ and any $t > 0$, $T_\varepsilon(t) (T_0(t) := T(t))$ is an order preserving semigroup; i.e., $T_\varepsilon(t)f_1 \leq T_\varepsilon(t)f_2$ whenever $f_1 \leq f_2$.

As far as the convergence of $T(t)f$ to $f$ as $t$ tends to 0 is concerned, in [3, Theorem 3.4] we have proved the following result.

Proposition 3.5. For any function $f \in C_b(\mathbb{R}^N)$ with compact support, $T(t)f$ tends to $f$ uniformly in $\mathbb{R}^N$ as $t$ tends to 0.

To conclude this section, we recall the following lemma, which will play a crucial role in the proof of the estimates of Theorem 5.3.

Lemma 3.6. Suppose that $Q = (q_{i,j})A \in L(\mathbb{R}^N)$ are two positive definite matrices. Further, assume that the submatrix $Q_0 = (q_{i,j})_{i,j=1}^r$ is strictly positive definite and $q_{ij} = 0$ if $i \lor j > r$. Then
\[ \text{Tr}(QA) \geq \lambda_{\min}(Q_0) \text{Tr}(A_1), \]
where $A_1$ is the submatrix obtained from $A$ by erasing the last $N - r$ rows and lines.
Taking Remark 3.4 into account, we deduce that 

\[ g \text{setting } g \text{3.1, it suffices to show that } g \text{to show that } T \text{for any compact set } F \text{c} \]

Without losing in generality throughout the proof, we assume that

\[ \text{Proof. See [3, Lemma 2.6].} \]

4. SOME PROPERTIES OF THE SEMIGROUP \( T(t) \)

In this section we first prove a continuity property of the semigroup \( \{T(t)\}_{t \geq 0} \) that will play a fundamental role in order to prove the Schauder estimates of Section 5. Then, we prove that \( \{T(t)\}_{t \geq 0} \) is strong Feller and, finally, we characterize the domain of the weak generator of the semigroup.

To begin with, we give the following definition.

**Proposition 4.1.** Let \( \{f_n\}_{n \in \mathbb{N}} \subset C_b(\mathbb{R}^N) \) be a bounded sequence of continuous functions converging to \( f \in C_b(\mathbb{R}^N) \) locally uniformly in \( \mathbb{R}^N \). Then \( T(\cdot)f_n \) converges to \( T(\cdot)f \) locally uniformly in \( [0, +\infty) \times \mathbb{R}^N \) and in \( C^{1,2}(F) \) for any compact set \( F \subset (0, +\infty) \times \mathbb{R}^N \).

**Proof.** Without losing in generality throughout the proof, we assume that \( f = 0 \) and \( \sup_{n \in \mathbb{N}} \|f_n\|_\infty \leq 1 \). As a first step, we show that \( T(\cdot)f_n \) tends to \( T(\cdot)f \) in \( C^{1,2}(F) \) for any compact set \( F \subset (0, +\infty) \times \mathbb{R}^N \). Since the function \( T(\cdot)f_n \) solves the differential equation \( Du - Au = 0 \) in \( (0, +\infty) \times \mathbb{R}^N \), by estimate (3.5) it follows that the sequence \( \{T(\cdot)f_n\}_{n \in \mathbb{N}} \) is bounded in \( C^{1+\alpha/2,2+\alpha}(K) \) for any compact set \( K \subset (0, +\infty) \times \mathbb{R}^N \) and any \( \alpha \in (0, 1) \). Therefore, the Ascoli-Arzelà theorem implies that there exists a subsequence \( \{T(\cdot)f_{n_k}\}_{k \in \mathbb{N}} \) converging in \( C^{1,2}(F) \), for any \( F \) as above, to some function \( g \), which, of course, solves the differential equation \( Du - Au = 0 \). We are going to show that \( g(t, \cdot) = T(t)f \) for any \( t > 0 \). For this purpose, by Proposition 3.1, it suffices to show that \( g \) can be extended by continuity at \( t = 0 \) by setting \( g(0, \cdot) = f \). So, let \( E \subset \mathbb{R}^N \) be a compact set and take any smooth function \( \eta \) compactly supported in \( \mathbb{R}^N \) such that \( \eta = 1 \) in \( E \) and \( 0 \leq \eta \leq 1 \).

Taking Remark 3.4 into account, we deduce that

\[
|T(t)f_{n_k}(x) - (T(t)(\eta f_{n_k}))(x)| = |T(t)((1-\eta)f_{n_k}))(x)| \\
\leq \|f_{n_k}\|_\infty \|1 - (T(t)\eta)(x)| \leq \|1 - (T(t)\eta)(x)|,
\]

(4.1)

for any \( t > 0 \) and any \( x \in \mathbb{R}^N \). Since \( f_{n_k} \) converges locally uniformly in \( \mathbb{R}^N \) as \( k \) tends to \( +\infty \), then the function \( T(t)(\eta f_{n_k}) \) converges uniformly in \( \mathbb{R}^N \). Letting \( k \) go to \( +\infty \) in (4.1), we get

\[
|g(t,x) - (T(t)(\eta f))(x)| \leq |1 - (T(t)\eta)(x)|, \quad t > 0, \ x \in \mathbb{R}^N.
\]

Since both \( \eta \) and \( \eta f \) are compactly supported in \( \mathbb{R}^N \), then, by Proposition 3.5, \( T(t)\eta \) and \( T(t)(\eta f) \) tend, respectively, to \( \eta \) and \( \eta f \) uniformly in \( \mathbb{R}^N \), as \( t \) tends to \( 0 \). Therefore, from (4.1) we deduce that \( g \) tends to \( f \) as \( t \) tends to \( 0 \), uniformly in \( E \). By the arbitrariness of \( E \), it follows that \( g \) can
be extended by continuity at \( t = 0 \) setting \( g(0, \cdot) = f \). Therefore, we have proved that \( T(\cdot) f_{n_k} \) converges to \( T(\cdot) f \) in \( C^{1,2}(F) \) as \( k \) tends to \(+\infty\), for any compact set \( F \subset (0, +\infty) \times \mathbb{R}^N \). Repeating the same arguments as above we can show that any sequence \( \{T(\cdot) f_{n_k}\}_{k \in \mathbb{N}} \) admits a subsequence \( \{T(t) f_{n_k}\}_{n \in \mathbb{N}} \) converging to \( T(\cdot) f \) in \( C^{1,2}(F) \). This implies that, actually, all the sequence \( \{T(\cdot) f_n\}_{n \in \mathbb{N}} \) converges in \( C^{1,2}(F) \) to \( T(\cdot) f \) for any compact set \( F \subset (0, +\infty) \times \mathbb{R}^N \). To conclude the proof, let us now check that \( T(\cdot) f_n \) converges to \( T(\cdot) f \) locally uniformly in \([0, +\infty) \times \mathbb{R}^N \). For this purpose, we adapt to our situation the proof given in [7, Proposition 4.6], adding more details. Let us fix \( R, T > 0 \) and prove that \( T(\cdot) f_n \) tends to \( T(\cdot) f \) uniformly in \([0, T] \times B(0, R) \). For any \( n \in \mathbb{N} \), let \( \varphi_n \in C_b(\mathbb{R}^N) \) be any nonnegative function such that \( \chi_{B(0,n-1)} \leq \varphi_n \leq \chi_{B(0,n)} \), where \( \chi_A \) denotes the characteristic function of the set \( A \). Moreover, for any \( \varepsilon > 0 \), let \( M_\varepsilon \) be the set defined by

\[
M_\varepsilon = \left\{ s \geq 0 : \exists n \in \mathbb{N} \text{ such that } \inf_{(t,x) \in [0,s] \times B(0,R)} (T(t) \varphi_n)(x) \geq 1 - \varepsilon \right\}.
\]

We claim that \( M_\varepsilon = [0, +\infty) \), for any \( \varepsilon > 0 \). To show this, we prove that \( M_\varepsilon \) is both a nonempty open and closed set in \([0, +\infty) \). Of course, \( M_\varepsilon \) is nonempty since \( 0 \in M_\varepsilon \). We first show that \( M_\varepsilon \) is closed. For this purpose, let \( \{s_n\}_{n \in \mathbb{N}} \subset M_\varepsilon \) converge to some point \( s \in [0, +\infty) \). We are going to prove that \( s \in M_\varepsilon \). Of course, if \( s = 0 \) we are done. So let us assume that \( s > 0 \). Then, \( s_n > 0 \) for some \( n_0 \in \mathbb{N} \) and, without losing in generality, we can assume that \( n_0 = 1 \). Since \( \{s_n\}_{n \in \mathbb{N}} \subset M_\varepsilon \), then for any \( n \in \mathbb{N} \) there exists \( k_n \in \mathbb{N} \) such that

\[
\inf_{(t,x) \in [0,s_n] \times B(0,R)} (T(t) \varphi_{k_n})(x) \geq 1 - \varepsilon.
\]

Moreover, since \( \varphi_n \leq \varphi_{n+1} \) for any \( n \in \mathbb{N} \), then \( \{T(t) \varphi_n\}_{n \in \mathbb{N}} \) is an increasing sequence. Hence, without losing in generality, we can assume that \( k_n \) is increasing to \(+\infty\). Suppose that \( s \leq s_1 \). Then

\[
\inf_{(t,x) \in [0,s] \times B(0,R)} (T(t) \varphi_{k_1})(x) \geq \inf_{(t,x) \in [0,s_1] \times B(0,R)} (T(t) \varphi_{k_1})(x) \geq 1 - \varepsilon,
\]

so that \( s \in M_\varepsilon \). Suppose now that \( s_1 < s \). Since \( \{\varphi_n\}_{n \in \mathbb{N}} \) converges to \( 1 \) locally uniformly in \( \mathbb{R}^N \), then, by the above results, \( T(\cdot) \varphi_n \) converges to \( 1 \) uniformly in \([s_1, s] \times B(0,R) \). Hence,

\[
\lim_{n \to +\infty} \inf_{(t,x) \in [s_1, s] \times B(0,R)} (T(t) \varphi_{k_n})(x) = 1,
\]
and, therefore, there exists \( n_0 \in \mathbb{N} \) such that
\[
\inf_{(t,x) \in [s_1, s] \times B(0,R)} (T(t)\varphi_{k_{n_0}})(x) \geq 1 - \varepsilon. \tag{4.2}
\]

Since
\[
\inf_{(t,x) \in [0, s_1] \times B(0,R)} (T(t)\varphi_{k_{n_0}})(x) \geq \inf_{(t,x) \in [0, s_1] \times B(0,R)} (T(t)\varphi_{k_1})(x) \geq 1 - \varepsilon, \tag{4.3}
\]
from (4.2) and (4.3) we deduce that
\[
\inf_{(t,x) \in [0, s] \times B(0,R)} (T(t)\varphi_{k_{n_0}})(x) \geq 1 - \varepsilon.
\]

It follows that \( s \in M_\varepsilon \) also in this case. Therefore, \( M_\varepsilon \) is closed. To show that \( M_\varepsilon \) is open in \([0, +\infty)\), we fix \( s \in M_\varepsilon \). Then, there exists \( n \in \mathbb{N} \) such that
\[
(T(t)\varphi_n)(x) \geq 1 - \varepsilon, \quad (t, x) \in [0, s] \times \overline{B(0,R)}.
\]

Since \( T(s)\varphi_n \to \mathbb{1} \) as \( n \to +\infty \), uniformly on compact subsets of \( \mathbb{R}^N \), possibly replacing \( n \) with a bigger integer, we can assume that
\[
(T(s)\varphi_n)(x) \geq 1 - \varepsilon/2 \quad \text{for every } x \in \overline{B(0,R)}.
\]
Moreover, if \( \delta_0 \) is sufficiently small, then \( (T(s + \delta)\varphi_n)(x) \geq 1 - \varepsilon \) for any \( (t, x) \in [s, s + \delta_0] \times \overline{B(0,R)} \). Indeed, for any \( \delta > 0 \)
\[
T(s + \delta)\varphi_n = T(s)\varphi_n + (T(s + \delta)\varphi_n - T(s)\varphi_n)
\geq T(s)\varphi_n - \|T(s)\|_{L(C_b(\mathbb{R}^N))}\|T(\delta)\varphi_n - \varphi_n\|_{\infty}
\geq 1 - \varepsilon/2 - \|T(\delta)\varphi_n - \varphi_n\|_{\infty},
\]
and, by Proposition 3.5, \( \|T(\delta)\varphi_n - \varphi_n\|_{\infty} \) vanishes as \( \delta \) tends to 0. Hence, if \( \delta_0 \) is sufficiently small, then \( (T(t)\varphi_n)(x) \geq 1 - \varepsilon \) for any \( (t, x) \in [0, s + \delta_0] \times \overline{B(0,R)} \). This implies that \( M_\varepsilon \) is open in \([0, +\infty)\). Being both an open and closed subset of \([0, +\infty)\), it follows that \( M_\varepsilon = [0, +\infty) \). Therefore, since
\[
(T(t)(\mathbb{1} - \varphi_m))(x) \leq \varepsilon, \quad (t, x) \in [0, T] \times \overline{B(0,R)}.
\]
Moreover, since the sequence \( \{\varphi_m\}_{m \in \mathbb{N}} \) is increasing, we can assume that \( m > R \). Now we are almost done. Indeed, splitting \( T(t)f_n = T(t)(f_n\varphi_m) + T(t)(f_n(\mathbb{1} - \varphi_m)) \), we get
\[
\sup_{(t,x) \in [0,T] \times \overline{B(0,R)}} |(T(t)f_n)(x)| \leq \sup_{(t,x) \in [0,T] \times \overline{B(0,R)}} |(T(t)(f_n\varphi_m))(x)|
+ \sup_{(t,x) \in [0,T] \times \overline{B(0,R)}} |(T(t)(f_n(\mathbb{1} - \varphi_m)))(x)|
\leq \|T(t)(f_n\varphi_m)\|_{\infty} + \|f_n\|_{\infty} \sup_{(t,x) \in [0,T] \times \overline{B(0,R)}} |(T(t)(\mathbb{1} - \varphi_m))(x)|
\]
\[ \leq \sup_{x \in B(0,m)} |f_n(x)| + \varepsilon \|f_n\|_\infty, \]

and the assertion follows letting \( n \) go to \(+\infty\), since \( \|f_n\|_\infty \leq K \) for any \( n \in \mathbb{N} \) and some positive constant \( K \).

**Corollary 4.2.** There exists a family of probability Borel measures \( \{p(t,x,dy)\} \) such that

\[ (T(t)f)(x) = \int_{\mathbb{R}^N} f(y)p(t,x,dy), \quad t > 0, \ x \in \mathbb{R}^N, \] (4.4)

for any \( f \in C_b(\mathbb{R}^N) \). It follows that if \( \{f_n\} \in C_b(\mathbb{R}^N) \) is a bounded sequence converging pointwise to \( f \in C_b(\mathbb{R}^N) \), then \( T(\cdot)f_n \) tends to \( T(\cdot)f \) pointwise.

**Proof.** We begin the proof observing that, for any fixed \( t > 0 \) and any \( x \in \mathbb{R}^N \), the operator \( f \mapsto (T(t)f)(x) \) is a linear functional on \( C_0(\mathbb{R}^N) \) (the space of all the functions \( f \in C_b(\mathbb{R}^N) \) vanishing at infinity), then by Riesz’s representation theorem, there exists a family of positive Borel measures \( \{p(t,x,\cdot)\}_{t \geq 0, \ x \in \mathbb{R}^N} \) such that (4.4) holds true for any \( f \in C_0(\mathbb{R}^N) \).

Such measures are finite for any \( t > 0 \) and any \( x \in \mathbb{R}^N \). Indeed, for any \( n \in \mathbb{N} \), let \( \varphi_n \in C_0(\mathbb{R}^N) \) be such that \( \chi_{B(0,n)} \leq \varphi_n \leq \chi_{B(0,n+1)} \). Then, from (4.4) and the positivity of the semigroup, we deduce that

\[ 1 \geq (T(t)\varphi_n)(x) = \int_{\mathbb{R}^N} \varphi_n(y)p(t,x,dy) \geq \int_{B(0,n)} p(t,x,dy). \]

Letting \( n \) go to \(+\infty\) yields \( p(t,x,\mathbb{R}^N) \leq 1 \). Now, with any \( f \in C_b(\mathbb{R}^N) \), we associate a bounded sequence \( \{f_n\} \in C_0(\mathbb{R}^N) \) converging to \( f \) locally uniformly as \( n \) tends to \(+\infty\). Writing (4.4) with \( f \) being replaced by \( f_n \) and letting \( n \) go to \(+\infty\), from Proposition 4.1 and the dominated convergence theorem, we conclude that (4.4) holds for any \( f \in C_b(\mathbb{R}^N) \). Finally, since \( T(\cdot)1 = 1 \), we easily see that for any \( t > 0 \) and any \( x \in \mathbb{R}^N \), \( p(t,x,\cdot) \) are probability measures.

Using formula (4.4) we can show that the semigroup \( \{T(t)\}_{t \geq 0} \) is strong Feller (i.e., it can be extended to a semigroup of bounded operators in \( B_b(\mathbb{R}^N) \), still denoted by \( \{T(t)\}_{t \geq 0} \), such that \( T(t) \) maps \( B_b(\mathbb{R}^N) \) into \( C_b(\mathbb{R}^N) \) for any \( t > 0 \)).

**Proposition 4.3.** The semigroup \( \{T(t)\}_{t \geq 0} \) is strong Feller.

**Proof.** The proof is similar to that of [2, Theorem 3.2]. Nevertheless, for the reader’s convenience we go into details. First of all we observe that using (4.4) we can extend the semigroup \( \{T(t)\}_{t \geq 0} \) to \( B_b(\mathbb{R}^N) \) by setting, for any
for any \( t > 0 \) and any \( x \in \mathbb{R}^N \), \( (T(t)f)(x) := \lim_{n \to +\infty}(T(t)f_n)(x) \), where \( \{f_n\}_{n \in \mathbb{N}} \subset C_b(\mathbb{R}^N) \) is any bounded sequence converging pointwise to \( f \). To show that the so extended semigroup is strong Feller, we fix \( f, f_n \) as above. By (3.5), with \( k = 1 \), we have
\[
\|T(t)f_n - T(t)f_m\|_{C^1_b(\mathbb{R}^N)} \leq C(t)\|f_n - f_m\|_{\infty} \leq \tilde{C}(t)\|f\|_{\infty},
\]
for any \( n, m \in \mathbb{N} \) and any \( t > 0 \), where \( C(\cdot) \) and \( \tilde{C}(\cdot) \) are positive functions. The Ascoli-Arzelà theorem implies that there exists a subsequence \( T(t)f_{n_k} \) converging locally uniformly in \( \mathbb{R}^N \) to a function \( v_t \in \operatorname{Lip}(\mathbb{R}^N) \), as \( k \) tends to \( +\infty \). Since \( T(t)f_n \) converges to \( T(t)f \) pointwise as \( n \) tends to \( +\infty \), we deduce that \( v_t(x) = (T(t)f)(x) \) for any \( t > 0 \) and any \( x \in \mathbb{R}^N \). This finishes the proof.

We now observe that since \( \{T(t)\}_{t \geq 0} \) is a contraction semigroup, then for any \( \lambda > 0 \) we can define the linear operator \( R(\lambda) \in L(C_b(\mathbb{R}^N)) \) by setting
\[
(R(\lambda)f)(x) = \int_0^{+\infty} e^{-\lambda t}(T(t)f)(x)dt, \quad x \in \mathbb{R}^N.
\]
As we can immediately check, \( R(\lambda) \) satisfies the resolvent identity. Moreover, the uniqueness of the real Laplace transform implies that \( R(\lambda) \) is one to one. Therefore, by a classic result (see, e.g., [8, Theorem VIII.4.1]), \( \{R(\lambda)\}_{\lambda > 0} \) is the resolvent family associated to some closed operator \( A : D(A) \subset C_b(\mathbb{R}^N) \to C_b(\mathbb{R}^N) \). From now on, we write \( R(\lambda, A) \) instead of \( R(\lambda) \). In next proposition we characterize \( D(A) \) and show that \( Au = \mathcal{A}u \) for any \( u \in D(A) \), where \( \mathcal{A}u \) is meant in the sense of distributions. This result is the keystone in order to prove the forthcoming Proposition 4.5. For this purpose, we first need to show the following lemma.

**Lemma 4.4.** For any \( f \in C^2_b(\mathbb{R}^N) \) such that \( \mathcal{A}f \in C^2_b(\mathbb{R}^N) \) it holds that
\[
T(t)\mathcal{A}f = \mathcal{A}T(t)f, \quad t > 0.
\]  

**Proof.** To prove (4.5) we first show that
\[
\mathcal{T}_\varepsilon(t)\mathcal{A}_\varepsilon f = \mathcal{A}_\varepsilon \mathcal{T}_\varepsilon(t)f, \quad \forall t > 0,
\]
for any \( \varepsilon > 0 \) and any \( f \in C^2_b(\mathbb{R}^N) \) such that \( \mathcal{A}f \in C_b(\mathbb{R}^N) \). Here \( \mathcal{A}_\varepsilon \) is given by (3.1). For this purpose, we recall that in [9, Proposition 3.2] the author proves that if \( g \) belongs to the domain of the weak generator \( \mathcal{A}_\varepsilon \) of the semigroup \( \{\mathcal{T}_\varepsilon(t)\}_{t \geq 0} \), i.e., to the set
\[
D(\mathcal{A}_\varepsilon) = \left\{ f \in C_b(\mathbb{R}^N) : \sup_{t > 0} \left\| \frac{T_\varepsilon(t)f - f}{t} \right\| < +\infty, \right\}
\]
Then $A_{\varepsilon}$ commutes with $T(t)$ on $D(A_{\varepsilon})$ for any $t > 0$. By [7, Proposition 3.5(i) and Theorem 3.7], we deduce that $D(A_{\varepsilon}) = D_{\max}(A_{\varepsilon}) := \{g \in C_b(\mathbb{R}^N) \cap \bigcap_{1 < p < +\infty} W^{2,p}_{\text{loc}}(\mathbb{R}^N) : A_{\varepsilon}g \in C_b(\mathbb{R}^N)\}$, and $A_{\varepsilon}g = A_{\varepsilon}f$ for any $g \in D_{\max}(A_{\varepsilon})$. Since $f \in D_{\max}(A_{\varepsilon})$ for any $\varepsilon > 0$, then (4.6) follows. Now we are almost done. Indeed, a straightforward computation shows that $A_{\varepsilon}f$ converges uniformly in $\mathbb{R}^N$ to $Af$ as $\varepsilon$ tends to 0. Moreover, by Proposition 3.3, $T_{\varepsilon}(t)f$ converges to $T(t)f$ in $C^2(K)$, as $\varepsilon$ tends to 0, for any $t > 0$. Therefore, since $\{T_{\varepsilon}(t)\}_{\varepsilon \geq 0}$ is a contraction semigroup (see (3.3)), we get

$$
\|T_{\varepsilon}(t)A_{\varepsilon}f - T(t)Af\|_{C(K)} \leq \|T_{\varepsilon}(t)(A_{\varepsilon}f - Af)\|_{C(K)} + \|(T_{\varepsilon}(t) - T(t))Af\|_{C(K)}
$$

and the last side of the previous chain of inequalities vanishes as $\varepsilon$ tends to 0. This implies that $T_{\varepsilon}(t)A_{\varepsilon}f$ tends to $T(t)Af$, locally uniformly in $\mathbb{R}^N$. Similarly, for any $t > 0 A_{\varepsilon}T_{\varepsilon}(t)f$ tends to $AT(t)f$ locally uniformly in $\mathbb{R}^N$ as $\varepsilon$ tends to 0. Hence, taking the limit as $\varepsilon$ tends to 0 in (4.6), we get (4.5).

We are now in a position to prove the following proposition.

**Proposition 4.5.** The following characterization holds true:

$$
D(A) = \left\{ f \in C_b(\mathbb{R}^N) : \exists \{f_n\}_{n \in \mathbb{N}} \subset C_b^2(\mathbb{R}^N), \exists g \in C_b(\mathbb{R}^N) : f_n \to f, \quad Af_n \to g \text{ locally uniformly in } \mathbb{R}^N \right. \\
\left. \quad \text{and } \sup_{n \in \mathbb{N}} (\|f_n\|_\infty + \|A f_n\|_\infty) < +\infty \right\}.
$$

Moreover, $Af = Af$ for any $f \in D(A)$. Here and above, $Af$ is meant in the sense of distributions.

**Proof.** We adapt to our situation the technique of [6, Theorem 6.2]. For this purpose, let us denote by $A_0$ the realization of the operator $A$ in $C_b(\mathbb{R}^N)$ with domain $D(A_0) = \{f \in C_b^2(\mathbb{R}^N) : Af \in C_b(\mathbb{R}^N)\}$. Taking Lemma 4.4 into account, it is easy to check that for any $\lambda > 0$ it holds that

$$
(R(\lambda, A)f)(x) = \int_0^{+\infty} e^{-\lambda t} (T(t)Af)(x) dt = \int_0^{+\infty} e^{-\lambda t} (AT(t)f)(x) dt
$$

$$
= \int_0^{+\infty} e^{-\lambda t} \left( \frac{\partial}{\partial t} T(t)f \right)(x) dt = -f(x) + \lambda (R(\lambda, A)f)(x), \quad x \in \mathbb{R}^N.
$$

(4.8)
Therefore, from (4.8) we deduce that \( f = R(\lambda, A)(\lambda f - Af) \) and, consequently, \( f \in D(A) \) and \( Af = Af \). Let now \( f \in D_0 \) (the function space defined by the right-hand side of (4.7)). By the above results we know that

\[
f_n = R(\lambda, A)(\lambda f_n - Af_n), \quad n \in \mathbb{N}, \quad \lambda > 0.
\]

(4.9)

Let us now show that for any bounded sequence \( \{h_n\}_{n \in \mathbb{N}} \subseteq C_b(\mathbb{R}^N) \) converging locally uniformly in \( \mathbb{R}^N \) to some function \( h \in C_b(\mathbb{R}^N) \), then \( R(\lambda, A)h_n \) converges to \( R(\lambda, A)h \), locally uniformly in \( \mathbb{R}^N \), as well. For this purpose, we observe that for any compact set \( K \subseteq \mathbb{R}^N \),

\[
\|R(\lambda, A)(h_n - h)\|_{C(K)} \leq \int_0^{+\infty} e^{-\lambda t}\|T(t)(h_n - h)\|_{C(K)}dt.
\]

By Proposition 4.1, \( \|T(t)(h_n - h)\|_{C(K)} \) converges pointwise in \([0, +\infty)\) to 0 as \( n \) tends to +\( \infty \). Moreover,

\[
\|T(t)(h_n - h)\|_{C(K)} \leq \|T(t)(h_n - h)\|_{\infty} \leq \|h_n - h\|_{\infty} \leq \|h_n\| + \|h\| \leq C,
\]

for any \( n \in \mathbb{N} \) and some positive constant \( C \), independent of \( n \). A straightforward application of the dominated convergence theorem shows that \( R(\lambda, A)h_n \) converges to \( R(\lambda, A)h \), locally uniformly in \( \mathbb{R}^N \). Hence, letting \( n \) go to +\( \infty \) in (4.9) gives \( f = R(\lambda, A)(\lambda f - g) \) and, consequently, \( f \in D(A) \) and \( g = Af \). Moreover, for any \( \varphi \in C^\infty_b(\mathbb{R}^N) \), with compact support, we have

\[
\int_{\mathbb{R}^N} Af_n \varphi dx = \int_{\mathbb{R}^N} f_n A^* \varphi dx, \quad n \in \mathbb{N},
\]

(4.10)

where by \( A^* \) we denote the adjoint of the operator \( A \); i.e., the operator formally defined by

\[
A^* \psi(x) = \sum_{i,j=1}^{r} D_{ij}(q_{ij} \psi)(x) - \sum_{i,j=1}^{N} b_{ij}x_j D_i \psi(x) - \psi(x)\text{Tr}(B), \quad x \in \mathbb{R}^N,
\]

(4.11)

for any smooth function \( \psi \). Hence, letting \( n \) go to +\( \infty \) in (4.10), we obtain that \( Af = Af \). We now prove that \( D(A) \subseteq D_0 \). For this purpose, let \( u \in D(A) \) and fix \( \lambda > 0 \). Since \( u \in D(A) \), there exists \( f \in C_b(\mathbb{R}^N) \) such that \( u = R(\lambda, A)f \). By convolution, we can determine a sequence of smooth functions \( \{f_n\}_{n \in \mathbb{N}} \subseteq C^2_b(\mathbb{R}^N) \), bounded in \( C_b(\mathbb{R}^N) \) and converging locally uniformly to \( f \) as \( n \) tends to +\( \infty \). Taking Proposition 3.3 into account, we
deduce that $R(\lambda, A)f_n$ belongs to $C^2_b(\mathbb{R}^N)$ for any $n \in \mathbb{N}$. Moreover,

$$\langle AR(\lambda, A)f_n\rangle(x) = \int_0^{+\infty} e^{-\lambda t} \langle AT(t)f_n\rangle(x) dt = -f_n(x) + \lambda(R(\lambda, A)f_n)(x),$$

(4.12)

$x \in \mathbb{R}^N$, $n \in \mathbb{N}$. Since, by the above results, $R(\lambda, A)f_n$ converges to $u$, locally uniformly in $\mathbb{R}^N$, letting $n$ go to $+\infty$ in (4.12) we deduce that $AR(\lambda, A)f_n$ converges to $\lambda u - f$ locally uniformly in $\mathbb{R}^N$. Moreover, as we can easily see, the sequences $\{R(\lambda, A)f_n\}_{n \in \mathbb{N}}$ and $\{AR(\lambda, A)f_n\}_{n \in \mathbb{N}}$ are bounded in $C^b(\mathbb{R}^N)$. Therefore, $u \in D_0$. □

5. Schauder estimates

In this section we prove Theorems A and B. For this purpose we first characterize the behavior of the semigroup in the anisotropic spaces $C^{3\theta, \theta}_b(\mathbb{R}^N)$ ($\theta \in [0, 1]$). More precisely, we show that for any $0 < \alpha \leq \beta \leq 1$ and any $\omega > 0$, there exists a positive constant $C = C(\alpha, \beta, \omega)$ such that

$$\|T(t)\|_{L(C^{3\alpha, \alpha}_b(\mathbb{R}^N), C^{3\beta, \beta}_b(\mathbb{R}^N))} \leq C e^{\omega t} t^{-3(\beta-\alpha)/2}, \quad t > 0. \quad (5.1)$$

There are two main ingredients to prove (5.1):

(i) the interpolation set equality

$$(C_b(\mathbb{R}^N), C^{3,1}_b(\mathbb{R}^N))_{\theta, \infty} = C^{3\theta, \theta}_b(\mathbb{R}^N), \quad (5.2)$$

with equivalence of the corresponding norms, for any $\theta \in (0, 1)$;

(ii) the estimate

$$\|T(t)\|_{L(C^{3k, k}_b(\mathbb{R}^N), C^{3,1}_b(\mathbb{R}^N))} \leq C e^{\omega t} t^{-3(1-k)/2}, \quad t > 0, \quad k = 0, 1. \quad (5.3)$$

Once properties (i) and (ii) are established, an interpolation argument allows us to prove (5.1). We begin by proving property (i). For this purpose we recall that for any pair of Banach spaces $X$ and $Y$, with $Y$ continuously embedded in $X$, and any $\theta \in (0, 1)$, the interpolation space $(X; Y)_{\theta, \infty}$ consists of all $x \in X$ such that

$$\|x\|_{\theta, \infty} := \sup_{t \in (0,1)} t^{-\theta} K(t, x) < +\infty,$$

where

$$K(t, x) = \inf_{x = a + b, \ a \in X, \ b \in Y} (\|a\|_X + t\|b\|_Y), \quad x \in X, \ t > 0.$$

Moreover, for any Banach space $D$ such that $Y \subset D \subset X$ with continuous embeddings, and any $x \in (0, 1]$, we say that $D$ belongs to the class $J_\alpha$.
between $X$ and $Y$ (in short $D \in J_\alpha(X,Y)$) if there exists a positive constant $C$ such that
\[
\|x\|_D \leq C\|x\|_{X}^{1-\alpha}\|x\|_{Y}^{\alpha}, \quad x \in Y.
\]
Similarly, we say that $D$ belongs to the class $K_\alpha$ between $X$ and $Y$ (in short, $D \in K_\alpha(X,Y)$) if $D \subset (X,Y)_{\alpha,\infty}$ with a continuous embedding.

**Lemma 5.1.** For any $\theta \in (0,1)$, formula (5.2) holds true with equivalence of the corresponding norms.

**Proof.** The proof follows from [6, Theorem 2.2] where the author proved that
\[
(BUC(\mathbb{R}^N), C_b^{3\alpha,\alpha}(\mathbb{R}^N))_{\gamma,\infty} = C_b^{3\alpha,\gamma,\gamma}(\mathbb{R}^N),
\]
for any $\alpha > 0$ and any $\gamma \in (0,1)$. Here $BUC(\mathbb{R}^N)$ denotes the space of all the bounded and uniformly continuous functions $f : \mathbb{R}^N \to \mathbb{R}$. As immediately seen, the characterization in (5.4) still holds if we replace the space $BUC(\mathbb{R}^N)$ with $C_b(\mathbb{R}^N)$. Hence, taking $\alpha \in (1,2)$ and $\gamma = 1/\alpha$ in (5.4) yields $(C_b(\mathbb{R}^N),C_b^{3\alpha,\alpha}(\mathbb{R}^N))_{1/\alpha,\infty} = C_b^{3,1}(\mathbb{R}^N)$, implying that $C_b^{3,1}(\mathbb{R}^N) \subset C_b^{3,1}(\mathbb{R}^N) \subset K_{1/\alpha}(C_b(\mathbb{R}^N),C_b^{3\alpha,\alpha}(\mathbb{R}^N))$. Since $C_b^{3,1}(\mathbb{R}) \in J_{1/\alpha}(C_b(\mathbb{R}^N),C_b^{3\alpha,\alpha}(\mathbb{R}^N))$ and $C_b^{1}(\mathbb{R}^{N-r}) \in J_{1/\alpha}(C_b(\mathbb{R}^{N-r}),C_b^{3\alpha,\alpha}(\mathbb{R}^{N-r}))$ (see [4, Proposition 1.1.3]), we obtain that $C_b^{3,1}(\mathbb{R}^N) \subset J_{1/\alpha}(C_b(\mathbb{R}^N),C_b^{3\alpha,\alpha}(\mathbb{R}^N))$. The reiteration theorem (see e.g. [4, Theorem 1.2.15]) and (5.4) imply that
\[
(C_b(\mathbb{R}^N),C_b^{3,1}(\mathbb{R}^N))_{\theta,\infty} = (C_b(\mathbb{R}^N),C_b^{3\alpha,\alpha}(\mathbb{R}^N))_{\theta/\alpha,\infty} = C_b^{3\theta,\theta}(\mathbb{R}^N),
\]
for any $\theta \in (0,1)$, with equivalence of the corresponding norms, and (5.2) follows.

By reiteration, from Lemma 5.1, we get the following corollary.

**Corollary 5.2.** For any $\alpha, \theta \in (0,1)$ and any $\alpha \leq \beta < 1$,
\[
(C_b^{3\alpha,\alpha}(\mathbb{R}^N),C_b^{3,1}(\mathbb{R}^N))_{\theta,\infty} = C_b^{3(\alpha+(1-\alpha)\theta),\alpha+(1-\alpha)\theta}(\mathbb{R}^N),
\]
with equivalence of the corresponding norms.

As far as property (ii) is concerned, we observe that when $k = 0$ estimate (5.3) follows easily from (3.7). Hence, we just need to show (5.3) when $k = 1$. For this purpose, we adapt Bernstein’s method (see [1]) to this anisotropic spaces. We introduce some notation. For any integer $k \in \{2,3,4\}$, any $k$-tuple $(i_1, \ldots, i_k)$ with $1 \leq i_1 \leq \cdots \leq i_k \leq N$, and any smooth function $w$, we set
\[
D_{i_1,\ldots,i_k}w = \frac{\partial^k w}{\partial x_{i_1} \cdots \partial x_{i_k}}.
\]
Moreover, we set
\[D_{s,1}w = (D_1w, \ldots, D_rw)\quad\text{and}\quad D_{s,2}w = (D_{r+1}w, \ldots, D_Nw).\]

Then, we introduce, instead of the corresponding tensors, the vectors \(D_i^k w\) \((k = 2, 3, 4)\) consisting of all the derivatives \(D_i \ldots D_k w\) ordered as follows: \(D_i \ldots D_k w\) precedes \(D_{j1} \ldots D_{jk}w\) if \(i_l \leq j_l\) for any \(l = 1, \ldots, k\) and \(i_l < j_l\) for some \(l_0 \in \{1, \ldots, k\}\), or \(\{j_1, \ldots, j_k\}\) contains more indexes \(j_l \geq r + 1\) than the set \(\{i_1, \ldots, i_k\}\). Finally, we set \((D_i^k w)^T = ((D_i^k w_1)^T, \ldots,(D_i^k w_{r+1})^T)\), where the vector \(D_i^k w\) \((j = 1, \ldots, k+1)\) contains all the derivatives \(D_i \ldots D_k w\) with \(i_{k+1-j} \leq r < i_{k-j}\) (the last inequality being meaningful if \(j > 1\)).

For instance, if \(N = 4, k = 3, \) and \(r = 2,\) then
\[
\begin{align*}
D_{s,1}^3 &= (D_{111}w, D_{112}w, D_{122}w, D_{222}w), \\
D_{s,2}^3 &= (D_{113}w, D_{114}w, D_{123}w, D_{124}w, D_{223}w, D_{224}w), \\
D_{s,3}^3 &= (D_{133}w, D_{134}w, D_{144}w, D_{233}w, D_{234}w, D_{244}w), \\
D_{s,4}^3 &= (D_{333}w, D_{334}w, D_{344}w, D_{444}w).
\end{align*}
\]

Moreover, for notational convenience, we also set \(n_m^1 = m(m + 1)/2, n_m^2 = m(m + 1)(m + 2)/6\) and \(n_m^3 = m(m + 1)(m + 2)(m + 3)/24\) for any \(m \in \mathbb{N}.\)

**Theorem 5.3.** Let \(\varepsilon > 0\) and assume that hypotheses H1-H3 are satisfied. Then, there exist two positive constants \(C\) and \(\omega\) such that (5.3) with \(k = 1\) holds true.

**Proof.** Without losing in generality we can restrict ourselves to showing that for any \(\omega > 0\) there exists a positive constant \(C,\) independent of \(\varepsilon,\) such that
\[
\|T_\varepsilon(t)\|_{L(C_b^{3,1}(\mathbb{R}^N), C_b^{3,1}(\mathbb{R}^N))} \leq Ce^{\omega t}, \quad t > 0. \tag{5.6}
\]
Here \(\{T_\varepsilon(t)\}_{t \geq 0}\) is the semigroup defined in Section 3. Indeed, suppose that (5.6) holds true. Then, since \(T_\varepsilon(\cdot)f\) tends to \(T(\cdot)f\) in \(C^{1,2}(F)\) for any compact set \(F \subset (0, +\infty) \times \mathbb{R}^N\) and any \(f \in C_b(\mathbb{R}^N)\) (see Proposition 3.3), it follows that, for any \(t > 0,\) the sup norms of both \(DT(\cdot)f\) and \(D_{s,1}^2 T(\cdot)f\) are bounded by the right-hand side of (5.6). As far as the vector \(D_{s,1}^3 T(\cdot)f\) is concerned, we observe that (5.6) implies that
\[
[(D_{s,1}^3 T_\varepsilon(t)f)(\cdot, x_{r+1}, \ldots, x_N)]_{\text{Lip}(\mathbb{R}^r)} \leq Ce^{\omega t}\|f\|_{C_b^{3,1}(\mathbb{R}^N)}, \quad t > 0,
\]
and, letting \(\varepsilon\) go to 0 yields
\[
[(D_{s,1}^3 T(t)f)(\cdot, x_{r+1}, \ldots, x_N)]_{\text{Lip}(\mathbb{R}^r)} \leq Ce^{\omega t}\|f\|_{C_b^{3,1}(\mathbb{R}^N)}, \tag{5.7}
\]
for any $t > 0$, and any $(x_{r+1}, \ldots, x_N) \in \mathbb{R}^{N-r}$, and we are done. Indeed, since $T(t)f \in C_0^0(\mathbb{R}^N)$ for any $t > 0$ (see Proposition 3.3), from (5.7) we deduce that

$$\|D_{*1}^3 T(t)f\|_{C_b^0(\mathbb{R}^N)} \leq C e^{ct} \|f\|_{C_b^{n+1}(\mathbb{R}^N)}, \quad t > 0,$$

and (5.3) follows. So, let us prove (5.6). For this purpose, we introduce the function $\xi_\varepsilon : [0, +\infty) \times \mathbb{R}^N \to \mathbb{R}$ defined by

$$\xi_\varepsilon(t, x) = \frac{1}{2} \alpha^3 (u_\varepsilon(t, x))^2 + \langle F Du_\varepsilon(t, x), Du_\varepsilon(t, x) \rangle$$

$$+ \langle G(t) D^2 u_\varepsilon(t, x), D^2 u_\varepsilon(t, x) \rangle + \langle H(t) D^3 u_\varepsilon(t, x), D^3 u_\varepsilon(t, x) \rangle,$$

for any $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^N$, where $u_\varepsilon = T_\varepsilon(\cdot)f$ and the matrices $F \in L(\mathbb{R}^N)$, $G(t) \in L(\mathbb{R}^{n_1})$ and $H(t) \in L(\mathbb{R}^{n_2})$ $(t > 0)$ are defined by

$$F = \begin{pmatrix} \alpha I_r & \beta F_1 \\ \beta F_1^* & -tI_{N-r} \end{pmatrix}, \quad G(t) = \begin{pmatrix} I_{n_1^1} & 0 & 0 \\ 0 & \alpha^{-7/16} t I_{(n-r)} & \alpha^{-4/5} t^2 G_1 \\ 0 & \alpha^{-4/5} t^2 G_1^* & \alpha^{-7/8} t^3 I_{n_2^{n_1-n-r}} \end{pmatrix},$$

$$H(t) = \begin{pmatrix} \alpha^{-4/3} I_{n_2^1} & t H_1 & t^2 H_2 & t^3 H_3 \\ t H_1^* & \alpha^{-8/9} t^2 I_{(n-r)} & 0 & 0 \\ t^2 H_2^* & 0 & \alpha^{-1} t^4 I_{n_2^{n_1-n-r}} & \alpha^{-13/12} t^5 H_4 \\ t^3 H_3^* & 0 & \alpha^{-13/12} t^5 H_4^* & \alpha^{-9/8} t^6 I_{n_2^{n_1-n-r}} \end{pmatrix}.$$

(5.8)

(5.9)

Here, $F_1 \in L(\mathbb{R}^{N-r}, \mathbb{R}^r)$ is any matrix such that $B_3 F_1 + F_1^* B_3^*$ is strictly negative and $-t$ denotes its maximum (negative) eigenvalue. Such a matrix exists since we have assumed that $\text{rank}(B_3) \geq N/2$. Finally, $G_1 \in L(\mathbb{R}^{n_1-n-r}, \mathbb{R}^{(N-r)})$, $H_1 \in L(\mathbb{R}^{(N-r)}, \mathbb{R}^{n_2})$, $H_2 \in L(\mathbb{R}^{n_1-n-r}, \mathbb{R}^{n_2})$, $H_3 \in L(\mathbb{R}^{n_1-n-r}, \mathbb{R}^{n_2})$ and $H_4 \in L(\mathbb{R}^{n_2-n-r}, \mathbb{R}^{n_2})$ are suitable matrices to be determined later on, as well as the constants $\alpha$ and $\beta$.

Taking Proposition 3.2 into account, it is easy to check that the function $\xi_\varepsilon$ is a classical solution of the Cauchy problem

$$\begin{cases}
D_\varepsilon \xi_\varepsilon(t, \cdot) = A_\varepsilon \xi_\varepsilon(t, \cdot) + g_\varepsilon(t, \cdot), & t > 0, \\
\xi_\varepsilon(0, \cdot) = \frac{1}{2} \alpha^3 f^2 + \langle F Df, Df \rangle + \langle ED^2_{*,1} f, D^2_{*,1} f \rangle + \alpha^{-\frac{2}{3}} \langle K_1 D^3_{*,1} f, D^3_{*,1} f \rangle, 
\end{cases}$$

where $g_\varepsilon = \sum_{j=1}^3 g_{j,\varepsilon}$, and $g_{j,\varepsilon}$ $(j = 1, 2, 3)$ are given by

$$g_{1,\varepsilon}(t, x)$$
\[
\begin{split}
= -\alpha^3 \langle Q^\varepsilon(x) Du_\varepsilon(t, x), Du_\varepsilon(t, x) \rangle - 2\mathrm{Tr}(Q^\varepsilon(x)D^2 u_\varepsilon(t, x)FD^2 u_\varepsilon(t, x)) \\
- 2 \sum_{i,j=1}^{N} q_{ij}^\varepsilon(x)\langle G(t)D_i^2D_iu_\varepsilon(t, x), D_i^2D_ju_\varepsilon(t, x) \rangle \\
- 2 \sum_{i,j=1}^{N} q_{ij}^\varepsilon(x)\langle H(t)D_i^3D_iu_\varepsilon(t, x), D_i^3D_ju_\varepsilon(t, x) \rangle; \\
g_{2,\varepsilon}(t, x) = \langle G'(t)D^2 u_\varepsilon(t, x), D^2 u_\varepsilon(t, x) \rangle + \langle H'(t)D^3 u_\varepsilon(t, x), D^3 u_\varepsilon(t, x) \rangle \\
+ \langle (BF + FB^*)Du_\varepsilon(t, x), Du_\varepsilon(t, x) \rangle \\
+ 2\langle G(t)|D^2_x, (Bx, D)|u_\varepsilon(t, x), D^2 u_\varepsilon(t, x) \rangle \\
+ 2\langle H(t)|D^3_x, (Bx, D)|u_\varepsilon(t, x), D^3 u_\varepsilon(t, x) \rangle; \\
g_{3,\varepsilon}(t, x) = 2 \sum_{i,j=1}^{r} D_{ij}u_\varepsilon(t, x)(FDq_{ij}(x), Du_\varepsilon(t, x)) \\
+ 2 \sum_{i,j=1}^{r} \langle G(t)|D^2_{x^i}, q_{ij}(x)D_{ij}|u_\varepsilon(t, x), D^2 u_\varepsilon(t, x) \rangle \\
+ 2 \sum_{i,j=1}^{r} \langle H(t)|D^3_{x^i}, q_{ij}(x)D_{ij}|u_\varepsilon(t, x), D^3 u_\varepsilon(t, x) \rangle. 
\end{split}
\]

Here by \([A, B]\) we have denoted the commutator between the operators \(A\) and \(B\), and \(Q^\varepsilon = (q_{ij}^\varepsilon)\) is such that \(q_{ij}^\varepsilon = q_{ij}\) if \(i \wedge j \leq r\), \(q_{ij} = \varepsilon \delta_{ij}\) if \(i \wedge j > r\).

Our aim consists in showing that we can suitably determine \(\alpha, \beta > 0\) and five matrices \(G_1, H_j\) \((j = 1, \ldots, 4)\) such that \(F, G(t)\) and \(H(t)\) are strictly positive definite for any \(t > 0\) and there exists \(T_0 > 0\), independent of \(\varepsilon\), such that \(g_\varepsilon \leq 0\) in \((0, T_0) \times \mathbb{R}^N\). The maximum principle in Proposition 3.1 (see (3.2)) then will imply that

\[
\xi_\varepsilon(t) \geq \frac{1}{2} \alpha^3 f^2 + \langle F Df, Df \rangle + |D^2_{x^1}f|^2 + \alpha^{-1/3}|D^3_{x^1}f|^2 \leq C_0\|f\|_{C^3,1(\mathbb{R}^N)}^2,
\]

in \((0, T_0) \times \mathbb{R}^N\), for some positive constant \(C_0\), independent of \(\varepsilon\). Since, of course,

\[
\xi_\varepsilon(t) \geq \lambda_{\min}(F)\|DT_\varepsilon(t)f\|^2_\infty + \lambda_{\min}(G(1))\left(\|D^2_{x^1}T_\varepsilon(t)f\|^2_\infty \\
+ t\|D^2_{x^2}T_\varepsilon(t)f\|^2_\infty + t^3\|D^2_{x^3}T_\varepsilon(t)f\|^2_\infty \right) + \lambda_{\min}(H(1))\left(\|D^2_{x^1}T_\varepsilon(t)f\|^2_\infty \\
+ t^2\|D^2_{x^2}T_\varepsilon(t)f\|^2_\infty + t^4\|D^2_{x^3}T_\varepsilon(t)f\|^2_\infty \right) + \lambda_{\min}(H(1))\left(\|D^2_{x^1}T_\varepsilon(t)f\|^2_\infty \\
+ t^2\|D^2_{x^2}T_\varepsilon(t)f\|^2_\infty + t^4\|D^2_{x^3}T_\varepsilon(t)f\|^2_\infty \right),
\]

where
for any $t \in (0, T_0]$, then, assuming that the matrices $F$, $G(t)$ and $H(t)$ are strictly positive definite for any $t > 0$, estimate (5.3) (with $k = 1$) follows in $(0, T_0]$, with $\omega = 0$ and $C$ replaced by a new constant $C_0$. The semigroup rule then will allow us to extend the previous estimates to all positive time. Indeed, if $t > T_0$ we split $T(t)f = T(t - T_0/2)T(T_0/2)f$. Taking (3.5) into account, we get
\[
\|T(t)f\|_{C_b^2(\mathbb{R}^N)} \leq C_0\|T(t - T_0/2)f\|_{C_b^2(\mathbb{R}^N)}
\leq C_0(t - T_0/2)^{-3/2}e^{\omega(t-T_0/2)}\|f\|_{\infty} \leq C_1e^{\omega t}\|f\|_{C_b^2(\mathbb{R}^N)},
\]
for any $\omega > 0$ and some positive constant $C_1 = C_1(\omega)$. We begin by showing that $g_\varepsilon \leq 0$ in $(0, T_0] \times \mathbb{R}^N$ for a suitable choice of $T_0$, the parameters $\alpha, \beta$, and the matrices $G_1, H_j$ ($j = 1, \ldots, 4$). We begin by estimating the function $g_{1,\varepsilon}$ in (5.10). As is immediately seen, $g_{1,\varepsilon} \leq g_1$, where $g_1$ is obtained from $g_{1,\varepsilon}$ by replacing everywhere the matrix $Q^\varepsilon$ with the matrix $Q$. Hence, by Lemma 3.6 and assumption H1, we deduce that
\[
g_{1}(t) \leq -\alpha^2\nu|D_{s,1}u(t)|^2 - 2\nu\sum_{j=1}^r(D^2u(t)FD^2u(t))_{jj}
\]
\[
- 2\nu\sum_{j=1}^r\langle G(t)D^2_eD_ju(t)\rangle + 2\nu|D^2_eu(t)|^2 - 2\nu\langle H(t)D^2_eD_ju(t)\rangle + 2\nu|D^2_eu(t)|^2
\]
\[
- 2\nu\langle K_2D^2_eu(t)\rangle + 2\nu|D^2_eu(t)|^2 - 2\nu\langle K_3D^3_eu(t)\rangle + 2\nu|D^3_eu(t)|^2
\]
\[
g_{1}(t) \leq -\alpha^2\nu|D_{s,1}u(t)|^2 - 2\alpha\nu\langle K_1D_{s,1}^2u(t)\rangle - 2\nu|D^3_{s,1}u(t)|^2 - 2\alpha\nu\langle K_2D_{s,1}^2u(t)\rangle - 2\nu|D^3_{s,1}u(t)|^2
\]
\[
ge - 2\alpha^2\nu\langle K_3D_{s,1}^3u(t)\rangle - 2\nu|D^3_{s,1}u(t)|^2 - 2\alpha^2\nu\langle K_4D_{s,1}^4u(t)\rangle - 2\nu|D^3_{s,1}u(t)|^2
\]
\[
+ 2\alpha^2\nu\langle K_5D_{s,1}^5u(t)\rangle + 2\nu|D^3_{s,1}u(t)|^2 - 2\alpha^2\nu\langle K_6D_{s,1}^6u(t)\rangle + 2\nu|D^3_{s,1}u(t)|^2
\]
\[
+ 2\alpha^2\nu\langle K_7D_{s,1}^7u(t)\rangle + 2\nu|D^3_{s,1}u(t)|^2 - 2\nu|D^3_{s,1}u(t)|^2
\]
\[
- 2\nu|D^3_{s,1}u(t)|^2 - 2\nu|D^3_{s,1}u(t)|^2 - 2\nu|D^3_{s,1}u(t)|^2 - 2\nu|D^3_{s,1}u(t)|^2
\]
\[
(5.13)
\]
where $K_1, K_3, K_4, K_6, K_7, K_8$ are diagonal matrices whose minimum eigenvalues are 1 and with entries independent of $\alpha$; the entries of the matrices $K_2$ and $K_5$ depend (linearly), respectively, only on the entries of the matrix $F_1$ and $G_1$, whereas the entries of the matrix $K_{j+8}$ depend linearly only on the entries of the matrix $H_j$ ($j = 1, \ldots, 4$). Therefore, taking assumption H2 into account, observing that $|\langle A\xi, \eta\rangle| \leq \|A\| |\xi||\eta|$ for any matrix
To estimate the function estimates for $g$, computation shows that of $\alpha, a, b > 0$, we deduce the inequality

$$2t^\delta \alpha^\gamma ab \leq t^{\delta_1} \alpha^\gamma a^2 + \alpha^\gamma t^{\delta_2} b^2,$$

where $\delta = \delta_1 + \delta_2, \ 2\gamma = \gamma_1 + \gamma_2). \quad (5.14)$

which holds for any $\alpha, a, b > 0$, we deduce the inequality

$$g_1(t) \leq -\alpha^3 \nu |D_{*,1} u_\varepsilon(t)|^2 - \nu(2\alpha - \alpha^{3/5} K_2) |D_{*,1}^2 u_\varepsilon(t)|^2$$

$$- \nu t(2\alpha - \alpha^{3/5} K_2) |D_{*,2} u_\varepsilon(t)|^2 - 2\nu |D_{*,1}^3 u_\varepsilon(t)|^2$$

$$- \nu t^3(2\alpha - \alpha^{3/5} K_2) |D_{*,3} u_\varepsilon(t)|^2 - 2\nu |D_{*,1}^4 u_\varepsilon(t)|^2$$

$$- 2\nu^2 t^{3/5} |D_{*,4} u_\varepsilon(t)|^2 - \nu^2 t^2(2\alpha - \alpha^{3/5} K_2) |D_{*,4}^3 u_\varepsilon(t)|^2$$

$$- \nu^3 t^2(2\alpha - \alpha^{3/5} K_2) |D_{*,4}^4 u_\varepsilon(t)|^2 - 2\nu^3 t^2(2\alpha - \alpha^{3/5} K_2) |D_{*,4}^5 u_\varepsilon(t)|^2$$

$$+ 2t^2 \nu (K_{10} D_{*,3}^4 u_\varepsilon(t), D_{*,1}^5 u_\varepsilon(t)) + 2t^3 \nu (K_{11} D_{*,4}^5 u_\varepsilon(t), D_{*,1}^6 u_\varepsilon(t)).$$

(5.15)

As we are going to show, the right-hand side of (5.15) gives the leading terms (as $t$ tends to $0$) of the function $g_\varepsilon$. For this purpose, in the following estimates for $g_{2,\varepsilon}$ and $g_{3,\varepsilon}$, we write explicitly only the terms which are not of higher order with respect to the corresponding ones in (5.15) and the terms which are not comparable with terms in (5.15). For the other terms we simply write $\alpha_t(t^k)$ for suitable $k \in \mathbb{N}$, where $\alpha_t(t^k)$ denotes any function of $t$ (possibly depending also on $\alpha$ and $\beta$) such that $\lim_{t \to 0} t^{-k} \alpha_t(t^k) = 0$.

To estimate the function $g_{2,\varepsilon}$ (see (5.11)) we observe that a straightforward computation shows that

$$[D_{*,1}^2, \langle Bx, D \rangle] u_\varepsilon(t, x) = \mathcal{L} D_{*,1}^2 u_\varepsilon(t, x), \quad [D_{*,1}^3, \langle Bx, D \rangle] u_\varepsilon(t, x) = \mathcal{M} D_{*,1}^3 u_\varepsilon(t, x),$$

for any $x \in \mathbb{R}^N$, where, using the same splitting into blocks as in (5.8) and (5.9), we have

$$\mathcal{L} = \begin{pmatrix} L_1 & L_2 & 0 \\ L_3 & L_4 & L_5 \\ 0 & L_6 & L_7 \end{pmatrix}, \quad \mathcal{M} = \begin{pmatrix} M_1 & M_2 & 0 & 0 \\ M_3 & M_4 & M_5 & 0 \\ 0 & M_6 & M_7 & M_8 \\ 0 & 0 & M_9 & M_{10} \end{pmatrix},$$

$L_j \ (j = 1, \ldots, 7)$ and $M_j \ (j = 1, \ldots, 10)$ being suitable matrices whose entries linearly depend on the entries of $B$, but are independent of $\alpha, \beta, G_1, H_j \ (j = 1, \ldots, 4)$.

Therefore, using inequality (5.14) and arguing as above, for any $t > 0$ we deduce that

$$2G(t)[D_{*,1}^2, \langle Bx, D \rangle] u_\varepsilon(t), D_{*,1}^2 u_\varepsilon(t)]$$
\[
\leq 2\|L_1\|D^2_{*,1}u_\varepsilon(t)\|^2 + \|L_2\|((\alpha^{1/2}|D^2_{*,1}u_\varepsilon(t)|)^2 + \alpha^{-1/2}|D^2_{*,2}u_\varepsilon(t)|^2) \\
+ \|L_5\|((\alpha^{1/16}|D^2_{*,2}u_\varepsilon(t)|)^2 + \alpha^{-13/16}t^2|D^2_{*,3}u_\varepsilon(t)|^2) \\
+ \alpha^{-4/5}t^2\lambda_{\max} (G_1^* L_5 + L_5^* G_1)|D^2_{*,3}u_\varepsilon(t)|^2 + \alpha_1(t)|D^2_{*,1}u_\varepsilon(t)|^2 \\
+ \alpha_1(t)|D^2_{*,2}u_\varepsilon(t)|^2 + \alpha_1(t^2)|D^2_{*,3}u_\varepsilon(t)|^2, \\
\text{(5.16)}
\]

\[
2\langle H(t)|D^3_{*,3}u_\varepsilon(t)\rangle \\
\leq 2\alpha^{-4/3}\|M_1\|D^3_{*,1}u_\varepsilon(t)\|^2 + 2t\|H_1\|\|M_2\||D^3_{*,2}u_\varepsilon(t)|^2 \\
+ \alpha^{-8/9}\|M_5\|\|t|D^3_{*,2}u_\varepsilon(t)|^2 + t^3|D^3_{*,3}u_\varepsilon(t)|^2 \\
+ \|H_2\|\|M_2\|\|t|D^3_{*,2}u_\varepsilon(t)|^2 + t^3|D^3_{*,3}u_\varepsilon(t)|^2) \\
+ \|M_8\|\{\alpha^{-9/10}t^3|D^3_{*,3}u_\varepsilon(t)|^2 + \alpha^{-11/10}t^5|D^3_{*,4}u_\varepsilon(t)|^2\} \\
+ \alpha^{-13/12}t^5\lambda_{\max}(H_4^* M_8 + M_8^* H_4)|D^3_{*,4}u_\varepsilon(t)|^2 \\
+ 2\alpha^{-4/3}|M_2D^3_{*,2}u_\varepsilon(t), D^3_{*,1}u_\varepsilon(t)| + 2t\langle H_1 M_5 D^3_{*,3}u_\varepsilon(t), D^3_{*,1}u_\varepsilon(t)\rangle \\
+ 2t^2\langle H_2 M_8 D^3_{*,4}u_\varepsilon(t), D^3_{*,1}u_\varepsilon(t)\rangle + \alpha_1(t)|D^3_{*,1}u_\varepsilon(t)|^2 \\
+ \alpha_1(t)|D^3_{*,2}u_\varepsilon(t)|^2 + \alpha_1(t^3)|D^3_{*,3}u_\varepsilon(t)|^2 + \alpha_1(t^5)|D^3_{*,4}u_\varepsilon(t)|^2
\]

and

\[
\langle (BF + FB^*)Du_\varepsilon(t), Du_\varepsilon(t)\rangle + \langle G'(t)D^2_{*,3}u_\varepsilon(t), D^2_{*,2}u_\varepsilon(t)\rangle \\
+ \langle H'(t)D^3_{*,3}u_\varepsilon(t), D^3_{*,1}u_\varepsilon(t)\rangle \leq 2(\alpha\|B_1\| + \beta\|B_2\|)|D^3_{*,1}u_\varepsilon(t)|^2 \\
+ (\beta\|B_1\|\|\|F_1\|\|\|B_2\|\|)|D^3_{*,1}u_\varepsilon(t)|^2 \\
+ \langle t\|D_1\| - \beta\|D_{*,2}u_\varepsilon(t)\|^2 + \alpha^{-7/16}|D^2_{*,2}u_\varepsilon(t)|^2 \\
+ \|G_1\|((\alpha^{-3/5}|D^2_{*,2}u_\varepsilon(t)|^2 + \alpha^{-1/2}|D^2_{*,3}u_\varepsilon(t)|^2) + 3t^2\alpha^{-7/8}|D^2_{*,3}u_\varepsilon(t)|^2 \\
+ 2\langle H_1 D^3_{*,2}u_\varepsilon(t), D^3_{*,1}u_\varepsilon(t)\rangle + 4t\langle H_2 D^3_{*,3}u_\varepsilon(t), D^3_{*,1}u_\varepsilon(t)\rangle \\
+ 6t^2\langle H_3 D^3_{*,4}u_\varepsilon(t), D^3_{*,1}u_\varepsilon(t)\rangle + 2t\alpha^{-8/9}|D^3_{*,2}u_\varepsilon(t)|^2 + 4\alpha^{-1}t^2|D^3_{*,3}u_\varepsilon(t)|^2 \\
+ 5\langle H_4\|((\alpha^{-25/24}t^3|D^3_{*,3}u_\varepsilon(t)|^2 + \alpha^{-9/8}t^5|D^3_{*,4}u_\varepsilon(t)|^2) + 6t^5\alpha^{-9/8}|D^3_{*,4}u_\varepsilon(t)|^2
\]

Therefore, combining (5.16)-(5.18), we get

\[
g_{2,\varepsilon}(t) \leq a_1|D_{*,1}u_\varepsilon(t)|^2 + a_2|D_{*,2}u_\varepsilon(t)|^2 + a_3(t)|D^2_{*,1}u_\varepsilon(t)|^2 + a_4(t)|D^2_{*,2}u_\varepsilon(t)|^2 \\
+ a_5(t)|D^2_{*,3}u_\varepsilon(t)|^2 + a_6(t)|D^3_{*,1}u_\varepsilon(t)|^2 + a_7(t)|D^3_{*,2}u_\varepsilon(t)|^2 + a_8(t)|D^3_{*,3}u_\varepsilon(t)|^2 \\
+ a_9(t)|D^3_{*,4}u_\varepsilon(t)|^2 + ((2\alpha^{-4/3}M_2 + 2H_1)D^2_{*,2}u_\varepsilon(t), D^3_{*,1}u_\varepsilon(t)) \\
+ 2t(H_1 M_5 + 2H_2)D^3_{*,3}u_\varepsilon(t), D^3_{*,1}u_\varepsilon(t))
\]
for any $t > 0$, where
\begin{align*}
a_1 &= 2\alpha \|B_1\| + \alpha^{1/2}(\beta \|B_1\|\|F_1\| + \varepsilon \|B_2\|) + 2\beta \|B_2\|\|F_1\|, \\
a_2 &= 2\varepsilon \|B_4\| - \beta t + \alpha^{-1/2}(\beta \|B_1\|\|F_1\| + \varepsilon \|B_2\|), \\
a_3(t) &= 2\|L_1\| + \alpha^{1/2}\|L_2\| + o_t(1), \\
a_4(t) &= \alpha^{-1/16}\|L_5\| + \alpha^{-7/16} + \alpha^{-1/2}\|L_2\| + a^{-3/5}\|G_1\| + o_t(1), \\
a_5(t) &= t^2\left(\frac{\alpha^{-4/5}\lambda_{\text{max}}(G_1^*L_5 + L_5^*G_1)}{\|G_1\| + o_t(1)} + a^{-1}\right), \\
a_6(t) &= 2\alpha^{-4/3}\|M_4\| + o_t(1), \\
a_7(t) &= t\{(2\|H_1\| + \|H_2\|)\|M_2\| + \alpha^{-8/9}\|M_5\| + 2\alpha^{-8/9} + o_t(1)\}, \\
a_8(t) &= t^3\left\{(\frac{\alpha^{-8/9}\|M_5\| + \|H_2\|\|M_2\| + \|H_2\|\|M_5\|}{\|H_4\| + o_t(1)} + 5\alpha^{-25/24}\|H_4\| + o_t(1)\}, \\
a_9(t) &= t^5\left\{(\frac{\alpha^{-13/12}\lambda_{\text{max}}(H_4^*M_8 + M_8^*H_4)}{\|H_4\| + 6\alpha^{-9/8} + o_t(1)}} + 5\alpha^{-9/8}\|H_4\| + 6\alpha^{-9/8} + o_t(1)\right\}.
\end{align*}

We now choose
\[H_1 = -\alpha^{-4/3}M_2, \quad H_2 = \frac{1}{2}\alpha^{-4/3}M_2M_5, \quad H_3 = -\frac{1}{6}\alpha^{-4/3}M_2M_5M_8,\]
to let the last two rows in (5.19) disappear. We stress that the previous choice of the matrices $H_j$, $j = 1, 2, 3$ implies $K_j = \alpha^{-4/3}\tilde{K}_j$, $(j = 9, 10, 11)$, for some matrices $\tilde{K}_j$ independent of $\alpha$. As a consequence, we can estimate the last three terms in (5.15) as follows:
\begin{align*}
2t\nu\langle K_9D^4_{\ast,2}u_\varepsilon(t), D^4_{\ast,1}u_\varepsilon(t) \rangle + 2t^2\nu\langle K_{10}D^4_{\ast,3}u_\varepsilon(t), D^4_{\ast,1}u_\varepsilon(t) \rangle \\
&\quad + 2t^3\nu\langle K_{11}D^4_{\ast,4}u_\varepsilon(t), D^4_{\ast,1}u_\varepsilon(t) \rangle
\leq \nu\alpha^{-3/2}(\|K_9\| + \|\tilde{K}_{10}\| + \|\tilde{K}_{11}\|)\|D^4_{\ast,1}u_\varepsilon(t)\|^2 + \nu\alpha^{-7/6}\|D^4_{\ast,2}u_\varepsilon(t)\|^2 + \nu\alpha^{-7/6}\|D^4_{\ast,3}u_\varepsilon(t)\|^2
\leq 2t^4\nu\|D^4_{\ast,1}u_\varepsilon(t)\|^2 + \nu\alpha^{-7/6}\|D^4_{\ast,2}u_\varepsilon(t)\|^2 + \nu\alpha^{-7/6}\|D^4_{\ast,3}u_\varepsilon(t)\|^2.
\end{align*}

Let us now consider the function $g_{3,\varepsilon}$ in (5.12). A long, but straightforward computation, yields, for any $x \in \mathbb{R}^N$, the two fundamental representations
\begin{align*}
[D^2_{\ast}, q_{ij}(x)D_{ij}]u_\varepsilon(t, x) &= D_{ij}u_\varepsilon(t, x)D^2_{\ast}q_{ij}(x) + ND^2_{\ast}u_\varepsilon(t, x), \\
[D^2_{\ast}, q_{ij}(x)D_{ij}]u_\varepsilon(t, x) &= D_{ij}u_\varepsilon(t, x)D^2_{\ast}q_{ij}(x) + P(x)D^2_{\ast}u_\varepsilon(t, x)
\end{align*}
where, for any $x \in \mathbb{R}^N$, the matrices $\mathcal{N}(x) \in L(\mathbb{R}^n_k, \mathbb{R}^n_N)$, $\mathcal{P}(x) \in L(\mathbb{R}^n_k)$, and $\mathcal{R}(x) \in L(\mathbb{R}^n_k, \mathbb{R}^n_N)$, split according to the splitting of the vectors $D^k_x u_\epsilon (k = 2, 3, 4)$, are given by

$$\mathcal{N}(x) = \begin{pmatrix} * & 0 & 0 & 0 \\ * & * & 0 & 0 \\ 0 & * & 0 & 0 \end{pmatrix}, \quad \mathcal{P}(x) = \begin{pmatrix} * & 0 & 0 & 0 \\ * & * & 0 & 0 \\ * & * & 0 & 0 \\ 0 & * & 0 & 0 \end{pmatrix},$$

$$\mathcal{R}(x) = \begin{pmatrix} * & 0 & 0 & 0 \\ * & * & 0 & 0 \\ 0 & * & 0 & 0 \\ 0 & 0 & * & 0 \end{pmatrix},$$

where "*" denote suitable matrices whose entries, depending on $x$, are linear combinations of the entries of the derivatives of the diffusion coefficients $q_{ij}$. In particular, their moduli can be estimated at any $x \in \mathbb{R}^N$ by $C \sqrt{v(x)}$, $C$ being a positive constant independent of $\alpha$, $\beta$, $H_j$ ($j = 1, 2, 3, 4$). Hence, taking assumption H2 into account, and applying (5.14) with $\alpha = \sqrt{v} |D^1_{x,\epsilon} u_\epsilon|$ and $b = |D^k_{x,\epsilon} u_\epsilon|$ ($i \leq k$, $j = 1, \ldots, k + 1$), it is easy to check that all the terms in the definition of the function $g_{3,\epsilon}$ are negligible (as $t$ tends to 0), with respect to the terms in (5.16) and (5.17). Now combining (5.13), (5.19), and (5.20), and taking $\beta = 2 \|B_4\| + 1$, we get

$$g_\epsilon(t) \leq \left\{ -\alpha^3 + o_\alpha(\alpha^3) \right\} |D_{x,1} u_\epsilon(t)|^2 + \left\{ -\nu + o_\alpha(1) \right\} |D_{x,2} u_\epsilon(t)|^2$$

$$+ \nu \left\{ -2\alpha + o_\alpha(\alpha) + o_\alpha(1) \right\} |D^2_{x,1} u_\epsilon(t)|^2 + \nu \left\{ -2\alpha + o_\alpha(\alpha) + o_\alpha(1) \right\} |D^2_{x,2} u_\epsilon(t)|^2$$

$$+ t^2 \left\{ \lambda_{\text{max}}(G^1_1 L_5 + L_5^1 G^1_1) \alpha^{-4/5} + o_\alpha(\alpha^{-4/5}) + o_\alpha(1) \right\} |D^3_{x,3} u_\epsilon(t)|^2$$

$$+ \{ -2\nu + o_\alpha(1) + o_\alpha(1) \} |D^2_{x,1} u_\epsilon(t)|^2 + t \left\{ -2\alpha^{-7/16} + o_\alpha(\alpha^{-7/16}) \right\} |D^3_{x,3} u_\epsilon(t)|^2$$

$$+ t \{ -2\nu + o_\alpha(1) + o_\alpha(1) \} |D^2_{x,2} u_\epsilon(t)|^2 + t^2 \{ -2\nu \alpha^{-7/8} + o_\alpha(\alpha^{-7/8}) + o_\alpha(1) \} |D^3_{x,3} u_\epsilon(t)|^2$$

$$+ t^5 \{ \alpha^{-13/12} \lambda_{\text{max}}(H^4_1 M_8 + M_8^1 H_4) + o_\alpha(\alpha^{-13/12}) + o_\alpha(1) \} |D^4_{x,4} u_\epsilon(t)|^2$$

$$- \nu \left\{ -2\alpha^{-4/3} + o_\alpha(\alpha^{-4/3}) + o_\alpha(1) \right\} |D^4_{x,1} u_\epsilon(t)|^2 - \nu t^2 \left\{ -2\alpha^{-8/9} + o_\alpha(\alpha^{-8/9}) \right\} |D^4_{x,1} u_\epsilon(t)|^2$$

$$+ \{ -2\alpha^{-1} + o_\alpha(\alpha^{-1}) + o_\alpha(1) \} |D^4_{x,4} u_\epsilon(t)|^2$$

$$- \nu t^6 \left\{ -2\alpha^{-9/8} + o_\alpha(\alpha^{-9/8}) \right\} |D^4_{x,4} u_\epsilon(t)|^2,$$  

(5.21)

where, to simplify the notation, we denote by $o_\alpha(\alpha^k)$ ($k \in \mathbb{R}_+$) any function depending only on $\alpha$ such that $\lim_{\alpha \to +\infty} \alpha^{-k} o_\alpha(\alpha^k) = 0$. We now show
that we can choose $T_0$ sufficiently close to 0 and $\alpha$ large enough, so as to make all the terms in (5.21) negative in $(0, T_0] \times \mathbb{R}^N$. We begin by observing that if we choose $G_1 = -L_5$ and $H_4 = -M_8$ the matrices $G_1^*L_5 + L_5^*G_1$ and $H_4^*M_8 + M_8^*H_4$ are strictly negative definite. Indeed, up to rearranging rows, we have

$$L_5 = S_0, \quad M_8 = \begin{pmatrix} S_0 & 0 & \cdots & 0 \\ \star & S_1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \star & \cdots & \cdots & \star & S_{N-r-1} \end{pmatrix},$$

where $S_j \in L(\mathbb{R}^{n_1^r, N-r-j, \cdots, \mathbb{R}^r (N-r-j)})$ is given by

$$S_j = \begin{pmatrix} B_3^j & 0 & \cdots & 0 \\ \star & B_3^{j+1} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \star & \cdots & \cdots & \star & B_3^{N-r-1} \end{pmatrix}, \quad j = 0, \ldots, N - r - 1,$$

$B_3^j$ being obtained from $B_3^3$ by removing the first $j$ columns. Here and above by “$\star$” we denote suitable matrices whose entries depend linearly only on the entries of $B_3$. Since, due to assumption $H3$, rank($B_3$) = $N - r$, it can be seen, because of the particular structure of the matrices $L_5$ and $M_8$, that rank($L_5$) = $n_1^{N-r}$ and rank($M_8$) = $n_2^{N-r}$. Therefore, it is immediate to check that the matrices $L_5^*L_5$ and $M_8^*M_8$ are strictly positive definite. Hence, from (5.21) it is clear that we can fix $\alpha > 0$ and $T_0 > 0$ such that all the terms in its right-hand side are negative. Finally, it is easy to see that, due to our choices of $\beta$, $G_1$ and $H_j$ ($j = 1, \ldots, 4$), for $\alpha$ large enough the matrices $F$, $G(t)$ and $H(t)$ are strictly positive definite for any $t > 0$. \[\square\]

We are now able to prove (5.1).

**Proposition 5.4.** For any $0 \leq \alpha \leq \beta \leq 3$ and any $\omega > 0$ there exists a positive constant $C = C(\alpha, \beta, \omega)$, such that (5.1) holds true.

**Proof.** As already mentioned, the proof is based on an interpolation argument. We recall that if $X_j$ and $Y_j$ ($j = 1, 2$) are two pairs of Banach spaces with $Y_j$ continuously embedded into $X_j$ for $j = 1, 2$, and
Moreover, we show that if 

\[ f \text{ fact implies that the function} \]

\[ \text{Let Proposition 5.5.} \]

\[ \text{given in} [5], \text{we omit it.} \]

\[ \text{with the regularity claimed in Theorem A. Since the proof is similar to that} \]

\[ \text{providing some regularity properties of the functions belonging to} \]

\[ \text{in [5, Theorem 2.1]} \]

\[ \text{we can prove the following proposition which, besides} \]

\[ \text{r} \]

\[ \text{solutions which are twice continuously differentiable with respect to the first} \]

\[ \text{(5.25) is the unique solution of equation (5.26) among all the distributional} \]

\[ \text{θ} \]

\[ \text{□} \]

\[ \text{(5.5) into account, we deduce (5.1) in its generality.} \]

\[ \text{Finally, interpolating (5.23) and (5.24), with} \]

\[ \text{and, still taking (5.2) into account, we deduce that} \]

\[ \text{Finally, interpolating (5.23) and (5.24), with} \]

\[ \text{taking} \]

\[ \text{into account, we deduce (5.1) in its generality.} \]

\[ \text{In order to prove Theorem A, we show that for any} f \in C_b(\mathbb{R}^N) \text{ and any} \]

\[ \lambda > 0 \text{ the function} \]

\[ u(x) = \int_0^{+\infty} e^{-\lambda t}(T(t)f)(x)dt, \quad x \in \mathbb{R}^N, \quad (5.25) \]

\[ \text{is a distributional solution to the elliptic equation} \]

\[ \lambda u - Au = f. \quad (5.26) \]

\[ \text{Moreover, we show that if} f \in C_b^{\theta,9/3}(\mathbb{R}^N) \text{ for some} \theta \in (0,1), \text{then} u \in C_b^{2+\theta,9/3}(\mathbb{R}^N) \text{ and satisfies} \]

\[ \text{Arguing as in [5, Theorem 2.1] we can prove the following proposition which, besides} \]

\[ \text{providing some regularity properties of the functions belonging to} \]

\[ D(A), \text{in} \]

\[ \text{fact implies that the function} u \text{ in} (5.25) \text{ is the unique solution of equation (5.26) among all the distributional} \]

\[ \text{solutions which are twice continuously differentiable with respect to the first} \]

\[ r \]

\[ \text{variables, and are bounded with their classical derivatives.} \]

\[ \text{Proposition 5.5. Let} u \in D(A) \text{ be such that} Au \in C_b^{\theta,9/3}(\mathbb{R}^N) \text{ for some} \theta \in [0,1). \text{ Then,} u \in C_b^{2+\theta,9/3}(\mathbb{R}^N) \text{ and there exists a positive constant} \]

\[ \text{C, independent of} u, \text{ such that} \]

\[ \|u\|_{C_b^{2+\theta,9/3}(\mathbb{R}^N)} \leq C(\|u\|_\infty + \|Au\|_{C_b^{\theta,9/3}(\mathbb{R}^N)}). \]
In particular, $D(A)$ is continuously embedded into $C^2_{\theta}(\mathbb{R}^N)$. Finally, if $f \in C^2_{\theta}(\mathbb{R}^N)$ and $u$ solves the equation $\lambda u - Au = f$, then

$$\|u\|_{C^2_{\theta}(\mathbb{R}^N)} \leq C\|f\|_{C^2_{\theta}(\mathbb{R}^N)}.$$ 

In order to prove that $u$ is actually the unique distributional solution to problem (5.26) which is twice continuously differentiable with respect to the first $r$ variables, we begin by proving the following lemma.

**Lemma 5.6.** Let $\mathcal{B}$ be the first-order differential operator formally defined by

$$\mathcal{B}u(x) = (Bx, Du(x)), \quad x \in \mathbb{R}^N.$$ 

Then, for any $u \in C_b(\mathbb{R}^N)$ such that $\mathcal{B}u \in C_b(\mathbb{R}^N)$, there exists a sequence of smooth functions $u_n$ such that $\mathcal{B}u_n \in C_b(\mathbb{R}^N)$ for any $n \in \mathbb{N}$ and converges to $\mathcal{B}u$ locally uniformly.

**Proof.** Let $\varphi$ be a smooth function compactly supported in $B(0,1)$, such that $0 \leq \varphi(x) \leq 1$ for any $x \in \mathbb{R}^N$ and with $\|\varphi\|_{L^1(\mathbb{R}^N)} = 1$. For any $n \in \mathbb{N}$ and any $x \in \mathbb{R}^N$, let $\varphi_n(x) = n^N \varphi(nx)$ and define the function $u_n : \mathbb{R}^N \to \mathbb{R}$ by setting

$$u_n(x) = (u \ast \varphi_n)(x) := \int_{\mathbb{R}^N} u(x-y)\varphi_n(y)dy, \quad x \in \mathbb{R}^N. \quad (5.27)$$

As is easily seen, $u_n$ converges to $u$ as $n$ tends to $+\infty$, locally uniformly in $\mathbb{R}^N$. Let us show that $\mathcal{B}u_n$ converges to $\mathcal{B}u$ locally uniformly in $\mathbb{R}^N$, as well. To compute $\mathcal{B}u_n$ we first assume that $u \in C^1_b(\mathbb{R}^N)$, and observe that for any $i, j = 1 \ldots N$, we have

$$x_jD_iu_n(x) = x_j\int_{\mathbb{R}^N} (D_i u)(x-y)\varphi_n(y)dy$$

$$= \int_{\mathbb{R}^N} (x_j - y_j)(D_i u)(x-y)\varphi_n(y)dy + \int_{\mathbb{R}^N} (D_i u)(x-y)y_j\varphi_n(y)dy$$

$$= \int_{\mathbb{R}^N} (x_j - y_j)(D_i u)(x-y)\varphi_n(y)dy + \int_{\mathbb{R}^N} u(x-y)D_i y_j\varphi_n(y)dy$$

$$= \int_{\mathbb{R}^N} (x_j - y_j)(D_i u)(x-y)\varphi_n(y)dy + \delta_{ij} \int_{\mathbb{R}^N} u(x-y)\varphi_n(y)dy$$

$$+ \int_{\mathbb{R}^N} u(x-y)y_j D_i\varphi_n(y)dy.$$ 

Therefore,

$$\mathcal{B}u_n = \mathcal{B}u \ast \varphi_n + \text{Tr}(B)u_n + u \ast \mathcal{B}\varphi_n. \quad (5.28)$$
Let now \( u \in C_b(\mathbb{R}^N) \) and let \( \{u_m\}_{m \in \mathbb{N}} \subset C^1_b(\mathbb{R}^N) \) be a sequence of smooth functions converging to \( u \) locally uniformly. Moreover, let \( u_m^* = u_m * \varphi_n \). As we can easily see, \( u_m^* \) and \( Bu_m^* \) converge, respectively, to \( u_n \) and \( Bu_n \), locally uniformly in \( \mathbb{R}^N \), as \( m \) tends to \( +\infty \). Moreover, \( Bu_m^* \) converges to \( Bu \) in the sense of distributions. This implies that \( Bu_m^* \) converges to \( Bu \) pointwise in \( \mathbb{R}^N \) as \( m \) tends to \( +\infty \). Therefore, writing (5.28) with \( u_n \) replaced by \( u_m^* \) and letting \( m \) go to \( +\infty \), we see that (5.28) holds true also in the case where \( u \in C_b(\mathbb{R}^N) \). Now we are almost done. Indeed, since \( Bu \in C^1_b(\mathbb{R}^N) \), then \( Bu \) converges locally uniformly to \( Bu \). Finally, \( (u \ast B \varphi_n)(x) = \sum_{i,j=1}^{N} b_{ij} \int_{\mathbb{R}^N} u(x - y) y_j D_i \varphi_n(y) dy \) and the last side of the previous chain of equalities tends to \( u(x) \sum_{i,j=1}^{N} b_{ij} \int_{\mathbb{R}^N} z_j D_i \varphi(z) dz = -u(x) \text{Tr}(B) \), as \( n \) tends to \( +\infty \), locally uniformly in \( \mathbb{R}^N \). Therefore, \( Bu_n \) converges locally uniformly to \( Bu \) as \( n \) tends to \( +\infty \). \( \square \)

We are now able to prove Theorem A.

**Proof of Theorem A. (Existence).** Let us show that, for any \( \lambda > 0 \), the function \( u = R(\lambda, A)f \) is a distributional solution to the equation (5.26). For this purpose, let \( \{f_n\}_{n \in \mathbb{N}} \subset C^2_b(\mathbb{R}^N) \) be a sequence of smooth functions compactly supported in \( \mathbb{R}^N \), converging locally uniformly to \( f \) as \( n \) tends to \( +\infty \) and set \( u_n = R(\lambda, A)f_n \). Using (3.6) with \( \omega = \lambda/2 \), we can prove that \( u_n \in C^2_b(\mathbb{R}^N) \). Moreover, by Proposition 4.5, \( Au_n = Au_n \) for any \( n \in \mathbb{N} \). It follows that \( u_n \) is a classical solution to the equation \( \lambda u_n - Au_n = f_n \). Let now \( A^* \) denote the formal adjoint of the operator \( A \) (see (4.11)). Then, for any \( n \in \mathbb{N} \) and any smooth function \( \varphi \in C^\infty_b(\mathbb{R}^N) \) with compact support, it holds that

\[
\int_{\mathbb{R}^N} f_n \varphi \, dx = \int_{\mathbb{R}^N} (\lambda u_n - Au_n) \varphi \, dx = \int_{\mathbb{R}^N} u_n (\lambda \varphi - A^* \varphi) \, dx. \tag{5.29}
\]

As we have already shown in the proof of Proposition 4.5, \( u_n \) converges to \( u \) as \( n \) tends to \( +\infty \), locally uniformly in \( \mathbb{R}^N \). Therefore, taking the limit as...
n tends to $+\infty$ in (5.29) gives
\[
\int_{\mathbb{R}^N} f \varphi \, dx = \int_{\mathbb{R}^N} u (\lambda \varphi - A^* \varphi) \, dx,
\]
implying that $u$ is a distributional solution to the equation (5.26). Now, Proposition 5.5 implies that $u$ enjoys all the claimed regularity properties.

**Uniqueness.** We now prove that $v = 0$ is the unique distributional solution to the equation $\lambda v - Av = 0$ which is bounded and continuous in $\mathbb{R}^N$ and admits classical derivatives $D_{ij} v \in C_b(\mathbb{R}^N)$ for any $i,j = 1, \ldots, r$. This is equivalent to proving that any solution $v$ to the previous equation, with the claimed regularity, belongs to $D(A)$. For this purpose, let $\{\hat{v}_m\}_{m \in \mathbb{N}}$ be the sequence defined by $\hat{v}_m = T(1/m) v$ for any $m \in \mathbb{N}$. By Proposition 3.3, each $\hat{v}_m$ belongs to $C_b^2(\mathbb{R}^N)$ and $\hat{v}_m$ converges locally uniformly to $v$ as $m$ tends to $+\infty$. Moreover, $\|\hat{v}_m\|_\infty \leq \|v\|_\infty$ for any $m \in \mathbb{N}$. We now show that
\[
A\hat{v}_m = T(1/m) Av,
\]
for any $m \in \mathbb{N}$. This will imply that $A\hat{v}_m$ converges locally uniformly to $Av$ and $\|A\hat{v}_m\|_\infty \leq \|Av\|_\infty$ for any $m \in \mathbb{N}$. Hence, by Proposition 4.5, $v \in D(A)$ and, consequently, $v = 0$. So, let us prove (5.30). For this purpose, we approximate $v$ by convolution, by the sequence of smooth functions $\{v_n\}_{n \in \mathbb{N}} \subset C_b^2(\mathbb{R}^N)$ defined as in (5.27). As we can easily see, $v_n$ converges locally uniformly to $v$ as $n$ tends to $+\infty$ and $D_i v_n, D_{ij} v_n$ converge locally uniformly, respectively, to $D_i v$ and $D_{ij} v$ for any $1 \leq i,j \leq r$. Moreover, since, by Lemma 5.6, $B v_n$ converges to $B v$ locally uniformly, then $A v_n$ converges to $A v$ locally uniformly as well. By Lemma 4.4, $T(1/m) Av_n = A T(1/m) v_n$ for any $m, n \in \mathbb{N}$. Now, by Proposition 4.1, $T(1/m) v_n$ converges to $T(1/m) v$ in $C^2(K)$ for any compact set $K \subset \mathbb{R}^N$ and $T(1/m) A v_n$ converges to $T(1/m) A v$, locally uniformly in $\mathbb{R}^N$. Hence, (5.30) follows. This finishes the proof.

The results of Theorem A can be improved if $f$ is smoother. In such a case it can be proved that the function $u = R(\lambda, A) f$ is a classical solution to equation (1.1), in the sense that all the derivatives involved in the definition of $A u$ exist in the classical sense and are bounded and continuous in $\mathbb{R}^N$. Since the proof of the next corollary is similar to, and even simpler than, the proof of Corollary 5.9, we do not go into details.

**Corollary 5.7.** Suppose that $f \in C_b^\theta(\mathbb{R}^N)$ for some $\theta > 1/3$. Then, the function $u = R(\lambda, A) f$ belongs to $C_b^{\theta+2/3}(\mathbb{R}^N)$ and there exists a positive constant $C$, independent of $f$, such that
\[
\|u\|_{C_b^{\theta+2/3}(\mathbb{R}^N)} \leq C \|f\|_{C_b^\theta(\mathbb{R}^N)}.
\]
We now consider the parabolic nonhomogeneous Cauchy problem (1.2). We are going to show that the unique distributional solution to such a Cauchy problem, with the regularity properties claimed in Theorem A, is the function $u$, the so-called mild solution, given by the variation of constants formula

$$u(t, x) = (T(t) f)(x) + \int_0^t (T(t-s) g(s, \cdot))(x) \, ds, \quad (t, x) \in [0, T] \times \mathbb{R}^N,$$

(5.31)

As a first step we observe that the convolution term in (5.31) is well defined. Of course, it suffices to check that for any $x \in \mathbb{R}^N$ the function $w : [0, T] \times [0, T] \to \mathbb{R}$, defined by $w(r, s) = (T(r) f(s))(x)$ for any $r, s \in [0, T]$, is continuous. For this purpose we observe that, by Proposition 3.3, the function $w(\cdot, s)$ is continuous for any $s \in [0, T]$, and by Proposition 4.1, $w(r, \cdot)$ is continuous in $[0, T]$, uniformly with respect to $r \in [0, T]$. Now taking (5.1) into account, and adapting to our case the abstract result in [5, Theorem 2.2], we can prove the following proposition, which provides the main smoothing properties of the function $v$.

**Proposition 5.8.** Let $g : [0, T] \times \mathbb{R}^N \to \mathbb{R}$ be a continuous function such that $\sup_{t \in [0, T]} \|g(t, \cdot)\|_{C_b^{2,\theta/3}(\mathbb{R}^N)} < +\infty$ for some $\theta \in [0, 1)$. Then, $v(t, \cdot) \in C^{2+\theta, (2+\theta)/3}(\mathbb{R}^N)$ for any $t \in [0, T]$. Moreover, there exists a positive constant $C$, independent of $g$, such that

$$\sup_{0 \leq t \leq T} \|v(t, \cdot)\|_{C_b^{2+\theta, (2+\theta)/3}(\mathbb{R}^N)} \leq C \sup_{0 \leq t \leq T} \|g(t, \cdot)\|_{C_b^{\theta/3}(\mathbb{R}^N)}.$$

Finally, the functions $u$, $D_i u$ and $D_{ij} u$ ($1 \leq i, j \leq r$) are continuous in $[0, T] \times \mathbb{R}^N$.

We are now in a position to prove Theorem B.

**Proof of Theorem B. (Existence).** Of course, by Proposition 3.3 we can restrict ourselves to proving the assertion in the case where $f \equiv 0$. For this purpose, we begin by proving that the function

$$v(t, x) = \int_0^t (T(t-s) g(s, \cdot))(x) \, ds, \quad (t, x) \in [0, T] \times \mathbb{R}^N,$$

is a distributional solution to (1.2) (with $f \equiv 0$). For this purpose, we introduce a sequence of smooth functions $g_n \in C_b^{1,2}([0, T] \times \mathbb{R}^N)$ ($n \in \mathbb{N}$) converging locally uniformly in $[0, T] \times \mathbb{R}^N$ to $g$ and we denote by $v_n$ the function defined by (5.31) with $g$ being replaced by $g_n$. Using estimate (3.6) with $k = 2$ we can show that $v_n$ is a classical solution to problem (1.2) (with $f \equiv 0$) where we replace $g$ by $g_n$ (see also the proof of Corollary 5.9).
Moreover, $v_n$ converges to $v$ uniformly in $[0, T] \times K$ for any compact set $K \subset \mathbb{R}^N$. Indeed, recalling that $\{T(t)\}_{t \geq 0}$ is an order preserving semigroup (see Remark 3.4), we deduce that
\[
\sup_{(t, x) \in [0, T] \times K} |v_n(t, x) - v(t, x)| = \sup_{(t, x) \in [0, T] \times K} \left| \int_0^t (T(s)(g_n(t - s, \cdot) - g(t - s, \cdot))) (x) ds \right| \\
\leq \sup_{x \in K} \int_0^T \left( \sup_{r \in [0, T]} |g_n(r, \cdot) - g(r, \cdot)| \right) (x) ds.
\]
Since $\sup_{t \in [0, T]} |g_n(r, \cdot) - g(r, \cdot)|$ converges to 0 locally uniformly in $\mathbb{R}^N$ as $n$ tends to $+\infty$, Proposition 4.1 implies that the last side of the previous chain of inequalities vanishes as $n$ tends to $+\infty$. Hence, $v_n$ tends to $v$ uniformly in $[0, T] \times K$. Now, we observe that for any smooth function $\varphi \in C_b^\infty((0, T) \times \mathbb{R}^N)$ with compact support in $(0, T) \times \mathbb{R}^N$, it holds that
\[
\int_{(0, T) \times \mathbb{R}^N} g_n \varphi \, dt \, dx = \int_{(0, T) \times \mathbb{R}^N} (D_t v_n - A v_n) \varphi \, dt \, dx \\
= \int_{(0, T) \times \mathbb{R}^N} v_n (-D_t \varphi - A^* \varphi) \, dt \, dx,
\]
where $A^*$ is given by (4.11). Letting $n$ go to $+\infty$ we deduce that $v$ is a distributional solution of (1.2) with $f \equiv 0$. The regularity properties of $u$ and (1.4) follow from Propositions 5.4 and 5.8.

**Uniqueness.** Let us now prove that the function $u$ in (5.31) is the unique distributional solution to problem (1.2) which is twice continuously differentiable in $[0, T] \times \mathbb{R}^N$ with respect to the space variables $x_j$ ($j = 1, \ldots, r$). We adapt to this situation the proof of Theorem A. For this purpose, let $w$ be a solution to the Cauchy problem (1.2) with $f = 0$ and $g = 0$, such that $w, D_i w, D_{ij} w \in C_b([0, T] \times \mathbb{R}^N)$ for any $i, j = 1, \ldots, r$. We extend $w$ by continuity to $\mathbb{R} \times \mathbb{R}^N$ by setting $\overline{w}(t, \cdot) = 0$ for any $t < 0$ and $\overline{w}(t, \cdot) = w(T, \cdot)$ for any $t > T$. The so extended function is as smooth as $w$ is; i.e., $D_i \overline{w}$ and $D_{ij} \overline{w}$ are continuous in $\mathbb{R} \times \mathbb{R}^N$ for any $i, j \leq r$. Then, we regularize $\overline{w}$ by convolution by setting
\[
w_n(t, x) = (\varphi_n \varphi_n)(t, x) := \int_{\mathbb{R} \times \mathbb{R}^N} \varphi(t-s, x-y) \varphi_n(y) \, ds \, dy, \quad (5.32)
\]
with $\varphi_n(s) = n \varphi(ns)$, $\varphi_n(x) = n^N \varphi(nx)$, where $\varphi \in C_b^\infty(\mathbb{R})$, and $\varphi \in C_b^\infty(\mathbb{R}^N)$ are such that $\chi(-1/2, 1/2) \leq \varphi \leq \chi(-1, 1)$, $\chi_{B(0, 1/2)} \leq \varphi \leq \chi_{B(0, 1)}$, and
\[ \|\varphi\|_{L^1(\mathbb{R})} = \|\varphi\|_{L^1(\mathbb{R}^N)} = 1. \]

Our aim consists in showing that \( D_t w_n - Aw_n \) converges to 0 locally uniformly in \((0,T) \times \mathbb{R}^N\). For this purpose, we observe that the smoothness of \( w \) implies that \( Aw \in C([0,T] \times \mathbb{R}^N) \) and that \( Aw_n \) converges to \( Aw \), locally uniformly in \((0,T) \times \mathbb{R}^N\). Therefore, we can restrict ourselves to showing that \((D_t - B)w_n\) converges to \((D_t - B)w \in C([0,T] \times \mathbb{R}^N)\) locally uniformly in \((0,T) \times \mathbb{R}^N\). To show such a property we first prove that

\[
(D_t - B)w_n = (D_t w - Bw) \ast \varrho_n \varphi_n + \text{Tr}(B)w_n + \varpi \ast (\varrho_n B \varphi_n), \quad (5.33)
\]

in \((1/n, T-1/n) \times \mathbb{R}^N\), where \(D_t w - Bw\) denotes any continuous extension of \(D_t w - Bw\) to the whole of \(\mathbb{R} \times \mathbb{R}^N\). An integration by parts shows that \((5.33)\) holds true in the case where \(\varpi \in C^{1,2}_b([0,T] \times \mathbb{R}^N)\). In the general case we approximate \(\varpi\) by a sequence of smooth functions \(\{\varpi_m\}_{m \in \mathbb{N}} \subset C^{1,2}_b(\mathbb{R} \times \mathbb{R}^N)\) converging to \(\varpi\) locally uniformly in \(\mathbb{R} \times \mathbb{R}^N\), and define the function \(\varpi_m\) \((m, n \in \mathbb{N})\) according to \((5.32)\), with \(\varpi\) being replaced by \(\varpi_m\).

As is immediately seen,

\[
(D_t - B)w_m^m = (D_t w^m - Bw^m) \ast \varrho_n \varphi_n + \text{Tr}(B)w_m^m + w^m \ast (\varrho_n B \varphi_n),
\]

in \([0,T] \times \mathbb{R}^N\). Since \(w^m\) converges to \(\varpi\) locally uniformly, then \(\text{Tr}(B)w_m^m + w^m \ast (\varrho_n B \varphi_n)\) converges to \(\text{Tr}(B)w_n + \varpi \ast (\varrho_n B \varphi_n)\), locally uniformly in \([0,T] \times \mathbb{R}^N\). As far as the term \((D_t w^m - Bw^m) \ast \varrho_n \varphi_n\) is concerned, we observe that

\[
((D_t w^m - Bw^m) \ast \varrho_n \varphi_n)(t,x) \quad (5.34)
\]

\[
= \int_{[t-1/n,t+1/n] \times B(x,1/n)} (D_t w^m - Bw^m)(s,y) \varrho_n(t-s) \varphi_n(x-y) \, ds \, dy,
\]

for any \((t,x) \in \mathbb{R} \times \mathbb{R}^N\). Suppose \(t \in (1/n, T-1/n)\); then the function \(\varrho_n(t-\cdot) \varphi_n(x-\cdot)\) is compactly supported in \((0,T) \times \mathbb{R}^N\). Therefore, since \(D_t w^m - Bw^m\) converges to \(D_t w - Bw \in C([0,T] \times \mathbb{R}^N)\) in the sense of distributions, letting \(m\) go to \(+\infty\) in \((5.34)\), we deduce that

\[
\lim_{m \to +\infty} ((D_t w^m - Bw^m) \ast \varrho_n \varphi_n)(t,x)
= \int_{[t-1/n,t+1/n] \times B(x,1/n)} (D_t w - Bw)(s,y) \varrho_n(t-s) \varphi_n(x-y) \, ds \, dy
= \int_{\mathbb{R} \times \mathbb{R}^N} (D_t w - Bw)(t-s,x-y) \varrho_n(s) \varphi_n(y) \, ds \, dy,
\]

for any \((t,x) \in (1/n, T-1/n) \times \mathbb{R}^N\) and \((5.33)\) follows. Now, arguing as in the proof of Lemma 5.6, from \((5.33)\) we can conclude that \(D_t w_n - Bw_n\).
converges locally uniformly in \((0, T) \times \mathbb{R}^N\) to \(D_t w - Bw\). It follows that 
\(D_t w_n - Aw_n\) converges to 0, locally uniformly in \((0, T) \times \mathbb{R}^N\), as \(n\) tends to 
\(+\infty\). To conclude the proof, we observe that for any \(n \in \mathbb{N}\) and any \(t_0 > 0\),
the function \(z_n = w_n(\cdot + t_0, \cdot)\) is a classical solution to the Cauchy problem
\[
\left\{
\begin{array}{l}
D_t z_n(t, x) = Az_n(t, x) + f_n(t + t_0, x), \quad t \in [0, T - t_0], \quad x \in \mathbb{R}^N, \\
z_n(0, x) = w_n(t_0, x), \quad x \in \mathbb{R}^N,
\end{array}
\right.
\]
(5.35)
where \(f_n := D_t z_n - Az_n\). Since \(f_n \in C_b^{1,2}([0, T] \times \mathbb{R}^N)\), the function \(h_n : [0, T - t_0] \times \mathbb{R}^N \to \mathbb{R}\) defined by
\[
h_n(t, x) = (T(t)w_n(t_0, \cdot))(x) + \int_0^t (T(t - s)f_n(t + t_0 - s, \cdot))(x)ds,
\]
for any \((t, x) \in [0, T - t_0] \times \mathbb{R}^N\), is a classical solution to the Cauchy problem
(5.35) (see also the proof of Corollary 5.9). The maximum principle (see Proposition 3.1) implies that
\(h_n = z_n\). Now, since \(z_n, f_n(\cdot + t_0, \cdot)\) converge, locally uniformly in \([0, T - t_0] \times \mathbb{R}^N\) to \(w(\cdot + t_0, \cdot)\) and 0, respectively, taking
Proposition 4.1 into account, we get
\[
w(t + t_0, x) = (T(t)w(t_0, \cdot))(x), \quad (t, x) \in [0, T - t_0] \times \mathbb{R}^N,
\]
for any \(t_0 \in (0, T)\). Letting \(t_0\) go to 0, and recalling that \(u(0, \cdot) = 0\), we deduce that \(w \equiv 0\). This finishes the proof.

To conclude this section we show that the results in Theorem B can be improved if \(f\) and \(g\) are smoother. In such a case the function \(u\) in (5.31) is
a classical solution to the Cauchy problem (1.2), in the sense that the time
derivative of \(u\) as well as all the space derivatives involved in the definition of \(Au\) exist in the classical sense and are bounded and continuous in \([0, T] \times \mathbb{R}^N\).

**Corollary 5.9.** Suppose, in addition to the assumptions of Theorem B, that \(f \in C_b^{\theta + 2/3}(\mathbb{R}^N)\), \(g(t, \cdot) \in C_b^\theta(\mathbb{R}^N)\) for some \(\theta > 1/3\) and any \(t \in [0, T]\),
with \(\sup_{t \in [0, T]} \|g(t, \cdot)\|_{C_b^\theta(\mathbb{R}^N)} < +\infty\). Then, the function \(u\) in (5.31) is once
continuously differentiable in \([0, T] \times \mathbb{R}^N\) with respect to the time and space
variables. Moreover, for any \(t \in [0, T]\), \(u(t, \cdot)\) belongs to \(C_b^{\theta + 2/3}(\mathbb{R}^N)\) and
there exists a positive constant \(C\) such that
\[
\sup_{t \in [0, T]} \|u(t, \cdot)\|_{C_b^{\theta + 2/3}(\mathbb{R}^N)} \leq C \sup_{t \in [0, T]} \|g(t, \cdot)\|_{C_b^\theta(\mathbb{R}^N)}.
\]
(5.36)

**Proof.** As in the proof of Theorem B, we can restrict ourselves to dealing
with the function \(v\) therein defined. Using an interpolation argument similar
to that used in the proof of Proposition 5.4, from (3.5) and (3.6) we deduce
that, for any \(0 \leq \alpha \leq \beta \leq 3\) and any \(\omega > 0\), there exists a positive constant \(C = C(\omega)\) such that

\[
\|T(t)\|_{L(C^\alpha_b(\mathbb{R}^N), C^\beta_b(\mathbb{R}^N))} \leq C \frac{e^{\omega t}}{t^{(\beta - \alpha)/2}}, \quad t > 0. \tag{5.37}
\]

Now, applying [5, Theorem 2.2] with \(Y_0 = C^\theta_b(\mathbb{R}^N), Y_1 = C^\alpha_b(\mathbb{R}^N), Y_2 = C^{2+\alpha}_b(\mathbb{R}^N)\), we deduce that \(v(t, \cdot) \in C^{\theta/2+3}_b(\mathbb{R}^N)\) for any \(t > 0\), and there exists a positive constant \(C_1\), independent of \(g\), such that

\[
\sup_{t \in [0, T]} \|v(t, \cdot)\|_{C^{\theta/2+3}_b(\mathbb{R}^N)} \leq C_1 \sup_{t \in [0, T]} \|g(t, \cdot)\|_{C^\beta_b(\mathbb{R}^N)}. \tag{5.38}
\]

Therefore, (5.36) follows from (3.6) and (5.38). To conclude the proof, let us show that \(v\) is continuously differentiable with respect to time in \((0, T] \times \mathbb{R}^N\) and \(D_t v = Av + g\). For this purpose we observe that for any \(h > 0\) sufficiently close to 0, one has

\[
\frac{v(t+h, x) - v(t, x)}{h} = \frac{1}{h} \int_t^{t+h} (T(t + h - s)g(s, \cdot))(x)ds 
+ \frac{1}{h} \int_0^t ((T(t + h - s) - T(t - s))g(s, \cdot))(x)ds.
\tag{5.39}
\]

As is easily seen, the first integral term in the right-hand side of (5.39) tends to \(g(t, x)\) as \(h\) tends to 0\(^+\). As far as the second term is concerned, we observe that

\[
((T(t + h - s) - T(t - s))g(s, \cdot))(x) = \int_0^h (D_tT(t - s + r)g(s, \cdot))(x)ds 
= \int_0^h (AT(t - s + r)g(s, \cdot))(x)ds.
\]

From (5.1) with \(\alpha = \theta/3\) and \(\beta = 2/3 + \theta/6\), we deduce that

\[
\|D_{ij}T(t - s + r)g(s, \cdot)\|_{\infty} \leq c(t-s) \sup_{t \in [0,T]} \|g(t, \cdot)\|_{C^\beta_b(\mathbb{R}^N)}, \tag{5.40}
\]

for any \(i, j \leq r\) and some function \(c \in L^1(0,T)\). Now, (5.37) (with \((\alpha, \beta) = (\theta, 1)\)) and (5.40) imply that, for any compact set \(K \subset \mathbb{R}^N\), there exists a function \(\hat{c}_K \in L^1(0,T)\) such that

\[
\|AT(t - s + r)g(s, \cdot)\|_{C^0(K)} \leq \hat{c}_K(t-s) \sup_{t \in [0,T]} \|g(t, \cdot)\|_{C^\beta_b(\mathbb{R}^N)}. \tag{5.41}
\]

Hence, by (5.41) and the dominated convergence theorem, we obtain that the second term in the right-hand side of (5.39) tends to \(Av(t, x)\) as \(h\) tends to 0\(^+\). Therefore, \(v(\cdot, x)\) is differentiable from the right in \((0, T)\) and
$D_t^+ v(t,x) = Av(t,x) + g(t,x)$. Since $v \in C_b([0,T] \times \mathbb{R}^N)$ and it satisfies (1.4) and (5.38), then arguing as in Remark 2.5, we can conclude that $Av \in C([0,T] \times \mathbb{R}^N)$. Therefore, $D_t^+ v$ is continuous in $[0,T] \times \mathbb{R}^N$ and this implies that $v$ is continuously differentiable with respect to time in $[0,T] \times \mathbb{R}^N$ and $D_t v = Av + g$. This completes the proof. \hfill \Box

References