CHAPTER 2

Non-Elliptic Regularity Results

Let $L$ be an operator of the form in (1.1.8). In the preceding chapter, we studied Kolmogorov’s forward equation $\partial_t \mu(t) = L^* \mu(t)$ and showed that, in great generality, solutions exist in the sense that, for each initial $\nu \in M_1(\mathbb{R}^N)$ there is a continuous $t \mapsto \mu(t) \in M_1(\mathbb{R}^N)$ such that (1.1.11) holds for all $\varphi \in C^2_c(\mathbb{R}^N; \mathbb{C})$. Moreover, in the generality which we have been working, this is the best sense in which one can hope to have solutions.

For example, if $a \equiv 0$ and $b(x) \equiv b$, then the only solution, in the sense of Schwartz distributions, to the corresponding forward equation with initial value $\nu$ is $t \mapsto \mu(t)$ where

$$\langle \varphi, \mu(t) \rangle = \int \varphi(x + tb) \nu(dx).$$

In particular, if $\nu = \delta_x$, then $\mu(t) = \delta_{x+tb}$.

Now suppose that $(t,x) \mapsto P(t,x)$ is a continuous transition probability which solves Kolmogorov’s forward equation in the sense that (1.2.29) is well-defined and holds for all $\varphi \in C^2(\mathbb{R}^N; \mathbb{C})$ with bounded second order derivatives. For example, by Theorems 1.2.25, this will be the case when $a = \sigma \sigma^\top$ and $\sigma$ and $b$ are uniformly Lipschitz continuous. Although we know that, for each $(t, x)$, $P(t, x)$ will, in general, be no better than a probability measure, when tested against smooth functions, it may, nonetheless, possess smoothness properties as a function of its backward variable $x$. For instance, in the preceding example, $\langle \varphi, P(t, x) \rangle = \varphi(x + bt)$, and so $x \mapsto \langle \varphi, P(t, x) \rangle$ will be just as smooth as $\varphi$. The reason why we are interested in this possibility is that it provides solutions to Kolmogorov’s backward equation

(2.0.1) \[ \partial_t u = Lu \quad \text{in} \quad (0, \infty) \times \mathbb{R}^N \quad \text{with} \quad \lim_{t \to 0} u(t) = \varphi. \]

To be precise, let $\varphi \in C_b(\mathbb{R}^N; \mathbb{C})$ be given, and set

(2.0.2) \[ u_\varphi(t, x) \equiv \langle \varphi, P(t, x) \rangle. \]

Further, assume that $u_\varphi(t, \cdot)$ has bounded second order derivatives for each $t \in (0, \infty)$. Then, by the Chapman–Kolmogorov equation and (1.2.29),

$$u_\varphi(t + h, x) - u_\varphi(t, x) = \langle u_\varphi(t), P(h, x) \rangle - u_\varphi(t, x)$$

$$= \int_0^h \langle Lu_\varphi(t), P(t + \tau, x) \rangle d\tau,$$
and so \( u_\varphi \) is differentiable with respect to \( t \in (0, \infty) \) and \( \partial_t u_\varphi = Lu_\varphi \). Hence, since \( u_\varphi(t, x) \to \varphi(x) \) as \( t \searrow 0 \), it follows that \( u_\varphi \) is a solution to (2.0.1).

In this chapter, we will find conditions under which \( P(t, x) \) has the smoothness property required by the preceding line of reasoning. Because, as we will see, a sufficiently good existence result for Kolmogorov’s backward equation implies uniqueness for his forward equation, we also will have found conditions under which the construction made in Chapter 1 yields the one and only solution his forward equation.

2.1 Smoothness of \( P(t, x) \)

Let \( \sigma : \mathbb{R}^N \to \text{Hom}(\mathbb{R}^M, \mathbb{R}^N) \) and \( b : \mathbb{R}^N \to \mathbb{R}^N \) be continuous maps which are \( \ell \)-times differentiable for some \( \ell \geq 1 \). Further, assume \( \sigma \) and \( b \) have bounded first order derivatives and that their higher order derivatives are slowly increasing (i.e., have at most polynomial growth). Let \( (t, x) \mapsto P(t, x) \) be the corresponding transition probability function constructed in § 1.1.3. In this section we will show that \( x \mapsto \langle \varphi, P_n(t, x) \rangle \) is \( \ell \)-times continuously differentiable whenever \( \varphi \) is a function which is \( \ell \)-times differentiable with derivatives that are slowly increasing.

2.1.1. Representation of Derivatives of \( P_n(t, x) \): Let \( \sigma \) and \( b \) be as above, and define \( \{ P_n : n \geq 0 \} \) accordingly, as in § 1.1.3. That is, \( P_n(0, x) = \delta_x \) and, for \( m2^{-n} \leq t \leq (m+1)2^{-n} \),

\[
P_n(t, x) = \int Q(t-m2^{-n}, y) P_n(m2^{-n}, x, dy),
\]

where \( Q(t, x) \) is the distribution of

\[
\omega \in \mathbb{R}^M \mapsto F(t, x, \omega) \equiv x + t^2 \sigma(x) \omega + tb(x) \in \mathbb{R}^N
\]

under \( \gamma \).

Next, set \( \Omega = (R^M)^{Z^+} \) and \( \Gamma = \gamma^{Z^+} \). Define \( X_n : [0, \infty) \times \mathbb{R}^N \times \Omega \to \mathbb{R}^N \) so that

\[
X_n(0, x, \omega) = x \& X_n(t, x, \omega) = F(t-m2^{-n}, X_n(m2^{-n}, \omega), \omega_{m+1})
\]

for \( m2^{-n} \leq t \leq (m+1)2^{-n} \) and \( \omega = (\omega_1, \ldots, \omega_m, \ldots) \in \Omega \).

Notice that, for \( t \leq m2^{-n} \), \( \omega \mapsto X_n(t, x, \omega) \) depends only on \( (\omega_1, \ldots, \omega_m) \) and is therefore independent of \( \sigma(\{ \omega_\tau : \tau \geq m+1 \}) \) under \( \Gamma \).

**Lemma 2.1.4.** For each \( n \geq 0 \) and \( (t, x) \in [0, \infty) \times \mathbb{R}^N \), \( P_n(t, x) \) is the distribution of \( \omega \mapsto X_n(t, x, \omega) \) under \( \Gamma \). That is, for any slowly increasing \( \varphi \in C(\mathbb{R}^N; \mathbb{C}) \),

\[
\langle \varphi, P_n(t, x) \rangle = \int \varphi \circ X_n(t, x, \omega) \Gamma(d\omega).
\]
PROOF: Obviously, there is nothing to do when \( t = 0 \). Now assume the result for \( t \in [0, m2^{-n}] \), and let \( t = m2^{-n} + \tau \), where \( \tau \in [0, 2^{-n}] \). Then, by the preceding remark,

\[
\langle \varphi, P_n(t, x) \rangle = \int \langle \varphi, Q(\tau, y) \rangle P_n(m2^{-n}, x, dy) 
= \int \langle \varphi, Q(\tau, X_n(m2^{-n}, x, \omega) \rangle \Gamma(d\omega)
= \int \left( \int \varphi \circ F(\tau, X_n(m2^{-n}, x, \omega), \omega_{m+1}) \gamma(d\omega_{m+1}) \right) \Gamma(d\omega)
= \int \varphi \circ X_n(t, x, \omega) \Gamma(d\omega).
\]

Hence, by induction, on \( m \geq 0 \), we are done. □

In order to facilitate the next step, we have to introduce some notation. Given \( \ell \geq 1 \) and \( \Phi \in C^\ell(\mathbb{R}^N; \mathbb{R}^N) \), we use \( \nabla^\ell \Phi \) to denote the mapping from \( \mathbb{R}^N \) into \( \text{Hom}(\mathbb{R}^N \otimes \cdots \otimes \mathbb{R}^N) \) determined by

\[
\nabla^\ell \Phi(x)\xi_1 \otimes \cdots \otimes \xi_\ell = \left. \frac{\partial^\ell}{\partial t_1 \cdots \partial t_\ell} \Phi \left( x + \sum_{k=1}^\ell t_k \xi_k \right) \right|_{t_1 = \cdots = t_\ell = 0}
\]

for \((\xi_1, \ldots, \xi_\ell) \in (\mathbb{R}^N)^\ell\). Using the chain rule and induction on \( \ell \), one can check that if \( \Phi \in C^\ell(\mathbb{R}^N; \mathbb{R}^N) \) and \( \Psi \in C^\ell(\mathbb{R}^N; \mathbb{R}^N) \), then

\[
\nabla^\ell (\Psi \circ \Phi) = \sum_{k=1}^\ell \sum_{\beta \in \mathcal{B}(k, \ell)} c_\beta (\nabla^k \Phi) \circ \Psi \nabla^\otimes \beta \Phi,
\]

where \( \mathcal{B}(k, \ell) = \{ \beta \in (\mathbb{Z}^+)^k : \sum_{j=1}^k \beta_j = \ell \} \), \( \nabla^\otimes \beta \Phi = \nabla^{\beta_1} \Phi \otimes \cdots \otimes \nabla^{\beta_k} \Phi \), and the coefficients \( c_\beta \) are given by the prescription: \( c_\beta = 1 \) if \( \beta \in \mathcal{B}(1, \ell) \cup \mathcal{B}(\ell, \ell) \) and, for \( \beta \in \mathcal{B}(k, k+1) \) with \( 2 \leq k \leq \ell \),

\[
c_\beta = \delta_{\beta_1,1} c_{(\beta_2, \ldots, \beta_k+1)} + \sum_{\alpha \in \mathcal{P}(\beta)} c_\alpha,
\]

where \( \mathcal{P}(\beta) \) is the set of \( \alpha \in \mathcal{B}(k, \ell) \) with the property that \( \beta_j = \delta_{i,j} + \alpha_j \), \( 1 \leq j \leq k \), for some \( 1 \leq i \leq k \).

Using (2.1.3), (2.1.5), and induction on \( m \geq 0 \), one sees that, for each \((t, \omega) \in [m2^{-n}, (m+1)2^{-n}] \times \Omega \), \( x \mapsto X_n(t, x, \omega) \) is \( \ell \)-times continuously differentiable. In fact, if \( X_n^\ell(t, x, \omega) = \nabla^\ell X_n(t, x, \omega) \) (the differentiation being with respect to the \( x \)-variables only), then

\[
X_n^\ell(t, x, \omega)
\]

(2.1.6) \[ \sum_{k=1}^\ell \sum_{\beta \in \mathcal{B}(k, \ell)} c_\beta \nabla^k F(\tau, X_n(m2^{-n}, x, \omega), \omega_{m+1}) X_n^\otimes (m2^{-n}, x, \omega) \]

for \( m2^{-n} \leq t \leq (m+1)2^{-n} \) and \( \tau = t - m2^{-n} \).
where \(X_n^{\otimes \beta} \equiv X_n^{\beta_1} \otimes \cdots \otimes X_n^{\beta_k}\). In addition, under the conditions we have imposed on \(\sigma\) and \(b\), we know that there exist \(C_1 < \infty\) and \(r \geq 1\) for which

\[
|F(\tau, x, \omega)| \vee \max_{1 \leq k, \ell \leq \ell} \|\nabla^k F(\tau, x, \omega)\|_{H.S.} \leq C_1(1 + |x|)^r(1 + |\omega|)
\]

when \((\tau, x, \omega) \in [0, 1] \times \mathbb{R}^N \times \mathbb{R}^M\), and so one can use (2.1.6) to check that, for each \(r \geq 1\) and \(R > 0\),

\[
\left( \int |X_n(t, x, \omega)| \vee \max_{1 \leq k, \ell \leq \ell} \|X^k_n(t, x, \omega)\|_{H.S.} \right)^{2r} \Gamma(d\omega) < \infty.
\]

**Lemma 2.1.7.** Given a slowly increasing \(\varphi \in C(\mathbb{R}^N; \mathbb{C})\), set

\[
u_{\varphi, n}(t, x) = \int \varphi(y) P_n(t, x, dy).
\]

If \(\varphi\) is \(\ell\)-times differentiable and its derivatives are slowly increasing, then, for each \(n \geq 0\) and \(t \in [0, \infty)\), \(\nu_{\varphi, n}(t, \cdot)\) is \(\ell\)-times continously differentiable and

\[
\nabla^\ell \nu_{\varphi, n}(t, x) = \sum_{k=1}^{\ell} \sum_{\beta \in \mathcal{B}(k, \ell)} c_\beta \int \nabla^k \varphi \left( X_n(t, x, \omega) \right) X_n^{\otimes \beta}(t, x, \omega) \Gamma(d\omega).
\]

**Proof:** The result is an immediate consequence of (2.1.5) and an application of the preceding integrability estimate to justify differentiation under the integral sign. \(\square\)

**2.1.2. Representation of Derivatives of \(P(t, x)\):** Starting from the result in Lemma 2.1.7, we now want to prove that (cf. (2.0.2)) \(u_{\varphi}(t, \cdot)\) is as smooth as \(\varphi\) and the coefficients \(\sigma\) and \(b\) are. For this purpose, we will use the contents of §1.3.1 to pass to a limit as \(n \to \infty\).

Given \(\ell \geq 1\), set \(E_\ell = \text{Hom}(\mathbb{R}^N; \mathbb{R}^N)\), \(E^{(\ell)} = E_1 \times \cdots \times E_\ell\), and

\[
\|J\|_{E^{(\ell)}} = \left( \sum_{k=1}^{\ell} \|J_k\|_{H.S.}^2 \right)^{\frac{1}{2}} \text{ for } J = (J_1, \ldots, J_\ell) \in E^{(\ell)}.
\]

Next, define \(\sigma_\ell : \mathbb{R}^N \times E^{(\ell)} \to \text{Hom}(\mathbb{R}^M; E_\ell)\) and \(b_\ell : \mathbb{R}^N \times E^{(\ell)} \to E_\ell\) so that

\[
\sigma_\ell(x, J) \omega = \sum_{k=1}^{\ell} \sum_{\beta \in \mathcal{B}(k, \ell)} c_\beta \nabla^k \sigma(x, \omega) J^{\otimes \beta}
\]

and

\[
b_\ell(x, J) = \sum_{k=1}^{\ell} \sum_{\beta \in \mathcal{B}(k, \ell)} c_\beta \nabla^k b(x) J^{\otimes \beta}
\]
for $x \in \mathbb{R}^N$ and $J = (J_1, \ldots, J_k) \in \mathcal{E}^{(t)}$. Notice that

$$
\|\sigma_t(x, J)\|_{H.}^2 + |b_t(x, J)|^2 
\leq C_t \left( 1 + |x|^{2r_t} + \sum_{k=1}^{\ell-1} \|J_k\|_{H.}^2 + \|J_\ell\|_{H.}^2 \right)
$$

(2.1.8)

for some $C_t < \infty$ and $r_t \in [1, \infty)$, where, when $\ell = 1$, $r_1 = 1$ and $\sum_{k=1}^{\ell-1}$ is taken to be 0. Finally, define $\sigma^{(t)} : \mathbb{R}^N \times \mathcal{E}^{(t)} \to \hom(\mathbb{R}^M; \mathbb{R}^N \times \mathcal{E}^{(t)})$ and $b^{(t)} : \mathbb{R}^N \times \mathcal{E}^{(t)} \to \mathbb{R}^N \times \mathcal{E}^{(t)}$ by

$$
\sigma^{(t)}(x, J) = \begin{pmatrix}
\sigma(x) \\
\sigma_1(x, J_1) \\
\vdots \\
\sigma_\ell(x, J) 
\end{pmatrix}
\quad \text{and} \quad
b^{(t)}(x, J) = \begin{pmatrix}
b(x) \\
b_1(x, J_1) \\
\vdots \\
b_\ell(x, J) 
\end{pmatrix}.
$$

**Lemma 2.1.9.** Let

$$(x, J) \in \mathbb{R}^N \times \mathcal{E}^{(t)} \to P_n^{(t)}(t, (x, J)) \in M_1(\mathbb{R}^N \times \mathcal{E})$$

be defined as in §1.1.3 relative to $\sigma^{(t)}$ and $b^{(t)}$. Then, there exists an $r^{(t)} \in [1, \infty)$ such that, if $r \in [r^{(t)}, \infty)$,

$$
\sup_{n \geq 0} \int \left( |y|^{2r} + \|J'\|_{E^{(t)}}^2 \right) P_n^{(t)}(t, (x, J), dy \times dJ') \leq A_r^{(t)} e^{A_r^{(t)} (1 + |x|^{2r} + \|J\|_{E^{(t)}}^{2r})}
$$

for some $A_r^{(t)} < \infty$. Moreover, there exists a continuous transition probability function $(t, (x, J)) \sim P^{(t)}(t, (x, J))$ which satisfies

$$
\int \left( |y|^{2r} + \|J'\|_{E^{(t)}}^2 \right) P^{(t)}(t, (x, J), dy \times dJ') \leq A_r^{(t)} e^{A_r^{(t)} (1 + |x|^{2r} + \|J\|_{E^{(t)}}^{2r})}
$$

for $r \geq r^{(t)}$ and to which $\{P_n^{(t)} : n \geq 0\}$ converges in the sense that

$$
\langle \Phi_n, P_n^{(t)}(t, (x_n, J_n)) \rangle \to \langle \Phi, P^{(t)}(t, (x, J)) \rangle
$$

if $(x_n, J_n) \to (x, J)$ in $\mathbb{R}^N \times \mathcal{E}^{(t)}$ and $\{\Phi_n : n \geq 0\} \subseteq C(\mathbb{R}^N \times \mathcal{E}^{(t)}; \mathbb{C})$ is a uniformly slowly increasing sequence (i.e., $|\Phi_n(x, J)| \leq C(1 + |x|^{r} + \|J\|_{E^{(t)}})$ for some $C < \infty$ and $r \in [1, \infty)$) which converges to $\Phi$ uniformly on compacts.

**Proof:** When $\ell = 1$, the coefficients satisfy (1.2.26), and therefore Theorem 1.2.25 and Lemma 1.3.36 apply. In particular, this allows one to prove the required result as a special case of Theorem 1.3.43. To prove the result for $\ell \geq 2$, use (2.1.8) and Theorem 1.3.43 at each step of an inductive argument. □
The connection between this and the preceding sections is provided by the following. Namely, referring to the notation in §2.1.1, define

\[(t, x, \omega) \in [0, \infty) \times \mathbb{R}^N \times \Omega \longrightarrow X_n^\ell(t, x, \omega) = \left( X_n(t, x, \omega), X_n^1(t, x, \omega), \ldots, X_n^\ell(t, x, \omega) \right) \in \mathbb{R}^N \times E^\ell, \]

and let 0 be the origin in \(E^\ell\).

**Lemma 2.1.10.** For each \(n \geq 0\), \(P_n^\ell(t, (x, 0))\) is the distribution of \(\omega \sim X_n^\ell(t, x, \omega)\) under \(\Gamma\).

**Proof:** Set

\[F(t, (x, J), \omega) = \left( x \atop J \right) + t^{\ell} \sigma^\ell(x, J)\omega + t b^\ell(x, J)\]

for \((x, J) \in \mathbb{R}^N \times E^\ell\) and \(\omega \in \mathbb{R}^M\). Then, by (2.1.3) and (2.1.6),

\[X_n^\ell(t, x, \omega) = F(t, X_n^\ell(t, x, \omega), \omega_{m+1})\]

for \(m2^{-n} \leq t \leq (m+1)2^{-n}\)

In addition, \(X_n^\ell(0, x, \omega) = \left( x \atop 0 \right)\). Hence, the identification of \(P_n^\ell(t, (x, 0))\) as the distribution of \(\omega \sim X_n^\ell(t, x, \omega)\) can be made by exactly the same argument as was used to prove Lemma 2.1.4. □

By combining Lemmas 2.1.9 and 2.1.10 with Lemma 2.1.7, we arrive at our goal.

**Theorem 2.1.11.** Let \(\sigma : \mathbb{R}^N \longrightarrow \text{Hom}(\mathbb{R}^M; \mathbb{R}^N)\) and \(b : \mathbb{R}^N \longrightarrow \mathbb{R}^N\) be \(\ell\)-times continuously differentiable functions for some \(\ell \geq 1\). Further, assume that

\[
\begin{align*}
\left(\|\sigma(0)\|_{\text{H.S.}} + |b(0)|^2\right) & \vee \left(\|\nabla \sigma(x)\|_{\text{H.S.}}^2 + \|\nabla b(x)\|_{\text{H.S.}}^2\right) \\
& \vee \max_{2 \leq k \leq \ell} \left(\|\nabla^k \sigma(x)\|_{\text{H.S.}}^2 + \|\nabla^k b(x)\|_{\text{H.S.}}^2\right) \\
& \leq C^{(\ell)}
\end{align*}
\]

for some \(r^{(\ell)} \in [0, \infty)\) and \(C^{(\ell)} < \infty\). Then, for each \(r \geq r^{(\ell)}\) and any \(\varphi \in C^r(\mathbb{R}^N; \mathbb{C})\) satisfying \(\max_{1 \leq k \leq \ell} \|\nabla^k \varphi(x)\|_{\text{H.S.}} \leq C(1 + |x|^{2r})\) for some \(C < \infty\), \(u_\varphi(t, \cdot) \in C^r(\mathbb{R}^N; \mathbb{C})\) for each \(t \geq 0\), and

\[
\nabla^r u_\varphi(t, x) = \sum_{k=1}^\ell \sum_{\beta \in \mathfrak{B}(k, \ell)} c_{\beta} \int \nabla^k \varphi(y) J^{\otimes \beta} P^\ell(t, (x, 0), dy \times dJ).
\]

Moreover, there exists an \(A < \infty\), depending only on \(N, \ell, r^{(\ell)}, C^{(\ell)}\), and \(r\), such that

\[\|\nabla^r u_\varphi(t, x)\|_{\text{H.S.}} \leq AC e^{At}(1 + |x|^r).\]

**Corollary 2.1.13.** Suppose that \(\sigma\) and \(b\) are as in the preceding and that \(\ell \geq 2\). Then, for each \(\varphi \in C^2(\mathbb{R}^N; \mathbb{C})\) with slowly increasing second derivatives, the function \(u_\varphi\) given by (2.0.2) is an element of \(C^{1, 2}(\mathbb{R}^N; \mathbb{C})\) which satisfies (2.0.1) and, together with its derivatives, is slowly increasing.
2.2 Kolmogorov’s Equations, the Existence–Uniqueness Duality

Proof: In view of Theorem 2.1.11, it suffices for us to check that $u_\varphi$ is differentiable with respect to $t$ and that it satisfies (2.0.1). But, using (1.3.46) and the Chapman–Kolmogorov equation, it is easy to check that

$$u_\varphi(t+h,x) - u_\varphi(t,x) = \langle u_\varphi(t), P(h,x) \rangle - u_\varphi(t,x) = \int_0^h \langle Lu_\varphi(\tau), P(\tau,x) \rangle \, d\tau,$$

and clearly this is all that we need. □

2.2 Kolmogorov’s Equations, the Existence–Uniqueness Duality

In this section we will show how existence results for one of Kolmogorov’s equations lead to uniqueness results for the other. For anyone familiar with operator theory, what is going on here is a manifestation of the duality between existence for an operator and uniqueness for its adjoint.

2.2.1. The Basic Duality Result: Let $L$ be the operator defined from $a : \mathbb{R}^N \to \text{Hom}(\mathbb{R}^N; \mathbb{R}^N)$ and $b : \mathbb{R}^N \to \mathbb{R}^N$, as in (1.1.2). Before stating our main result, we will need the contents of the following lemma.

Lemma 2.2.14. Assume that $a$ and $b$ satisfy the condition in (1.1.21), and let $t \mapsto \mu(t)$ be any continuous solution to (1.1.11) for $\varphi \in C^2_c(\mathbb{R}^N; \mathbb{C})$. Then, there is a non-decreasing function $r \in (0,\infty) \mapsto \lambda_r \in [0,\infty)$, depending only on $r$ and the $\lambda$ in (1.1.21), such that

$$\lambda_0 = 0 \quad \text{and} \quad \int (1 + |y|^{2r}) \mu(t,dy) \leq e^{\lambda_r t} \int (1 + |x|^{2r}) \nu(dx).$$

In particular, if $\int (1 + |x|^{2r}) \nu(dx) < \infty$, then (1.1.11) continues to hold for any $\varphi \in C^2(\mathbb{R}^N; \mathbb{C})$ satisfying

$$\sup_{x \in \mathbb{R}^N} \frac{|\varphi(x)| + \left(\nabla \varphi(x), a(x) \nabla \varphi(x)\right)^{\frac{1}{2}}_{\mathbb{R}^N} + |L\varphi(x)|}{1 + |x|^{2r}} < \infty.$$

Proof: We begin by proving the last statement under the assumption that $t \mapsto \int (1 + |y|^{2r}) \mu(t,dy)$ is bounded on compact intervals. To this end, choose $\psi \in C^\infty_c(\mathbb{R}^N; [0,1])$ so that $\psi = 1$ on $B(0,1)$ and $\psi = 0$ off of $B(0,2)$, and set $\psi_R(x) = \psi(R^{-1}x)$ and $\varphi_R = \psi_R \varphi$ for $R \geq 1$. Clearly, by Lebesgue’s Dominated Convergence Theorem,

$$\langle \varphi, \mu(t) \rangle - \langle \varphi, \nu \rangle = \lim_{R \to \infty} \int_0^t \langle L\varphi_R, \mu(\tau) \rangle \, d\tau,$$

and so we need only show that the limit on the right equals $\int_0^t \langle \varphi, \mu(\tau) \rangle \, d\tau$.

But

$$L\varphi_R(x) = \psi_R(x) L\varphi + \left(\nabla \varphi(x), a(x) \nabla \psi_R(x)\right)_{\mathbb{R}^N} + \varphi(x) L\psi_R(x),$$
and so, again by Lebesgue’s Dominated Convergence Theorem, it suffices to show that
\[
\int_0^t \left( (\nabla \varphi(y), a(y) \nabla \psi_R(y))_{\mathbb{R}^N} + \varphi(x) L\psi_R(y) \right) \mu(\tau, dy) \, d\tau = 0.
\]
For this purpose, first note that, because \(a(y)\) is symmetric and non-negative definite,
\[
\left| (\nabla \varphi(y), a(y) \nabla \psi_R(y))_{\mathbb{R}^N} \right|^2 
\leq (\nabla \varphi(y), a(y) \nabla \phi(y))_{\mathbb{R}^N} (\nabla \psi_R(y), a(y) \nabla \psi_R(y))_{\mathbb{R}^N}.
\]
Because
\[
(\nabla \psi_R(y), a(y) \nabla \psi_R(y))_{\mathbb{R}^N} 
\leq \|\nabla \psi\|_2^2 \sup_{a \leq |y| \leq 2R} \text{Trace}(a(y')) \mathbf{1}_{B(0,R)}(y),
\]
we now see that there is a \(C < \infty\) such that
\[
\left| (\nabla \varphi(y), a(y) \nabla \psi_R(y))_{\mathbb{R}^N} \right| \leq C(1 + |y|^{2r}) \mathbf{1}_{B(0,1)}(y).
\]
Similarly, one can show \(|L\psi_R(y)|\) is also bounded by a constant, independent of \(R \geq 1\), times \((1 + |y|^{2r}) \mathbf{1}_{B(0,1)}(y)\) Thus, the desired conclusion follows after yet another application of Lebesgue’s Dominated Convergence Theorem.

Returning to the proof of (2.2.15), consider the functions
\[
\varphi_{r,\epsilon}(x) = \left( \frac{1 + |x|^2}{1 + \epsilon |x|^2} \right)^r \quad \text{for } r \in (0, \infty) \text{ and } \epsilon \in (0, 1).
\]
Because
\[
\nabla \varphi_{r,\epsilon}(x) = \frac{2r(1 - \epsilon)\varphi_{r,\epsilon}(x)}{(1 + \epsilon |x|^2)(1 + |x|^2)^{-\frac{r}{2}}} \frac{x}{(1 + \epsilon |x|^2)^{\frac{r}{2}}}
\]
and
\[
\nabla^2 \varphi_{r,\epsilon}(x) = \frac{2r(1 - \epsilon)\varphi_{r,\epsilon}(x)}{1 + |x|^2} \times \left[ \frac{I}{1 + \epsilon |x|^2} + \left( \frac{4(1 - \epsilon)}{(1 + \epsilon |x|^2)^3} + \frac{2(r - 1)}{(1 + \epsilon |x|^2)^2} \right) \frac{x \otimes x}{1 + |x|^2} \right].
\]
In particular, by the preceding result (applied with the \(r\) there equal to 0) and shows that
\[
\langle \varphi_{r,\epsilon}, \mu(t) \rangle \leq \langle \varphi_{r,\epsilon}, \nu \rangle + \lambda_e \int_0^t \langle \varphi_{r,\epsilon}, \mu(\tau) \rangle \, d\tau
\]
for some \(\lambda_e < \infty\) which is independent of \(\epsilon\) and depends only on the \(\lambda\) in (1.1.21). Hence, by Gronwall’s inequality,
\[
\int (1 + |y|^{2r}) \mu(t, dy) = \lim_{\epsilon \searrow 0} \langle \varphi_{r,\epsilon}, \mu(t) \rangle \leq e^{\lambda_e t} \int (1 + |x|^2)^r \nu(dx). \quad \square
\]
2.2 Kolmogorov’s Equations, the Existence–Uniqueness Duality  

**Theorem 2.2.16.** Assume that \(a\) and \(b\) satisfy the growth condition in (1.1.21), and suppose that \(t \in [0, \infty) \rightarrow \mu(t) \in \mathcal{M}_1(\mathbb{R}^N)\) is a continuous solution to (1.1.11) for all \(\varphi \in C_c^2(\mathbb{R}^N; \mathbb{C})\). If \(\int |x|^{2r} \nu(dx) < \infty\) for some \(r \in [0, \infty)\) and if, for some \(T \in (0, \infty)\), \(u \in C^{1,2}(\mathbb{R}^N; \mathbb{C})\) is a solution to \(\partial_t u = Lu\) which satisfies

\[
\sup_{t \in [0,T] \times \mathbb{R}^N} \frac{|u(t,x)| + (\nabla u(t,x), a(x)\nabla u(t,x))_{\mathbb{R}^N}^{\frac{1}{2}}}{1 + |x|^{2r}} < \infty,
\]

then \(\langle u(0), \mu(T) \rangle = \langle u(T), \nu \rangle\).

**Proof:** By Lemma 2.2.14 we know that (2.2.15) holds. Now define \(\psi_R, R \geq 1\), as in the preceding proof and set \(u_R = \psi_R u\). Then, by Lebesgue’s Dominated Convergence Theorem,

\[
\langle u(0), \mu(T) \rangle - \langle u(T), \nu \rangle = \lim_{R \to \infty} \left( \langle u_R(0), \mu(T) \rangle - \langle u_R(T), \nu \rangle \right)
\]

\[
= \lim_{R \to \infty} \int_0^T \frac{d}{dt} \langle u_R(T-t), \mu(t) \rangle \, dt
\]

\[
= \lim_{R \to \infty} \int_0^T \langle Lu_R(T-t) - \psi_R Lu(T-t), \mu(t) \rangle \, dt
\]

\[
= \lim_{R \to \infty} \int_0^T \langle u(T-t) \psi_R + (\nabla u(T-t), a \nabla \psi_R)_{\mathbb{R}^N}, \mu(t) \rangle \, dt.
\]

By the argument used in the preceding proof, both \(u(T-t,x) \psi_R(x)\) and \(|\nabla u(T-t,x), a(x) \nabla \psi_R(x)|_{\mathbb{R}^N}\) are bounded by a constant, independent of \(t \in [0,T]\) and \(R \geq 1\), times \((1 + |x|^{2r})1_{R(0,R) \in x}\). Hence, by Lebesgue’s Dominated Convergence Theorem,

\[
\lim_{R \to \infty} \int_0^T \langle u(T-t) \psi_R + (\nabla u(t), a \nabla \psi_R)_{\mathbb{R}^N}, \mu(t) \rangle \, dt = 0. \quad \square
\]

**Remark 2.2.17.** By considering the Markov process whose transition probability function is \((t,x) \rightarrow P(t,x)\) and using Doob’s Stopping Time Theorem, one can dispense with the growth condition on \((\nabla u, a \nabla u)_{\mathbb{R}^N}\).

**Corollary 2.2.18.** Let \(a\) and \(b\) be as in the preceding, and assume that \((t,x) \rightarrow P(t,x)\) is an associated continuous transition probability function. Then, for each slowly increasing \(\varphi \in C(\mathbb{R}^N; \mathbb{C})\) there is at most one solution \(u \in C^{1,2}(0, \infty) \times \mathbb{R}^N; \mathbb{C})\) which satisfies the growth condition

\[
\sup_{t \in [0,T] \times \mathbb{R}^N} \frac{|u(t,x)| + (\nabla u(t,x), a(x)\nabla u(t,x))_{\mathbb{R}^N}^{\frac{1}{2}}}{1 + |x|^{2r}} < \infty, \quad T \in (0, \infty),
\]

for some \(r \in [0, \infty)\). In fact, \(u(t,x) = \langle \varphi, P(t,x) \rangle\).
2.2.19, one sees that

\[ \text{Proof: Using the same argument as we used in the proof of Corollary 2.1.13.} \]

\[ \text{Again let } a \text{ and } b \text{ be as in Theorem 2.2.16, and assume that for each } \varphi \in C^2_c(\mathbb{R}^N; \mathbb{R}) \text{ there is a solution } u \in C^{1,2}((0, \infty) \times \mathbb{R}^N; \mathbb{R}) \text{ to (2.0.1) satisfying} \]

\[ \sup_{t \in [0,T]} \frac{\|\nabla^2 u(t,x)\|_{\text{H.S.}}}{1 + |x|^{2r}} < \infty \]

for some \( r \in [0, \infty) \) and all \( T \in (0, \infty) \). Then, for each \( \nu \in M_1(\mathbb{R}^N) \) with moments of all orders, there is at most one \( t \rightarrow \mu(t) \) satisfying (1.1.11) for all \( \varphi \in C^2_c(\mathbb{R}^N; \mathbb{R}) \). In particular, if \( \{P_n : n \geq 0\} \) is defined relative to \( a \) and \( b \) as in §1.2, then there is a unique continuous transition probability function \( (t,x) \rightarrow P(t,x) \) to which \( \{P_n : n \geq 0\} \) converges in the sense that

\[ \langle \varphi_n, P_n(t_n,x_n) \rangle \longrightarrow \langle \varphi, P(t,x) \rangle \]

whenever \( (t_n,x_n) \rightarrow (t,x) \) in \( [0,\infty) \times \mathbb{R}^N \) and \( \{\varphi_n : n \geq 0\} \subseteq C(\mathbb{R}^N; \mathbb{C}) \) is a uniformly slowly increasing sequence which converges to \( \varphi \) uniformly on compacts.

\[ \text{Proof: Using the same argument as we used in the proof of Corollary 2.2.19, one sees that} \]

\[ \langle u(T), \nu \rangle = \lim_{s \searrow 0} \langle u(s), \mu(T) \rangle = \langle \varphi, \mu(T) \rangle. \]

Thus \( \mu(T) \) is uniquely determined for each \( T \in [0, \infty) \). Given this uniqueness statement and (1.3.44), we know that, for each \( (t,x) \), \( \{P_n(t,x) : n \geq 0\} \) can have at most limit and therefore must converge in the required sense to a continuous transition function. \( \square \)

The following is an immediate consequence of our present considerations. See Corollary 2.4.29 for a related result in which the hypotheses here on \( \sigma \) are replaced by hypotheses on \( a \) itself.

\[ \text{Corollary 2.2.20. Assume that } a = \sigma \sigma^\top \text{ and that } \sigma \text{ and } b \text{ are twice continuously differentiable functions with the properties that } \|\nabla \sigma\|_{\text{H.S.}} \text{ and } \|\nabla^2 \sigma\|_{\text{H.S.}} \text{ are bounded and } \|\nabla^2 \sigma\|_{\text{H.S.}} \text{ are slowly increasing. Then, for each } \nu \in M_1(\mathbb{R}^N) \text{ with moments of all orders, there is precisely one continuous } t \rightarrow \mu(t) \text{ which satisfies (1.1.11) for all } \varphi \in C^2_c(\mathbb{R}^N; \mathbb{R}). \text{ In particular, there is precisely one continuous transition probability function } (t,x) \rightarrow P(t,x) \text{ which solves (1.2.29) for all } \varphi \in C^2_c(\mathbb{R}^N; \mathbb{R}).} \]

\[ \text{Proof: This is a simple conclusion from the preceding together with Corollary 2.1.13.} \]
2.3 Square Roots

One of the most severe weaknesses of the theory developed in §§ 1.2 and 2.1 is that it relies on our being able to take a good square root of the matrix $a$. Indeed, when $a$ can degenerate and we need its square root to be smooth, this flaw is fatal: in general, there simply is no way to take a smooth square root of $a$, even when $a$ is analytic. Nonetheless, as we are about to see, there is no problem when $a$ is non-degenerate and, if all that we require is Lipschitz continuity, there is no problem as long as $a$ is twice differentiable.

2.3.1. The Non-Degenerate Case: In the following, when $a$ is non-negative definite and symmetric matrix, $a^{1/2}$ denotes the non-negative definite, symmetric square root of $a$.

**Lemma 2.3.21.** Assume that $a : \mathbb{R}^N \rightarrow \text{Hom} (\mathbb{R}^N; \mathbb{R}^N)$ is an $\ell$-times continuously differentiable, symmetric, non-negative definite matrix-valued function for some $\ell \geq 1$. Further, assume that $a(x)$ is strictly positive definite, and let $\lambda_{\text{min}}(x)$ denote its smallest eigenvalue. Then $a^{1/2}$ is $\ell$-times continuously differentiable at $x$ and there is a universal constant $C_\ell < \infty$ such that

$$\|\nabla^\ell a^{1/2}(x)\|_{\text{H.S.}} \leq C_\ell \lambda_{\text{min}}(x)^{1/2} \sum_{k=1}^\ell \left( \frac{\|a(x)\|_{\text{H.S.}}}{\lambda_{\text{min}}(x)} \right)^k,$$

where $\|a(x)\|_{\text{H.S.}} = \max_{1 \leq k \leq \ell} \|\nabla^k a(x)\|_{\text{H.S.}}$.

**Proof:** Given $R \in (1, \infty)$, let $A(R)$ be the space of symmetric, $N \times N$-matrices $a$ satisfying $R^{-1}I < a < RI$, and define $\Phi_R : A(R) \rightarrow A(R^{1/2})$ by

$$\Phi_R(a) = R^{1/2} \sum_{m=0}^\infty \left( \begin{array}{c} 1 \\ m \end{array} \right) (R^{-1}a - I)^m,$$

where $\left( \begin{array}{c} 1 \\ m \end{array} \right)$ is the coefficient of $x^m$ in the Taylor’s expansion of $x \mapsto (1+x)^{1/2}$ at 0. By testing its action on the eigenvectors of $a$, one sees that $a^{1/2} = \Phi_R(a)$. Hence, by (2.1.5), in order to prove the result, it suffices to check that there exists a $B_\ell < \infty$ such that, for any symmetric matrices $E_1, \ldots, E_\ell \in \text{Hom}(\mathbb{R}^N; \mathbb{R}^N)$ with $\|E_k\|_{\text{H.S.}} = 1$,

$$\left\| \frac{\partial^\ell}{\partial E_1 \cdots \partial E_\ell} \Phi_R(a) \right\| \leq B_\ell \lambda(a)^{1/2-\ell}.$$

But clearly

$$\frac{\partial^\ell}{\partial E_1 \cdots \partial E_\ell} \Phi_R(a) = R^{1/2} \sum_{m=\ell}^\infty \left( \begin{array}{c} 1 \\ m \end{array} \right) D^{(m)}(a),$$

where $D^{(m)}(a)$ denotes the $m$th derivative of $a$.
where \( D^{(m)}(a) \) is the sum of \( \frac{m!}{m-\ell!} \) terms each of which is the product of \( m \) factors: \( \ell \) being the \( E_k \)'s and the other \( (m-\ell) \) being \( \frac{a}{R} - I \). In particular,

\[
\|D^{(m)}(a)\|_{\text{H.S.}} \leq \|\frac{a}{R} - I\|_{\text{op}} \prod_{k=1}^{m-\ell} \|E_k\|_{\text{H.S.}} \leq (1 - \frac{\lambda}{R})^{m-\ell},
\]

where \( \lambda \) is the smallest eigenvalue of \( a \).

Hence

\[
\left\| \frac{\partial^\ell}{\partial E_1 \cdots \partial E_{\ell}} \Phi_R(a) \right\| \leq R^\frac{1}{2} \sum_{m=\ell}^{\infty} \left( \frac{1}{2} \right) \left( \frac{m!}{(m-\ell)!} \right) (1 - \frac{\lambda}{R})^{m-\ell}
\]

\[
= (-1)^{\ell+1} \frac{d}{d\xi} \left( R^\frac{1}{2} \sum_{m=0}^{\infty} \left( \frac{1}{2} \right) \left( \frac{m}{m} \right) \left( \frac{1}{R} - 1 \right)^m \right) \bigg|_{\xi=\lambda}
\]

\[
= (-1)^{\ell+1} \frac{d}{d\xi} \xi^\frac{1}{2} \bigg|_{\xi=\lambda} = \ell! \left( \frac{1}{2} \right) \lambda^{\frac{1}{2}-\ell}. \]
In particular, by Lemma 2.3.21, this means that $a^{\frac{1}{2}}$ is differentiable and that the required estimate will follow once we show that

(*) \[ \|\partial_x a^{\frac{1}{2}}(x)\|_{\text{L.S.}} \leq \sqrt{2K} \]

for all $x \in \mathbb{R}^N$ and $\xi \in \mathbb{S}^{N-1}$.

To prove (*), let $x$ be given, and work with an orthonormal coordinate system in which $a(x)$ is diagonal. Then, from $a = a^{\frac{1}{2}} a^{\frac{1}{2}}$ and Leibnitz’s rule, one obtains

\[ \partial_x a_{ij} = \partial_x a_{ij}^{\frac{1}{2}} \left( \sqrt{a_{ii}(x)} + \sqrt{a_{jj}(x)} \right). \]

Hence, because $\sqrt{\alpha} + \sqrt{\beta} \geq \sqrt{\alpha + \beta}$ for all $\alpha, \beta \geq 0$,

\[ |\partial_x a_{ij}^{\frac{1}{2}}(x)| \leq \frac{|\partial_x a_{ij}(x)|}{\sqrt{a_{ii}(x) + a_{jj}(x)}}. \]

To complete the proof of (*), set

\[ f_{\pm}(t) = a_{ii}(x + t\xi) \pm 2a_{ij}(x + t\xi) + a_{jj}(x + t\xi), \]

and apply (2.3.22) to get

(2.3.24) \[ |\partial_x a_{ij}(x)| \leq \frac{|f'_{\pm}(0)| + |f''_{\pm}(0)|}{4} \leq \frac{2\|\partial^2 a_{ij}\|_u}{2\|\partial^2 a_{ij}\|_u}, \]

which, in conjunction with the preceding, leads first to

\[ |\partial_x a_{ij}^{\frac{1}{2}}(x)| \leq \sqrt{2\|\partial^2 a_{ij}\|_u}, \]

and thence to (*). \qed

As we mentioned earlier, even if $a$ is real analytic as a function of $x$, it will not in always be possible to find a smooth $\sigma$ such that $a = \sigma \sigma^{\top}$. The reasons for this have their origins in classical algebraic geometry. Indeed, D. Hilbert showed that it is not possible to express every non-negative polynomial as a finite sum of squares of polynomials. After combining this fact with Taylor’s theorem, one realizes that it rules out the existence of a smooth choice of $\sigma$. Of course, the problems arise only at the places where $a$ degenerates: away from degeneracies, as the proof of Lemma 2.3.21 shows, the entries of $a^{\frac{1}{2}}$ are analytic functions of the entries of $a$.

In spite of the preceding, it should be recognized that, although Lipschitz continuity is the best one can do in general, there are circumstances in which one can better, especially if one is willing to consider square roots other than the non-negative, symmetric one. For example, consider $a : \mathbb{R}^2 \rightarrow \text{Hom}(\mathbb{R}^2; \mathbb{R}^2)$ given by $a(x)_{ij} = x_i x_j$. Because $a(x)^2 = |x|^2a(x)$, $a^{\frac{1}{2}}(x) = |x|^{-1}a(x)$, which is not smooth at the origin. On the other hand, $\sigma(x) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ is smooth everywhere and $a(x) = \sigma(x)\sigma(x)^\top$. 
2.4 Another Approach

As the preceding section makes clear, basing our theory on properties of $\sigma$ instead of $a$ creates a serious weakness in degenerate situations. Thus, it is important to know that there is a way of getting estimates of the sort in §2.1 without relying on the existence of a smooth $\sigma$. The key ideas here is due to O. Oleinik.

2.4.1. The Weak Minimum Principle: In this subsection, $a : \mathbb{R}^N \rightarrow \text{Hom}(\mathbb{R}^N; \mathbb{R}^N)$ and $b : \mathbb{R}^N \rightarrow \mathbb{R}^N$ are bounded, measurable functions, $a(x)$ is symmetric and non-negative definite for each $x \in \mathbb{R}^N$, and $L$ is defined from $a$ and $b$ as in (1.1.2).

**Lemma 2.4.25.** If $u \in C^{1,2}((0, T] \times \mathbb{R}^N; \mathbb{R})$ is bounded below and satisfies

$$(L - \partial_t)u \leq 0 \text{ on } (0, T] \times \mathbb{R}^N \text{ and } \lim_{R \to 0} \inf_{x \in B(0, R)} u(t, x) \geq 0 \text{ for } R > 0,$$

then $u \geq 0$ on $(0, T] \times \mathbb{R}^N$.

**Proof:** We first show that if $(L - \partial_t)u < 0$ on $(0, T] \times \mathbb{R}^N$, then $u$ cannot achieve a minimum there. To this end, suppose that $u$ achieves a minimum at $(t_0, x_0)$. Then, by the first and second derivative tests, $\partial_t u(t_0, x_0) \leq 0$, $\nabla u(t_0, x_0) = 0$, and $\nabla^2 u(t_0, x_0)$ would be symmetric and non-negative definite. But this would lead to the contradiction that $(L - \partial_t)u(t_0, x_0) \leq 0$, since the product of two symmetric, non-negative matrices has a non-negative trace.

We now return to the original assumptions. For $\delta > 0$ and $\epsilon > 0$, set

$$u_{\epsilon, \delta}(t, x) = u(t, x) + \delta t + e^t|x|^2.$$

Then

$$(L - \partial_t)u(t, x) \leq -\delta - \epsilon e^t \left[ |x|^2 - \text{Trace}(a(x)) + 2(x, b(x))_{\mathbb{R}^N} \right]
\leq -\delta + \epsilon e^t \left[ |b(x)|^2 + \text{Trace}(a(x)) \right].$$

Thus, for each $\delta > 0$, there exists an $\epsilon(\delta) > 0$ such that $(L - \partial_t)u_{\epsilon, \delta} < 0$ when $\epsilon < \epsilon(\delta)$. In addition, for any $\delta > 0$ and $\epsilon > 0$, $u_{\epsilon, \delta}(t, x) \rightarrow \infty$ uniformly $t \in (0, T]$ as $|x| \rightarrow \infty$. Hence, if $u_{\epsilon, \delta}$ were to become negative somewhere in $(0, T] \times \mathbb{R}^N$, then there would exist a sequence $\{(t_n, x_n) : n \geq 1\} \subseteq (0, T] \times \mathbb{R}^N$ which converges to some $(t_0, x_0) \in [0, T] \times \mathbb{R}^N$ and for which

$$0 > \inf\{u_{\epsilon, \delta}(t, x) : (t, x) \in (0, T] \times \mathbb{R}^N\} = \lim_{n \to \infty} u_{\epsilon, \delta}(t_n, x_n).$$

Moreover, by the second assumption made about $u$, $t_0 > 0$, and so $(t_0, x_0) \in (0, T] \times \mathbb{R}^N$ would be a point at which $u_{\epsilon, \delta}$ achieves its minimum value. In particular, because this cannot happen when $\epsilon < \epsilon(\delta)$, we have now shown that $u_{\epsilon, \delta} \geq 0$ for $\epsilon < \epsilon(\delta)$. Now let $\epsilon \searrow 0$ and then $\delta \searrow 0$. □
Theorem 2.4.26. Assume that \( u \in C^{1,2}((0, T] \times \mathbb{R}^N; \mathbb{R}) \cap C_b((0, T] \times \mathbb{R}^N; \mathbb{R}) \) satisfies
\[
(L - \partial_t + c(t))u \geq -g(t) \quad \text{on } (0, T] \times \mathbb{R}^N,
\]
where \( c \) and \( g \) are continuous, \( \mathbb{R} \)-valued functions on \([0, T]\). Then
\[
u(t, x) \leq \|u(0, \cdot)\|_u e^{C(t)} + \int_0^t g(\tau)e^{C(t) - C(\tau)} d\tau,
\]
where \( C(t) \equiv \int_0^t c(\tau) d\tau \).

Proof: Set
\[
v(t, x) = \|u(0, \cdot)\|_u - u(t, x)e^{-C(t)} - \int_0^t g(\tau)e^{-C(\tau)} d\tau,
\]
and check that \((L - \partial_t)v \leq 0\) on \((0, T] \times \mathbb{R}^N\) and \( \lim_{t \downarrow 0} \inf_{x \in \mathbb{R}^N} v(t, x) \geq 0 \). Hence, by Lemma 2.4.25, \( v \geq 0 \), which is equivalent to the asserted estimate. \( \square \)

2.4.2. Oleinik’s Estimate:
Recall the notation
\[
\partial^\alpha f = \frac{\partial\|\alpha\| f}{\partial_{x_1}^{\alpha_1} \cdots \partial_{x_N}^{\alpha_N}} \quad \text{for } \alpha \in \mathbb{N}^N,
\]
where \( \|\alpha\| \equiv \sum_1^N \alpha_i \). In addition, define
\[
\|f\|_{u}^{(\ell)} = \left( \sum_{\|\alpha\| \leq \ell} \|\partial^\alpha f\|_u^2 \right)^{\frac{1}{2}}
\]
for \( \ell \geq 0 \). Finally, write \( \beta \leq \alpha \) to mean that \( \beta_k \leq \alpha_k \) for each \( 1 \leq k \leq N \).

Let \( a, b, \) and \( L \) be as in the preceding section. An important role in Oleinik’s argument is played by the following.

Lemma 2.4.27. Given any symmetric \( H \in \text{Hom}(\mathbb{R}^N; \mathbb{R}^N) \) and \( e \in \mathbb{S}^{N-1} \),
\[
\left( \sum_{i,j=1}^N \partial_{x} a(x)_{ij} H_{ij} \right)^2 \leq 4N^2 \Lambda(x, e) \text{Trace}(Ha(x)H),
\]
where \( \Lambda(e) \equiv \sup_{x \in \mathbb{R}^N} \|\partial_x^2 a(x)\|_{\text{op}} \).
Let $x \in \mathbb{R}^N$ be given, choose an orthogonal $O$ so that $O \alpha(x)O$ is diagonal, and set $\tilde{a}(y) = O^T a(y)O$ and $H = OHO^T$. Then

$$\sum_{ij} \partial_{\alpha} \tilde{a}_{ij}(x) H_{ij} = \sum_{ij} \partial_{\alpha} \tilde{a}_{ij}(x) H_{ij},$$

and so, by Schwarz's inequality and (2.3.24),

$$\left( \sum_{ij} \partial_{\alpha} \tilde{a}_{ij}(x) H_{ij} \right) \leq N^2 \sum_{ij} \left( \partial_{\alpha} \tilde{a}_{ij}(x) \right)^2 H_{ij}^2$$

$$\leq 4N^2 \Lambda(e) \sum_{ij} \tilde{a}(x) H_{ij}^2 = 4N^2 \Lambda(e) = 4N^2 \Lambda(e) \text{Trace}(H \alpha(x) H). \quad \square$$

**Theorem 2.4.28.** Assume that $a$ and $b$ have $\ell \geq 2$ bounded continuous derivatives, and let $c \in \mathcal{C}_{W}^{\ell}(\mathbb{R}; \mathbb{R})$ be given. Given $u \in \mathcal{C}_{W}^{\ell+2}(\{0, T\} \times \mathbb{R}^N; \mathbb{R})$ with $\partial^\alpha u \in \mathcal{C}_{W}^{1,0}(\{0, T\} \times \mathbb{R}^N; \mathbb{R})$ for each $\|\alpha\| \leq \ell$, set

$$g(t, x) \equiv (\partial_t - L - c) u(t, x).$$

Then

$$\|u(t, \cdot)\|^{(\ell)}_{\mathcal{U}} \leq A_{\ell} \left( \|u(0, \cdot)\|^{(\ell)}_{\mathcal{U}} + \sup_{\tau \in [0, t]} \|g(\tau, \cdot)\|^{(\ell)}_{\mathcal{U}} \right) e^{B_{\ell} t},$$

where $A_{\ell} < \infty$ and $B_{\ell} < \infty$ can be chosen to depend only on $N, \ell, \|a\|^{(\ell)}_{\mathcal{U}}, \|b\|^{(\ell)}_{\mathcal{U}},$ and $\|c\|^{(\ell)}_{\mathcal{U}}$.

**Proof:** For $\|\alpha\| \leq \ell$, set $u_\alpha = \partial^\alpha u$ and $g_\alpha = \partial^\alpha g$, and use Leibnitz’s rule and induction to check that

$$\partial_t u_\alpha = Lu_\alpha + \frac{1}{2} \sum_{k}^{k'} \sum_{i,j=1}^{N} \alpha_k (\partial_k a_{ij}) \partial_x \partial_x u_{\alpha_k} + \sum_{\gamma \leq \alpha} c_{\alpha, \gamma} u_{\gamma} = -g_\alpha,$$

where $\sum_{k}^{k'}$ denotes summation over $1 \leq k \leq N$ with $\alpha_k > 0$, the $c_{\alpha, \gamma}$’s are linear combinations of $a, b, c$, and their derivatives up to order $\|\alpha\|$, and, when $\alpha_k > 0, \alpha_k^{\gamma}$ is obtained from $\alpha$ by replacing $\alpha_k$ by $\alpha_k - 1$ and leaving the other coordinates unchanged. Now remember that $Lf^2 = 2fLf + (\nabla f, a \nabla f)\mathbb{R}^N$, and use this to see that if $1 \leq \ell' \leq \ell$ and $w_{\ell'} \equiv \sum_{\|\alpha\| = \ell'} u_{\alpha}^2$, then

$$\partial_t w_{\ell'} = Lw_{\ell'} + \sum_{\|\alpha\| = \ell'} \sum_{k}^{k'} \sum_{i,j=1}^{N} \alpha_k u_{\alpha} (\partial_x a_{ij}) \partial_x \partial_x u_{\alpha_k}$$

$$- \sum_{\|\alpha\| = \ell'} (\nabla u_{\alpha}, a \nabla u_{\alpha})_{\mathbb{R}^N} + 2 \sum_{\|\alpha\| = \ell'} \left( u_{\alpha} g_{\alpha} + \sum_{\beta \leq \alpha} c_{\alpha, \beta} u_{\alpha} u_{\beta} \right).$$
At first sight, the preceding looks bad: there are $\ell' + 1$ order derivatives appearing where they should not be. Oleinik’s crucial observation is that Lemma 2.4.27 allows us to eliminate these. Namely, set $\Lambda = \sup_{e \in \mathbb{R}^{N-1}} \Lambda(e)$, and conclude that

$$
\left( \sum_{k}^{N} \sum_{i,j=1}^{N'} \alpha_k u_\alpha (\partial_{x_k} a_{ij}) \partial_{z_i} \partial_{x_j} u_\alpha^k \right)^2
\leq u_\alpha^2 \left( \sum_{k}^{N} \sum_{i,j=1}^{N'} \alpha_k^2 \right) \left( \sum_{k}^{N} \sum_{i,j=1}^{N'} (\partial_{x_k} a_{ij}) \partial_{z_i} \partial_{x_j} u_\alpha^k \right)^2
\leq 4N^2(\ell')^2 \Lambda u_\alpha^2 \sum_{k} \text{Trace}(\nabla^2 u_\alpha^k a \nabla^2 u_\alpha^k)_{\mathbb{R}^N}.
$$

In particular, this shows that there is a $C_{\ell'} < \infty$, depending only on $N$, $\ell'$, and $\|a\|_u^{(2)}$, such that

$$\left( \sum_{\|\alpha\| = \ell'} \sum_{i,j=1}^{N'} \alpha_k u_\alpha (\partial_{x_k} a_{ij}) \partial_{z_i} \partial_{x_j} u_\alpha^k \right)^2
\leq 4C_{\ell'} w_{\ell'} \sum_{\|\alpha\| = \ell'} (\nabla u_\alpha, a \nabla u_\alpha)_{\mathbb{R}^N},$$

and so

$$\sum_{\|\alpha\| = \ell'} \sum_{i,j=1}^{N'} \alpha_k u_\alpha (\partial_{x_k} a_{ij}) \partial_{z_i} \partial_{x_j} u_\alpha^k \leq 4 \sum_{\|\alpha\| = \ell'} (\nabla u_\alpha, a \nabla u_\alpha)_{\mathbb{R}^N} \leq C_{\ell'} w_{\ell'}.$$

Using the preceding estimate in the equation for $\partial_t w_{\ell'}$, we arrive at

$$(L - \partial_t + C_{\ell'}) w_{\ell'} \geq -2 \sum_{\|\alpha\| = \ell'} \sum_{\beta \leq \alpha} c_{\alpha,\beta} u_\alpha u_\beta - 2 \sum_{\|\alpha\| = \ell'} u_\alpha g_\alpha.$$

Note that

$$2 \left| \sum_{\|\alpha\| = \ell'} \sum_{\beta \leq \alpha} c_{\alpha,\beta} u_\alpha u_\beta \right| \leq 2 \left( \max_{\|\alpha\| = \ell'} \sum_{\beta \leq \alpha} \|c_{\alpha,\beta}\|_u \right) w_{\ell'}$$

$$+ \left( \max_{\|\beta\| < \ell'} \sum_{\|\alpha\| - \ell'} \|c_{\alpha,\beta}\|_u^2 \right) \left( \|u(t, \cdot)\|_u^{(\ell' - 1)} \right)^2,$$

whereas

$$2 \left| \sum_{\|\alpha\| = \ell'} u_\alpha g_\alpha \right| \leq w_{\ell'} + \left( \|g(t, \cdot)\|_u^{(\ell')} \right)^2.$$
Hence, after adjusting $C_{\ell'}$, we have

$$\tag{**} (L - \partial_t + C_{\ell'} u) v \geq -C_{\ell'} \left( \|u(t, \cdot)\|_{H^{(\ell'-1)}_a} \right)^2 - \left( \|g(t, \cdot)\|_{H^{(\ell')}_{\ell}} \right)^2. $$

Given (**), one proceeds by induction on $\ell'$. By Theorem 2.4.26, we know that

$$\|u(t, \cdot)\|_a \leq e^{B_0 t} \left( \|u(0, \cdot)\|_a + \sup_{\tau \in [0, t]} \|g(\tau, \cdot)\|_a \right),$$

where $B_0 = \|c\|_a$. Next, one applies Theorem 2.4.26 to (**) to carry out each inductive step. \(\square\)

**Corollary 2.4.29.** Under the conditions in Theorem 2.4.28, for each $\varphi \in C^1_c(\mathbb{R}^N; \mathbb{R})$ there is a bounded sequence $\{u_n : n \geq 1\} \subseteq C^{1,\ell}(0, \infty \times \mathbb{R}^N; \mathbb{R})$ with the properties that $\lim_{n \to \infty} \|u_n(0, \cdot) - \varphi\|_a = 0$,

$$\sup_{t \in [0, T]} \sup_{n \geq 1} \|u_n(t, \cdot)\|^{(\ell)}_a < \infty$$

$$\sup_{t \in [0, T]} \| (L - \partial_t) u_n(t, \cdot)\|_a \longrightarrow 0 \quad \text{for all } T \in (0, \infty).$$

Moreover, there exists a unique $u_\varphi \in C_b(\mathbb{R}^N; \mathbb{R})$ to which any such sequence converges uniformly on $[0, T] \times \mathbb{R}^N$ for each $T \in (0, \infty)$. Finally, for any $\nu \in M_1(\mathbb{R}^N)$ and any continuous $t \to \mu_t$ which satisfies (1.1.11) for all $\varphi \in C^\infty_c(\mathbb{R}^N; \mathbb{R})$, $\langle \varphi, \mu_t \rangle = \langle \varphi_n(t), \nu \rangle$, $t \geq 0$. In particular, there is precisely one $t \to \mu_t$ which satisfies (1.1.11) for all $\varphi \in C^\infty_c(\mathbb{R}^N; \mathbb{R})$.

**Proof:** Choose $\rho \in C^\infty_c(B(0, 1); [0, \infty))$ with total integral 1, set $\rho_\epsilon(x) = \epsilon^{-N} \rho(\epsilon^{-1} x)$, take $a_n(x) = \rho_\frac{1}{\epsilon} \ast a(x) + \frac{1}{\epsilon} I$, $b_n = \rho_\frac{1}{\epsilon} \ast b$, and define $L_n$ accordingly. Because, by Lemma 2.3.21, the positive definite, symmetric square root $\sigma_n$ of $a_n$ has bounded derivatives of all orders, Corollary 2.2.20 guarantees that $L_n$ determines a transition probability function $(t, x) \to P_n(t, x)$ and Corollary 2.1.13 says that $(t, x) \to u_n(t, x) = \langle \varphi_n, P_n(t, x) \rangle$ is smooth and solves $\partial_t u_n = L_n u_n$. Furthermore, because $\varphi_n, \sigma_n, \text{ and } b_n$ and their derivatives are all bounded, when applied to $u_n$, (2.1.12) holds with $r = 0$. Thus, for each $n \geq 1$, $\sup_{t \in [0, T]} \|u_n(t, \cdot)\|^{(\ell)}_a < \infty$. But, because the derivatives of $a_n$, $b_n$, and $\varphi_n$ up to order $\ell$ are bounded independent of $n \geq 1$, Theorem 2.4.28 allows us to say that $\sup_{n \geq 1} \sup_{t \in [0, T]} \|u_n(t, \cdot)\|^{(\ell)}_a < \infty$. In particular, $g_n \equiv (L - \partial_t) u_n \longrightarrow 0$ uniformly on $[0, T] \times \mathbb{R}^N$. Finally, by Theorem 2.4.26, for $1 \leq m \leq n$,

$$|u_n(t, x) - u_m(t, x)| \leq \int_0^t \|g_n(\tau) - g_m(\tau)\|_a \, d\tau,$$

and so there is a $u_\varphi \in C([0, \infty) \times \mathbb{R}^N; \mathbb{R})$ to which $\{u_n : n \geq 1\}$ converges uniformly on $[0, T] \times \mathbb{R}^N$ for each $T \in (0, \infty)$.  

40 2 Non-Elliptic Regularity Results
To complete the proof at this point, all that we have to check is that 
\[ \langle \varphi, \mu_t \rangle = \langle u_\varphi(t), \nu \rangle \]
for any solution to (1.1.11) for all \( \varphi \in C^\infty_c(\mathbb{R}^N; \mathbb{R}) \).
But, because \( a \) and \( b \) are bounded, (1.1.11) will continue to hold for all \( \varphi \in C^2_b(\mathbb{R}^N; \mathbb{R}) \), and so, just as in the proof of Theorem 2.2.16, one can check that
\[
\frac{d}{dt} \langle u_n(T-t), \mu_t \rangle = \langle g_n(T-t), \mu_t \rangle, \quad t \in [0, T],
\]
which leads to
\[
\langle \varphi, \mu_T \rangle - \langle u_\varphi(T), \nu \rangle = \lim_{n \to \infty} \left( \langle \varphi, \mu_T \rangle - \langle u_n(T), \nu \rangle \right)
= \lim_{n \to \infty} \int_0^T (g_n(T-t), \mu_t) \, dt = 0. \quad \square
\]

The following is an essentially immediate consequence of the preceding.

**Corollary 2.4.30.** The function \( u_\varphi \) described in Corollary 2.4.29 is in \( C^1,\ell^{-1}_b(\mathbb{R}^N; \mathbb{R}) \) and \( \nabla^{\ell-1} u_\varphi \) is uniformly Lipschitz continuous with respect to \( x \) on \( [0, T] \times \mathbb{R}^N \) for each \( T \in (0, \infty) \). In particular, if \( \ell \geq 3 \), then \( \partial_t u_\varphi = Lu_\varphi \).