# Analysis & PDE

Volume 8 No. 1 2015

msp

# Analysis & PDE

msp.org/apde





See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2015 is US $205/year for the electronic version, and $390/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

**mathematical sciences publishers**

**nonprofit scientific publishing**

http://msp.org/

ANALYSIS AND PDE
Vol. 8, No. 1, 2015
dx.doi.org/10.2140/apde.2015.8.1

# HÖLDER CONTINUITY AND BOUNDS FOR FUNDAMENTAL SOLUTIONS TO NONDIVERGENCE FORM PARABOLIC EQUATIONS

SEIICHIRO KUSUOKA

We consider nondegenerate second-order parabolic partial differential equations in nondivergence form with bounded measurable coefficients (not necessary continuous). Under certain assumptions weaker than the Hölder continuity of the coefficients, we obtain Gaussian bounds and Hölder continuity of the fundamental solution with respect to the initial point. Our proofs employ pinned diffusion processes for the probabilistic representation of fundamental solutions and the coupling method.

## 1. Introduction and main result

Let $a(t,x)=(a_{ij}(t,x))$ be a symmetric $d\times d$-matrix-valued bounded measurable function on $[0,\infty)\times\mathbb{R}^d$ which is uniformly positive-definite, i.e.,

$$\Lambda^{-1}I \le a(t,x) \le \Lambda I, \tag{1-1}$$

where $\Lambda$ is a positive constant and $I$ is the unit matrix. Let $b(t,x)=(b_i(t,x))$ be an $\mathbb{R}^d$-valued bounded measurable function on $[0,\infty)\times\mathbb{R}^d$ and $c(t,x)$ a bounded measurable function on $[0,\infty)\times\mathbb{R}^d$. Consider the parabolic partial differential equation

$$\begin{cases} \dfrac{\partial}{\partial t}u(t,x) = \dfrac{1}{2}\displaystyle\sum_{i,j=1}^d a_{ij}(t,x)\frac{\partial^2}{\partial x_i\partial x_j}u(t,x) + \sum_{i=1}^d b_i(t,x)\frac{\partial}{\partial x_i}u(t,x) + c(t,x)u(t,x), \\ u(0,x)=f(x). \end{cases} \tag{1-2}$$

Generally, (1-2) does not have a unique solution. We will assume the continuity of $a$ in spatial components uniformly in $t$, and this implies the uniqueness of the weak solution; see [Stroock and Varadhan 1979]. In the present paper, we always consider cases where the uniqueness of the weak solution holds. Set

$$L_t f(x) := \frac{1}{2}\sum_{i,j=1}^d a_{ij}(t,x)\frac{\partial^2}{\partial x_i\partial x_j}f(x) + \sum_{i=1}^d b_i(t,x)\frac{\partial}{\partial x_i}f(x) + c(t,x)f(x), \qquad f\in C_b^2(\mathbb{R}^d)$$

and denote the fundamental solution to (1-2) by $p(s,x;t,y)$, i.e., $p(s,x;t,y)$ is a measurable function defined for $s,t\in[0,\infty)$ such that $s<t$ and $x,y\in\mathbb{R}^d$ which satisfies

$$\frac{\partial}{\partial t}\int_{\mathbb{R}^d} f(y)p(s,\,\cdot\,;t,y)\,dy = L_t\left(\int_{\mathbb{R}^d} f(y)p(s,\,\cdot\,;t,y)\,dy\right) \quad\text{and}\quad \lim_{r\downarrow s}\int_{\mathbb{R}^d} f(y)p(s,\,\cdot\,;r,y)\,dy = f$$

*MSC2010:* primary 35B65, 35K10, 60H30; secondary 60H10, 60J60.
*Keywords:* parabolic partial differential equation, diffusion, fundamental solution, Hölder continuity, Gaussian estimate, stochastic differential equation, coupling method.

for $s, t \in [0, \infty)$ such that $s < t$ and a continuous function $f$ with a compact support. In the present paper, we consider the existence and the regularity of $p(0, x; t, y)$.

The problem of regularity of the fundamental solutions to parabolic partial differential equations with bounded measurable coefficients has a long history. Parabolic equations in divergence form has been investigated more thoroughly than that those in nondivergence form, because the variational method is applicable to them. The Hölder continuity of the fundamental solution to $\partial_t u = \nabla \cdot a \nabla u$ for a matrix-valued bounded measurable function $a$ with ellipticity condition $\Lambda^{-1} I \leq a \leq \Lambda I$ was originally obtained by De Giorgi [1957] and Nash [1958] independently. Precisely speaking, in their results the $\alpha$-Hölder continuity of the fundamental solution, with some positive number $\alpha \in (0, 1]$, is obtained. The index $\alpha$ depends on many constants appearing in the Harnack inequality and so on. These results have been extended to the case of the more general equations $\partial_t u = \nabla \cdot a \nabla u + b \cdot \nabla u - cu$, where $b, c$ are bounded measurable; see [Aronson 1967; Stroock 1988]. The case for unbounded coefficients is also studied; see, for example, [Metafune et al. 2009; Porper and Eidelman 1984; 1992]. An analogy to the case of a type of nonlocal generators (the associated stochastic processes are called stable-like processes) is given by Chen and Kumagai [2003]. In the results above, the index of the Hölder continuity of the fundamental solution depends on many constants appearing in the estimates, and it is difficult to calculate its exact value. Moreover, it is difficult to obtain even a lower bound for the index.

The fundamental solutions to parabolic equations in nondivergence form with low-regular coefficients have been studied mainly in the case of Hölder-continuous coefficients. One of the most powerful tools for the problem is the parametrix method, and it yields the existence, uniqueness and Hölder continuity of the fundamental solution; see [Friedman 1964; Ladyženskaja et al. 1967; Porper and Eidelman 1984, Chapter I]. Furthermore, an a priori estimate (the so-called Schauder estimate) is known for the solutions, and twice-continuous differentiability in $x$ of the fundamental solution $p(s, x; t, y)$ to (1-2) is obtained; see, for example, [Ladyženskaja et al. 1967; Krylov 1996; Bogachev et al. 2009; Bogachev et al. 2005]. We remark that all the coefficients $a, b, c$ need to be Hölder-continuous to apply the Schauder estimate. Even in the case that $a$ is the unit matrix and $b$ is not continuous, we cannot expect the continuous differentiability of the fundamental solution; see [Karatzas and Shreve 1991, Remark 5.2, Chapter 6].

In the present paper, we consider the Gaussian estimate and the lower bound of the index for the Hölder continuity in $x$ of the fundamental solution $p(s, x; t, y)$ to (1-2) by a probabilistic approach.

Now we fix some assumptions. Let $B(x, R)$ be the open ball in $\mathbb{R}^d$ centered at $x$ with radius $R$ for $x \in \mathbb{R}^d$ and $R > 0$. We assume that

$$\sum_{i,j=1}^{d} \sup_{t \in [0,\infty)} \int_{\mathbb{R}^d} \left| \frac{\partial}{\partial x_j} a_{ij}(s, x) \right|^{\theta} e^{-m|x|} \, dx \leq M, \tag{1-3}$$

where the derivatives are in the weak sense, $\theta$ is a constant in $[d, \infty) \cap (2, \infty)$, and $m$ and $M$ are nonnegative constants. We also assume the continuity of $a$ in spatial component uniformly in $t$, i.e., for any $R > 0$ there exists a continuous and nondecreasing function $\rho_R$ on $[0, \infty)$ such that $\rho_R(0) = 0$ and

$$\sup_{t \in [0,\infty)} \sup_{i,j} |a_{ij}(t, x) - a_{ij}(t, y)| \leq \rho_R(|x - y|), \quad x, y \in B(0; R). \tag{1-4}$$

We remark that under the assumptions (1-1) and (1-4), the equation (1-2) under consideration is well-posed (see [Stroock and Varadhan 1979, Chapter 7]), and for fixed $s \in [0, \infty)$ the fundamental solution $p(s, \cdot\,; t, \cdot\,)$ exists for almost all $t \in (s, \infty)$ (see [Stroock and Varadhan 1979, Theorem 9.1.9]). However, the fundamental solution does not always exist for all $t \in (s, \infty)$ under assumptions (1-1) and (1-4); see [Fabes and Kenig 1981]. We remark that under assumptions (1-1), (1-3) and (1-4), neither existence of the fundamental solutions nor examples where the fundamental solution does not exist are known. In the case that $a$ does not depend on the time $t$, the fundamental solution exists for all $t$; see [Stroock and Varadhan 1979, Theorem 9.2.6]. We also remark that (1-3) and (1-4) do not imply the local Hölder continuity of $a$ in the spatial component.

Let $p^X(s, x; t, y)$ be the fundamental solution to the parabolic equation

$$\frac{\partial}{\partial t}u(t, x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t, x)\frac{\partial^2}{\partial x_i \partial x_j}u(t, x),$$

and let

$$L_t^X = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t, x)\frac{\partial^2}{\partial x_i \partial x_j}. \tag{1-5}$$

Suppose that $\{a^{(n)}(t, x)\}$ is a sequence of symmetric $d \times d$-matrix-valued functions with components in $C_b^\infty([0, \infty) \times \mathbb{R}^d)$ such that $a^{(n)}(t, x)$ converges to $a(t, x)$ for each $(t, x) \in [0, \infty) \times \mathbb{R}^d$. We also assume that (1-1), (1-3) and (1-4) hold for $a^{(n)}$ instead of $a$, with the same constants $m$, $M$, $\theta$, $R$ and $\Lambda$, and the same function $\rho_R$. Denote the fundamental solution to the parabolic equation associated with the generator

$$\frac{1}{2}\sum_{i,j=1}^{d} a_{ij}^{(n)}(t, x)\frac{\partial^2}{\partial x_i \partial x_j}$$

by $p^{X,(n)}$. We assume the uniform Gaussian estimates for the fundamental solutions to $p^{X,(n)}$, i.e., there exist positive constants $\gamma_G^-$, $\gamma_G^+$, $C_G^-$ and $C_G^+$ such that

$$\frac{C_G^-}{(t-s)^{d/2}}\exp\biggl(-\frac{\gamma_G^-|x-y|^2}{t-s}\biggr) \le p^{X,(n)}(s, x; t, y) \le \frac{C_G^+}{(t-s)^{d/2}}\exp\biggl(-\frac{\gamma_G^+|x-y|^2}{t-s}\biggr) \tag{1-6}$$

for $s, t \in [0, \infty)$ such that $s < t$, $x, y \in \mathbb{R}^d$, and $n \in \mathbb{N}$. The Gaussian estimates for the fundamental solutions to parabolic equations in divergence form have been well investigated; see [Aronson 1967; Karrmann 2001; Porper and Eidelman 1984; 1992]. However, there are not many known results in the case of nondivergence form. A sufficient condition for the Gaussian estimate is obtained in [Porper and Eidelman 1992, Theorem 19] by means of Dini's continuity condition. The result includes the case of Hölder-continuous coefficients. We remark that two-sided estimates similar to the Gaussian estimates for the equations with general coefficients are obtained in [Escauriaza 2000].

Now we state the main theorem of this paper:

**Theorem 1.1.** *Assume* (1-1), (1-3), (1-4) *and* (1-6). *Then, there exist constants* $C_1$, $C_2$, $\gamma_1$ *and* $\gamma_2$ *depending on* $d$, $\gamma_{\mathrm{G}}^-$, $\gamma_{\mathrm{G}}^+$, $C_{\mathrm{G}}^-$, $C_{\mathrm{G}}^+$, $m$, $M$, $\theta$, $\Lambda$, $\|b\|_\infty$ *and* $\|c\|_\infty$ *such that*

$$\frac{C_1 e^{-C_1(t-s)}}{(t-s)^{d/2}} \exp\biggl(-\frac{\gamma_1|x-y|^2}{t-s}\biggr) \le p(s,x;t,y) \le \frac{C_2 e^{C_2(t-s)}}{(t-s)^{d/2}} \exp\biggl(-\frac{\gamma_2|x-y|^2}{t-s}\biggr)$$

*for* $s, t \in [0,\infty)$ *such that* $s < t$ *and* $x, y \in \mathbb{R}^d$. *Moreover, for any* $R > 0$ *and sufficiently small* $\varepsilon > 0$, *there exists a constant* $C$ *depending on* $d$, $\varepsilon$, $\gamma_{\mathrm{G}}^-$, $\gamma_{\mathrm{G}}^+$, $C_{\mathrm{G}}^-$, $C_{\mathrm{G}}^+$, $m$, $M$, $\theta$, $R$, $\rho_R$, $\Lambda$, $\|b\|_\infty$ *and* $\|c\|_\infty$ *such that*

$$|p(0,x;t,y) - p(0,z;t,y)| \le C t^{-d/2-1} e^{Ct} |x-z|^{1-\varepsilon}$$

*for* $t \in (0,\infty)$, $x, z \in B(0; R/2)$ *and* $y \in \mathbb{R}^d$.

The first assertion of Theorem 1.1 is the Gaussian estimate for $p$. The advantage of the result is obtaining the Gaussian estimate of the fundamental solution to the parabolic equation in nondivergence form without the continuity of $b$ and $c$. Such a result seems difficult to obtain via the parametrix method. The second assertion of Theorem 1.1 implies that $p(0,x;t,y)$ is $(1-\varepsilon)$-Hölder continuous in $x$, and this is a clear lower bound for the continuity. The approach in this paper is mainly probabilistic. The key method to prove Theorem 1.1 is the coupling method introduced by Lindvall and Rogers [1986]. This method enables us to discuss the Hölder continuity of $p(0,x;t,y)$ in $x$ from the oscillation of the diffusion processes without regularity of the coefficients.

If $a$ is uniformly continuous in the spatial component, our proof below follows without restriction on $x, z$, and the following corollary holds:

**Corollary 1.2.** *Assume* (1-1), (1-3), (1-6) *and that there exists a continuous and nondecreasing function* $\rho$ *on* $[0,\infty)$ *such that* $\rho(0) = 0$ *and*

$$\sup_{t\in[0,\infty)} \sup_{i,j} |a_{ij}(t,x) - a_{ij}(t,y)| \le \rho(|x-y|), \quad x, y \in \mathbb{R}^d.$$

*Then, for sufficiently small* $\varepsilon > 0$, *there exists a constant* $C$ *such that*

$$|p(0,x;t,y) - p(0,z;s,y)| \le C t^{-d/2-1} e^{Ct} |x-z|^{1-\varepsilon}$$

*for* $t \in (0,\infty)$ *and* $x, y, z \in \mathbb{R}^d$.

The assumption (1-6) may seem strict. However, as mentioned above, Theorem 19 of [Porper and Eidelman 1992] gives a Gaussian estimate for the parabolic equations with coefficients which satisfy a version of Dini's continuity condition. From this sufficient condition and Theorem 1.1, we have the following corollary:

**Corollary 1.3.** *Assume* (1-1), (1-3), *and that there exists a continuous and nondecreasing function* $\rho$ *on* $[0,\infty)$ *such that* $\rho(0) = 0$,

$$\int_0^1 \frac{1}{r_2}\biggl(\int_0^{r_2} \frac{\rho(r_1)}{r_1}\,dr_1\biggr) dr_2 < \infty \ \ \textit{and} \ \sup_{t\in[0,\infty)} \sup_{i,j} |a_{ij}(t,x) - a_{ij}(t,y)| \le \rho(|x-y|), \ \ x, y \in \mathbb{R}^d. \tag{1-7}$$

*Then, for sufficiently small* $\varepsilon > 0$, *there exists a constant* $C$ *such that*

$$|p(0, x; t, y) - p(0, z; s, y)| \leq Ct^{-d/2-1}e^{Ct}|x - z|^{1-\varepsilon}$$

*for* $t \in (0, \infty)$ *and* $x, y, z \in \mathbb{R}^d$.

We remark that, for $\alpha \in (0, 1]$ and a positive constant $C$, $\rho(r) = Cr^{\alpha}$ satisfies (1-7). Furthermore, $\rho(r) = C \min\{1, (-\log r)^{-\alpha}\}$ satisfies (1-7) for $\alpha \in (2, \infty)$. We also remark that continuity of $b$ and $c$ are not assumed in Corollary 1.3.

The organization of the paper is as follows:

In Section 2, we prepare the probabilistic representation of the fundamental solution to (1-2). It should be remarked that we consider the case where $a$ is smooth in Sections 2–4, and the general case is considered only in Section 5. The representation enable us to consider the Hölder continuity of the fundamental solution by a probabilistic way, and actually in Section 4 we prove the constant appearing in the Hölder continuity of $p(0, x; t, y)$ in $x$ depends only on the suitable constants. The representation is obtained by the Feynman–Kac formula and the Girsanov transformation, and in the end of this section $p(s, x; t, y)$ is represented by the functional of the pinned diffusion process.

In Section 3, we prepare some estimates. The goal of this section is Lemma 3.5, which concerns the integrability of a functional of the pinned diffusion process. Generally speaking, it is much harder to see the integrability with respect to conditional probability measures than with respect to the original probability measure. In our case, conditioning generates a singularity, and this fact makes the estimate difficult. To overcome the difficulty, we begin with Lemma 3.1, which is an estimate of the derivative of $p(s, x; t, y)$. The proof of this lemma is analytic, and (1-3) is assumed for the lemma. In this section, we also have the Gaussian estimate for $p(s, x; t, y)$.

In Section 4, we prove that the constant appearing in the $(1 - \varepsilon)$-Hölder continuity of $p(0, x; t, y)$ in $x$ depends only on the suitable constants. This section is the main part of our argument. To show this, we apply the coupling method to diffusion processes. By virtue of the coupling method, the continuity problem of the fundamental solution is reduced to the problem of the local behavior of the pinned diffusion processes. To see the local behavior, (1-4) is needed. Finally, by showing an estimate of the coupling time, we obtain the $(1 - \varepsilon)$-Hölder continuity of $p(0, x; t, y)$ in $x$ and the suitable dependence of the constant appearing in the Hölder continuity.

In Section 5, we consider the case of general $a$ and prove Theorem 1.1. Our approach is just smoothly approximating $a$ and using the result obtained in Section 4.

Throughout this paper, we denote the inner product in the Euclidean space $\mathbb{R}^d$ by $\langle \cdot , \cdot \rangle$, and all random variables are considered on a probability space $(\Omega, \mathscr{F}, P)$. We denote the expectation of random variables by $E[\,\cdot\,]$ and the expectation on the event $A \in \mathscr{F}$ (i.e., $\int_A \cdot \, dP$) by $E[\,\cdot\,; A]$. We denote the smooth functions with bounded derivatives on $S$ by $C_b^{\infty}(S)$ and the smooth functions on $S$ with compact support by $C_0^{\infty}(S)$.

## 2. Probabilistic representation of the fundamental solution

In this section, we assume that $a_{ij}(t, x) \in C_b^{0,\infty}([0, \infty) \times \mathbb{R}^d)$. Define a $d \times d$-matrix-valued function $\sigma(t, x)$ as the square root of $a(t, x)$. Then, (1-1) implies that $\sigma_{ij}(t, x) \in C_b^{0,\infty}([0, \infty) \times \mathbb{R}^d)$, $a(t, x) =$

$\sigma(t,x)\sigma(t,x)^T$ and

$$\sup_{t\in[0,\infty)}\sup_{i,j}|\sigma_{ij}(t,x)-\sigma_{ij}(t,y)|\le C\rho_R(|x-y|),\quad x,y\in B(0;R),\tag{2-1}$$

where $C$ is a constant depending on $\Lambda$. Note that (1-1) implies that

$$\Lambda^{-1/2}I\le\sigma(t,x)\le\Lambda^{1/2}I.\tag{2-2}$$

Consider the stochastic differential equation:

$$\begin{cases}dX_t^x=\sigma(t,X_t^x)\,dB_t,\\ X_0^x=x.\end{cases}\tag{2-3}$$

Lipschitz continuity of $\sigma$ implies the existence of a solution and its pathwise uniqueness. Let $(\mathscr{F}_t)$ be the $\sigma$-field generated by $(B_s : s\in[0,t])$. Then, pathwise uniqueness implies that the solution $X_t^x$ is $\mathscr{F}_t$-measurable. All stopping times appearing in this paper are associated with $(\mathscr{F}_t)$. We remark that the generator of $(X_t^x)$ is given by (1-5), and therefore the transition probability density of $(X_t^x)$ coincides with the fundamental solution $p^X$ of the parabolic equation generated by $(L_t^X)$. The smoothness of $\sigma$ implies the smoothness of $p^X(s,x;t,y)$ on $(0,\infty)\times\mathbb{R}^d\times(0,\infty)\times\mathbb{R}^d$; see, for example, [Kusuoka and Stroock 1985] for the probabilistic proof and [Lax and Milgram 1954] for the analytic proof.

There is a relation between the fundamental solution and the generator, as follows. Since $p^X$ is smooth, by the definition of $p^X$ we have

$$\frac{\partial}{\partial t}p^X(s,x;t,y)=[L_t^X p^X(s,\cdot\,;t,y)](x)\tag{2-4}$$

for $s,t\in[0,\infty)$ such that $s<t$ and $x,y\in\mathbb{R}^d$. Let $(L_t^X)^*$ be the dual operator of $L_t^X$ on $L^2(\mathbb{R}^d)$. Define $T_{s,t}^X$ and $(T_{s,t}^X)^*$ as the semigroups generated by $L_t^X$ and $(L_t^X)^*$, respectively. Since

$$\int_{\mathbb{R}^d}\phi(x)(T_{s,t}^X\psi)(x)\,dx=\int_{\mathbb{R}^d}\psi(x)[(T_{s,t}^X)^*\phi](x)\,dx,$$

we have

$$\int_{\mathbb{R}^d}\phi(x)\biggl(\int_{\mathbb{R}^d}\psi(y)p^X(s,x;t,y)\,dy\biggr)dx=\int_{\mathbb{R}^d}\psi(x)\biggl(\int_{\mathbb{R}^d}\phi(y)(p^X)^*(s,x;t,y)\,dy\biggr)dx,$$

where $(p^X)^*(s,x;t,y)$ is the fundamental solution associated with $(L_t^X)^*$. Hence, it holds that

$$p^X(s,x;t,y)=(p^X)^*(s,y;t,x)$$

for $s,t\in(0,\infty)$ such that $s<t$ and $x,y\in\mathbb{R}^d$. Differentiating both sides of this equation with respect to $t$, we obtain

$$[L_t^X p^X(s,\cdot\,;t,y)](x)=[(L_t^X)^*(p^X)^*(s,\cdot\,;t,x)](y)=[(L_t^X)^*p^X(s,x;t,\cdot)](y)\tag{2-5}$$

for $s, t \in [0, \infty)$ such that $s < t$, and $x, y \in \mathbb{R}^d$. By the Chapman–Kolmogorov equation, we have, for $s, t, u \in [0, \infty)$ such that $u < s < t$ and $x, y \in \mathbb{R}^d$, that

$$p^X(u, x; t, y) = \int_{\mathbb{R}^d} p^X(u, x; s, \xi) p^X(s, \xi; t, y)\, d\xi.$$

Differentiating both sides of this equation with respect to $s$, we have

$$0 = \int_{\mathbb{R}^d} \left( \frac{\partial}{\partial s} p^X(u, x; s, \xi) \right) p^X(s, \xi; t, y)\, d\xi + \int_{\mathbb{R}^d} p^X(u, x; s, \xi) \left( \frac{\partial}{\partial s} p^X(s, \xi; t, y) \right) d\xi$$

for $s, u \in [0, \infty)$ such that $u < s$ and $x, y \in \mathbb{R}^d$. Since (2-4) and (2-5) imply that

$$\frac{\partial}{\partial s} p^X(u, x; s, \xi) = [L_s^X p^X(u, \cdot\, ; s, \xi)](x) = [(L_s^X)^* p^X(u, x; s, \cdot)](\xi),$$

we have, for $s, t, u \in [0, \infty)$ such that $u < s < t$ and $x, y \in \mathbb{R}^d$, that

$$\begin{aligned} \int_{\mathbb{R}^d} p^X(u, x; s, \xi) \left( \frac{\partial}{\partial s} p^X(s, \xi; t, y) \right) d\xi &= - \int_{\mathbb{R}^d} [(L_s^X)^* p^X(u, x; s, \cdot)](\xi) p^X(s, \xi; t, y)\, d\xi \\ &= - \int_{\mathbb{R}^d} p^X(u, x; s, \xi) [L_s^X p^X(s, \cdot\, ; t, y)](\xi)\, d\xi. \end{aligned}$$

Noting that $p^X(u, x; s, \xi)$ converges to $\delta_x(\xi)$ as $u \uparrow s$ in the sense of Schwartz distributions, we obtain

$$\frac{\partial}{\partial s} p^X(s, x; t, y) = -[L_s^X p^X(s, \cdot\, ; t, y)](x) \tag{2-6}$$

for $s, t \in (0, \infty)$ such that $s < t$ and $x, y \in \mathbb{R}^d$.

Next we study the probabilistic representation of $p(s, x; t, y)$ by $p^X(s, x; t, y)$. By the Feynman–Kac formula (see, for example, [Revuz and Yor 1999, Proposition 3.10, Chapter VIII]) and the Girsanov transformation (see, for example, [Ikeda and Watanabe 1989, Theorem 4.2, Chapter IV]), we have the following representation of $u(t, x)$ by $X_t^x$:

$$u(t, x) = E\left[ f(X_t^x) \exp\left( \int_0^t \langle b_\sigma(s, X_s^x), dB_s \rangle - \frac{1}{2} \int_0^t |b_\sigma(s, X_s^x)|^2\, ds + \int_0^t c(s, X_s^x)\, ds \right) \right], \tag{2-7}$$

where $b_\sigma(t, x) := \sigma(t, x)^{-1} b(t, x)$. For $s \le t$ and $x \in \mathbb{R}^d$, let

$$\mathscr{E}(s, t; X^x) := \exp\left( \int_s^t \langle b_\sigma(u, X_u^x), dB_u \rangle - \frac{1}{2} \int_s^t |b_\sigma(u, X_u^x)|^2\, du + \int_s^t c(u, X_u^x)\, du \right).$$

Then, by the definition of the fundamental solution and (2-7), we obtain the probabilistic representation of the fundamental solution:

$$p(0, x; t, y) = p^X(0, x; t, y) E^{X_t^x = y}[\mathscr{E}(0, t; X^x)], \tag{2-8}$$

where $P^{X_t^x=y}$ is the conditional probability measure of $P$ on $X_t^x = y$ and $E^{X_t^x=y}[\,\cdot\,]$ is the expectation with respect to $P^{X_t^x=y}$. Hence, to see the regularity of $p(0,x;t,y)$ in $x$, it is sufficient to see the regularity of the function $x \mapsto p^X(0,x;t,y)E^{X_t^x=y}[\mathscr{E}(0,t;X^x)]$. We prove Theorem 1.1 by studying the regularity of this function. The definition of $\mathscr{E}$ implies that

$$\mathscr{E}(0,t;X^x) - \mathscr{E}(\tau\wedge t,t;X^x) = \mathscr{E}(\tau\wedge t,t;X^x)(\mathscr{E}(0,\tau\wedge t;X^x)-1) \tag{2-9}$$

for any stopping time $\tau$ and $t\in[0,\infty)$, and by Itō's formula we have

$$\mathscr{E}(s,t;X^x) - 1 = \int_s^t \mathscr{E}(s,u;X^x)\langle b_\sigma(u,X_u^x), dB_u\rangle + \int_s^t \mathscr{E}(s,u;X^x)c(u,X_u^x)\,du \tag{2-10}$$

for $s,t\in[0,\infty)$ such that $s\le t$. We use these equations in the proof.

Now we consider the diffusion process $X^x$ pinned at $y$ at time $t$. Let $s,t\in[0,\infty)$ such that $s<t$, $x,y\in\mathbb{R}^d$ and $\varepsilon>0$. By the Markov property of $X$, we have for $A\in\mathscr{F}_s$ that

$$P(A\cap\{X_t^x\in B(y;\varepsilon)\}_{B(y;\varepsilon)}\biggl(\int_{\mathbb{R}^d} p^X(s,\xi;t,\xi')P(A\cap\{X_s^x\in d\xi\})\biggr)\,d\xi'.$$

Hence, we obtain

$$P^{X_t^x=y}(A) = \frac{1}{p^X(0,x;t,y)}\int_{\mathbb{R}^d} p^X(s,\xi;t,y)P(A\cap\{X_s^x\in d\xi\}) \tag{2-11}$$

for $s,t\in(0,\infty)$ such that $s<t$, $A\in\mathscr{F}_s$ and $x,y\in\mathbb{R}^d$. This formula enables us to see the generator of the pinned diffusion process. By Itō's formula, (2-6) and (2-11) we have, for $f\in C_b^2(\mathbb{R}^d)$, $s,t\in[0,\infty)$ such that $s<t$ and $x,y\in\mathbb{R}^d$,

$$\begin{aligned}
&p^X(0,x;t,y)E^{X_t^x=y}[f(X_s^x)] - p^X(0,x;t,y)f(x)\\
&\quad= E[f(X_s^x)p^X(s,X_s^x;t,y)] - E[f(X_0^x)p^X(0,X_0^x;t,y)]\\
&\quad= E\biggl[\int_0^s (L_u^X f)(X_u^x)p^X(u,X_u^x;t,y)u\biggr] + E\biggl[\int_0^s f(X_u^x)\Bigl(\frac{\partial}{\partial u}p^X(u,\xi;t,y)\Bigr)\Big|_{\xi=X_u^x}du\biggr]\\
&\qquad\qquad + E\biggl[\int_0^s f(X_u^x)(L_u^X p^X(u,\cdot\,;t,y))(X_u^x)\,du\biggr]\\
&\qquad\qquad + \frac{1}{2}E\biggl[\int_0^s \langle\sigma(u,X_u^x)^T\nabla f(X_u^x), \sigma(u,X_u^x)^T(\nabla p^X(u,\cdot\,;t,y))(X_u^x)\rangle\,du\biggr]\\
&\quad= p^X(0,x;t,y)\int_0^s E^{X_t^x=y}[(L_u^X f)(X_u^x)]\,du\\
&\qquad\qquad + \frac{1}{2}p^X(0,x;t,y)\int_0^s E^{X_t^x=y}\biggl[\biggl\langle\nabla f(X_u^x), a(u,X_u^x)\frac{(\nabla p^X(u,\cdot\,;t,y))(X_u^x)}{p^X(u,X_u^x;t,y)}\biggr\rangle\biggr]\,du.
\end{aligned}$$

Hence, the generator of $X$ pinned at $y$ at time $t$ is

$$\frac{1}{2}\sum_{i,j=1}^d a_{ij}(s,x)\frac{\partial^2}{\partial x_i\partial x_j} + \biggl\langle\frac{1}{2}a(s,x)\frac{\nabla p^X(s,\cdot\,;t,y)(x)}{p^X(s,x;t,y)}, \nabla\biggr\rangle$$

for $s \in [0, t)$ and $x \in \mathbb{R}^d$. Of course, pinned Brownian motion is an example of pinned diffusion processes; see [Ikeda and Watanabe 1989, Example 8.5, Chapter IV].

## 3. Estimates

In this section we prepare some estimates for the proof of the main theorem. Assume that $a$ is smooth and fix notation as in Section 2.

**Lemma 3.1.** *Let $t \in (0, \infty)$ and $\phi$ be a nonnegative continuous function on $(0, t) \times \mathbb{R}^d$ such that $\phi(\,\cdot\,, x) \in W^{1,1}_{\mathrm{loc}}((0, t), ds)$ for $x \in \mathbb{R}^d$ and $\phi(s, \cdot\,) \in W^{1,2}_{\mathrm{loc}}(\mathbb{R}^d, dx)$ for $s \in (0, t)$. Then, for $s_1, s_2 \in (0, t)$ such that $s_1 \le s_2$,*

$$\begin{aligned}
&\int_{s_1}^{s_2}\int_{\mathbb{R}^d} \frac{\langle a(u,\xi)\nabla_\xi p^X(u,\xi;t,y), \nabla_\xi p^X(u,\xi;t,y)\rangle}{p^X(u,\xi;t,y)^2}\phi(u,\xi)\,d\xi\,du \\
&\le C(1+|\log(t-s_1)|)\int_{\mathbb{R}^d}\phi(s_1,\xi)\,d\xi + C(t-s_1)^{-1}\int_{\mathbb{R}^d}|y-\xi|^2\phi(s_1,\xi)\,d\xi \\
&\quad + C(1+|\log(t-s_2)|)\int_{\mathbb{R}^d}\phi(s_2,\xi)\,d\xi + C(t-s_2)^{-1}\int_{\mathbb{R}^d}|y-\xi|^2\phi(s_2,\xi)\,d\xi + C\int_{s_1}^{s_2}\int_{\mathbb{R}^d}\phi(u,\xi)\,d\xi\,du \\
&\quad + C\int_{s_1}^{s_2}\int_{\operatorname{supp}\phi}\frac{|\nabla_\xi\phi(u,\xi)|^2}{\phi(u,\xi)}\,d\xi\,du + C\int_{s_1}^{s_2}(1+|\log(t-u)|)\int_{\mathbb{R}^d}\left|\frac{\partial}{\partial u}\phi(u,\xi)\right|d\xi\,du \\
&\quad + C\int_{s_1}^{s_2}(t-u)^{-1}\int_{\mathbb{R}^d}|y-\xi|^2\left|\frac{\partial}{\partial u}\phi(u,\xi)\right|d\xi\,du + C\sum_{i,j=1}^{d}\int_0^t\int_{\mathbb{R}^d}\left|\frac{\partial}{\partial\xi_j}a_{ij}(u,\xi)\right|^2\phi(u,\xi)d\xi\,du,
\end{aligned}$$

*where $C$ is a constant depending on $d$, $\gamma_G^-$, $\gamma_G^+$, $C_G^-$, $C_G^+$, and $\Lambda$, and $\operatorname{supp}\phi$ is the support of $\phi$.*

**Remark 3.2.** If $\phi$ is a continuous function on $\mathbb{R}^d$, the Lebesgue measure of $\operatorname{supp}\phi \setminus \{x \in \mathbb{R}^d : \phi(x) > 0\}$ is zero.

*Proof of Lemma 3.1.* It is sufficient to show the theorem for $\phi \in C_0^\infty([0, t] \times \mathbb{R}^d)$, because the general case is obtained by approximation. Let $u \in (0, t)$. Recall that the components of the coefficient $\sigma$ of (2-3) are in $C_b^{0,\infty}([0, \infty) \times \mathbb{R}^d)$ and $\sigma$ is uniformly positive-definite. Hence, the associated transition probability density $p^X(\,\cdot\,, \cdot\,; t, \cdot\,)$ is smooth on $(0, t) \times \mathbb{R}^d \times \mathbb{R}^d$, and $p^X(s, x; t, y) > 0$ for $s \in (0, t)$ and $x, y \in \mathbb{R}^d$; see, for example, [Aronson 1967]. By the Leibniz rule, we have

$$\begin{aligned}
&\frac{\partial}{\partial u}\int_{\mathbb{R}^d}(\log p^X(u,\xi;t,y))\phi(u,\xi)\,d\xi \\
&\qquad = \int_{\mathbb{R}^d}\frac{(\partial/\partial u)p^X(u,\xi;t,y)}{p^X(u,\xi;t,y)}\phi(u,\xi)\,d\xi + \int_{\mathbb{R}^d}(\log p^X(u,\xi;t,y))\frac{\partial}{\partial u}\phi(u,\xi)\,d\xi. \qquad (3\text{-}1)
\end{aligned}$$

The equality (2-6) and integration by parts imply

$$\begin{aligned}
\int_{\mathbb{R}^d} &\frac{(\partial/\partial u)p^X(u,\xi;t,y)}{p^X(u,\xi;t,y)}\phi(u,\xi)\,d\xi \\
&= -\frac{1}{2}\sum_{i,j=1}^d \int_{\mathbb{R}^d} \frac{a_{ij}(u,\xi)(\partial^2/\partial\xi_i\partial\xi_j)p^X(u,\xi;t,y)}{p^X(u,\xi;t,y)}\phi(u,\xi)\,d\xi \\
&= -\frac{1}{2}\sum_{i,j=1}^d \int_{\mathbb{R}^d} \frac{(\partial/\partial\xi_i)(a_{ij}(u,\xi)(\partial/\partial\xi_j)p^X(u,\xi;t,y))}{p^X(u,\xi;t,y)}\phi(u,\xi)\,d\xi \\
&\qquad + \frac{1}{2}\sum_{i,j=1}^d \int_{\mathbb{R}^d} \frac{((\partial/\partial\xi_i)a_{ij}(u,\xi))(\partial/\partial\xi_j)p^X(u,\xi;t,y)}{p^X(u,\xi;t,y)}\phi(u,\xi)\,d\xi \\
&= -\frac{1}{2}\int_{\mathbb{R}^d} \frac{\langle a(u,\xi)\nabla_\xi p^X(u,\xi;t,y), \nabla_\xi p^X(u,\xi;t,y)\rangle}{p^X(u,\xi;t,y)^2}\phi(u,\xi)\,d\xi \\
&\qquad + \frac{1}{2}\int_{\mathbb{R}^d} \left\langle \frac{a(u,\xi)\nabla_\xi p^X(u,\xi;t,y)}{p^X(u,\xi;t,y)}, \nabla_\xi\phi(u,\xi)\right\rangle d\xi \\
&\qquad + \frac{1}{2}\sum_{i,j=1}^d \int_{\mathbb{R}^d} \frac{((\partial/\partial\xi_i)a_{ij}(u,\xi))(\partial/\partial\xi_j)p^X(u,\xi;t,y)}{p^X(u,\xi;t,y)}\phi(u,\xi)\,d\xi.
\end{aligned}$$

Hence, by (3-1) we have

$$\begin{aligned}
\frac{\partial}{\partial u}\int_{\mathbb{R}^d}(\log p^X(u,\xi;t,y))\phi(u,\xi)\,d\xi &= -\frac{1}{2}\int_{\mathbb{R}^d} \frac{\langle a(u,\xi)\nabla_\xi p^X(u,\xi;t,y), \nabla_\xi p^X(u,\xi;t,y)\rangle}{p^X(u,\xi;t,y)^2}\phi(u,\xi)\,d\xi \\
&\quad + \frac{1}{2}\int_{\mathbb{R}^d} \left\langle \frac{a(u,\xi)\nabla_\xi p^X(u,\xi;t,y)}{p^X(u,\xi;t,y)}, \nabla_\xi\phi(u,\xi)\right\rangle d\xi \\
&\quad + \frac{1}{2}\sum_{i,j=1}^d \int_{\mathbb{R}^d} \frac{((\partial/\partial\xi_i)a_{ij}(u,\xi))(\partial/\partial\xi_j)p^X(u,\xi;t,y)}{p^X(u,\xi;t,y)}\phi(u,\xi)\,d\xi \\
&\quad + \int_{\mathbb{R}^d}(\log p^X(u,\xi;t,y))\frac{\partial}{\partial u}\phi(u,\xi)\,d\xi.
\end{aligned}$$

Integrating both sides from $s_1$ to $s_2$ with respect to $u$, we obtain

$$\begin{aligned}
\frac{1}{2}\int_{s_1}^{s_2}\int_{\mathbb{R}^d} &\frac{\langle a(u,\xi)\nabla_\xi p^X(u,\xi;t,y), \nabla_\xi p^X(u,\xi;t,y)\rangle}{p^X(u,\xi;t,y)^2}\phi(u,\xi)\,d\xi\,du \\
&= \int_{\mathbb{R}^d}(\log p^X(s_1,\xi;t,y))\phi(s_1,\xi)\,d\xi - \int_{\mathbb{R}^d}(\log p^X(s_2,\xi;t,y))\phi(s_2,\xi)\,d\xi \\
&\quad + \frac{1}{2}\int_{s_1}^{s_2}\int_{\mathbb{R}^d} \left\langle \frac{a(u,\xi)\nabla_\xi p^X(u,\xi;t,y)}{p^X(u,\xi;t,y)}, \nabla_\xi\phi(u,\xi)\right\rangle d\xi\,du \\
&\quad + \frac{1}{2}\sum_{i,j=1}^d \int_{s_1}^{s_2}\int_{\mathbb{R}^d} \frac{((\partial/\partial\xi_i)a_{ij}(u,\xi))(\partial/\partial\xi_j)p^X(u,\xi;t,y)}{p^X(u,\xi;t,y)}\phi(u,\xi)\,d\xi\,du \\
&\quad + \int_{s_1}^{s_2}\int_{\mathbb{R}^d}(\log p^X(u,\xi;t,y))\frac{\partial}{\partial u}\phi(u,\xi)\,d\xi\,du. \qquad (3\text{-}2)
\end{aligned}$$

Now we consider the estimates for the terms on the right-hand side of this equation. By (1-6), we have for $s \in (0, t)$ that

$$\left|\int_{\mathbb{R}^d} (\log p^X(s, \xi; t, y))\phi(s, \xi)\, d\xi\right| \le \int_{\mathbb{R}^d} \left( |\log C_{\mathrm{G}}^{+}| + |\log C_{\mathrm{G}}^{-}| + \frac{d}{2} |\log(t-s)| + \frac{\gamma_{\mathrm{G}}^{-} |y-\xi|^2}{t-s} \right) \phi(s, \xi)\, d\xi.$$

Hence, there exists a constant $C$ depending on $d$, $\gamma_{\mathrm{G}}^{-}$, $\gamma_{\mathrm{G}}^{+}$, $C_{\mathrm{G}}^{-}$, $C_{\mathrm{G}}^{+}$, and $\Lambda$ such that, for $s \in (0, t)$,

$$\begin{aligned}\left|\int_{\mathbb{R}^d} (\log p^X(s, \xi; t, y))\phi(s, \xi)\, d\xi\right| \\ \le C(1 + |\log(t-s)|) \int_{\mathbb{R}^d} \phi(s, \xi) d\xi + C(t-s)^{-1} \int_{\mathbb{R}^d} |y-\xi|^2 \phi(s, \xi)\, d\xi. \end{aligned} \tag{3-3}$$

The third term of the right-hand side of (3-2) can be estimated as follows:

$$\begin{aligned}\frac{1}{2}\left|\int_{s_1}^{s_2} \int_{\mathbb{R}^d} \left\langle \frac{a(u, \xi)\nabla_z p^X(u, \xi; t, y)}{p^X(u, \xi; t, y)}, \nabla_\xi \phi(u, \xi)\right\rangle d\xi\, du\right| \\ \le \frac{1}{8} \int_{s_1}^{s_2} \int_{\mathbb{R}^d} \frac{\langle a(u, \xi)\nabla_\xi p^X(u, \xi; t, y), \nabla_\xi p^X(u, \xi; t, y)\rangle}{p^X(u, \xi; t, y)^2} \phi(u, \xi)\, d\xi\, du \\ + 8 \int_{s_1}^{s_2} \int_{\operatorname{supp}\phi} \frac{\langle a(u, \xi)\nabla_\xi \phi(u, \xi), \nabla_\xi \phi(u, \xi)\rangle}{\phi(u, \xi)}\, d\xi\, du. \end{aligned} \tag{3-4}$$

To estimate the fourth term of the right-hand side of (3-2), we observe that

$$\begin{aligned}\frac{1}{2}\left|\sum_{i,j=1}^{d} \int_{s_1}^{s_2} \int_{\mathbb{R}^d} \frac{((\partial/\partial\xi_i) a_{ij}(u, \xi))(\partial/\partial\xi_j) p^X(u, \xi; t, y)}{p^X(u, \xi; t, y)} \phi(u, \xi)\, d\xi\, du\right| \\ \le \frac{1}{8\Lambda} \sum_{j=1}^{d} \int_{s_1}^{s_2} \int_{\mathbb{R}^d} \frac{|(\partial/\partial\xi_j) p^X(u, \xi; t, y)|^2}{p^X(u, \xi; t, y)^2} \phi(u, \xi)\, d\xi\, du \\ + 8d\Lambda \sum_{i,j=1}^{d} \int_{s_1}^{s_2} \int_{\mathbb{R}^d} |\frac{\partial}{\partial\xi_j} a_{ij}(u, \xi)|^2 \phi(u, \xi)\, d\xi\, du. \end{aligned}$$

Hence, by (1-1) we have

$$\begin{aligned}\frac{1}{2}\left|\sum_{i,j=1}^{d} \int_{s_1}^{s_2} \int_{\mathbb{R}^d} \frac{(\frac{\partial}{\partial\xi_i} a_{ij}(t, \xi)) \frac{\partial}{\partial\xi_j} p^X(u, \xi; t, y)}{p^X(u, \xi; t, y)} \phi(u, \xi)\, d\xi\, du\right| \\ \le \frac{1}{8} \int_{s_1}^{s_2} \int_{\mathbb{R}^d} \frac{\langle a(u, \xi)\nabla_\xi p^X(u, \xi; t, y), \nabla_\xi p^X(u, \xi; t, y)\rangle}{p^X(u, \xi; t, y)^2} \phi(u, \xi)\, d\xi\, du \\ + C \sum_{i,j=1}^{d} \int_{s_1}^{s_2} \int_{\mathbb{R}^d} \left|\frac{\partial}{\partial\xi_j} a_{ij}(u, \xi)\right|^2 \phi(u, \xi)\, d\xi\, du, \end{aligned} \tag{3-5}$$

where $C$ depends on $d$, $\gamma_{\mathrm{G}}^{-}$, $\gamma_{\mathrm{G}}^{+}$, $C_{\mathrm{G}}^{-}$, $C_{\mathrm{G}}^{+}$ and $\Lambda$. By using (1-6), we estimate the final term of the right-hand side of (3-2) as follows:

$$\left|\int_{s_1}^{s_2}\int_{\mathbb{R}^d}(\log p^X(u,\xi;t,y))\frac{\partial}{\partial u}\phi(u,\xi)\,d\xi\,du\right|$$
$$\leq \int_{s_1}^{s_2}\int_{\mathbb{R}^d}\left(|\log C_{\mathrm{G}}^{+}|+|\log C_{\mathrm{G}}^{-}|+\frac{d}{2}|\log(t-u)|+\frac{\gamma_{\mathrm{G}}^{-}|y-\xi|^2}{t-u}\right)\left|\frac{\partial}{\partial u}\phi(u,\xi)\right|d\xi\,du.$$

Hence, there exists a constant $C$ depending on $d$, $\gamma_{\mathrm{G}}^{-}$, $\gamma_{\mathrm{G}}^{+}$, $C_{\mathrm{G}}^{-}$, $C_{\mathrm{G}}^{+}$ and $\Lambda$ such that

$$\left|\int_{\mathbb{R}^d}(\log p^X(s,\xi;t,y))\phi(s_1,\xi)\,d\xi\right|$$
$$\leq C\int_{s_1}^{s_2}(1+|\log(t-u)|)\int_{\mathbb{R}^d}\left|\frac{\partial}{\partial u}\phi(u,\xi)\right|d\xi\,du+C\int_{s_1}^{s_2}(t-u)^{-1}\int_{\mathbb{R}^d}|y-\xi|^2\left|\frac{\partial}{\partial u}\phi(u,\xi)\right|d\xi\,du. \quad (3\text{-}6)$$

Therefore, by (3-2), (3-3), (3-4), (3-5) and (3-6) we obtain the lemma. □

Next, we state the fact on the integrability of $\mathscr{E}$ as a lemma. The proof is obtained by the standard argument; see, for example, [Stroock and Varadhan 1979, Theorem 4.2.1]. So, we omit it.

**Lemma 3.3.** *Let $\tau_1$, $\tau_2$ be stopping times such that $0 \leq \tau_1 \leq \tau_2$ almost surely. It holds that, for any $q \in \mathbb{R}$,*

$$E[\mathscr{E}(t\wedge\tau_1, t\wedge\tau_2; X^x)^q] \leq e^{C(1+q^2)t}, \quad t>0,\ x,y\in\mathbb{R}^d,$$

*where $C$ is a constant depending on $d$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$.*

**Lemma 3.4.** *Let $\tau_1$, $\tau_2$ be stopping times such that $0 \leq \tau_1 \leq \tau_2 \leq t$ almost surely. It holds that*

$$p^X(0,x;t,y)E^{X_t^x=y}\left[\int_0^t \mathscr{E}(u\wedge\tau_1, u\wedge\tau_2; X^x)^q\,du\right] \leq Ct^{-d/2+1-\varepsilon}e^{C(1+q^2)t}\exp\left(-\frac{\gamma|x-y|^2}{t}\right)$$

*for $t \in (0,\infty)$, $x, y \in \mathbb{R}^d$, $q \in \mathbb{R}$ and sufficiently small $\varepsilon > 0$, where $C$ and $\gamma$ are positive constants depending on $d$, $\varepsilon$, $\gamma_{\mathrm{G}}^{+}$, $C_{\mathrm{G}}^{+}$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$.*

*Proof.* In view of Fubini's theorem and (2-11), it is sufficient to show that there exist positive constants $C$ and $\gamma$, depending on $d$, $\varepsilon$, $\gamma_{\mathrm{G}}^{+}$, $C_{\mathrm{G}}^{+}$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$, such that

$$\int_0^t E[\mathscr{E}(u\wedge\tau_1, u\wedge\tau_2; X^x)^q p^X(u,X_u;t,y)]\,du \leq Ct^{-d/2+1-\varepsilon}e^{C(1+q^2)t}\exp\left(-\frac{\gamma|x-y|^2}{t}\right) \quad (3\text{-}7)$$

for $t \in (0, \infty)$ and $x, y \in \mathbb{R}^d$. By (1-6) and Hölder's inequality, we have

$$\begin{aligned}
&\int_0^t E[\mathscr{E}(u \wedge \tau_1, u \wedge \tau_2; X^x)^q p^X(u, X_u; t, y)]\, du \\
&\qquad \le C_G^+ \int_0^t E\biggl[\mathscr{E}(u \wedge \tau_1, u \wedge \tau_2; X^x)^q (t-u)^{-\frac{d}{2}} \exp\biggl(-\frac{\gamma_G^+ |X_u - y|^2}{t-u}\biggr)\biggr] du \\
&\qquad \le C_G^+ \biggl(\int_0^t E[\mathscr{E}(u \wedge \tau_1, u \wedge \tau_2; X^x)^{(d+\varepsilon)q/\varepsilon}]\, du\biggr)^{\frac{\varepsilon}{d+\varepsilon}} \\
&\qquad\qquad \times \biggl(\int_0^t E\biggl[(t-u)^{-(d+\varepsilon)/2} \exp\biggl(-\frac{(d+\varepsilon)\gamma_G^+ |X_u^x - y|^2}{d(t-u)}\biggr)\biggr] du\biggr)^{\frac{d}{d+\varepsilon}}.
\end{aligned}$$

Hence, in view of Lemma 3.3, to show (3-7) it is sufficient to prove that

$$\int_0^t E\biggl[(t-u)^{-(d+\varepsilon)/2} \exp\biggl(-\frac{(d+\varepsilon)\gamma_G^+ |X_u^x - y|^2}{d(t-u)}\biggr)\biggr] du \le C t^{-(d+\varepsilon)/2+1} \exp\biggl(-\gamma \frac{|x-y|^2}{t}\biggr) \tag{3-8}$$

for $t \in (0, \infty)$ and $x, y \in \mathbb{R}^d$, where $C$ and $\gamma$ are constants depending on $d$, $\varepsilon$, $\gamma_G^+$, $C_G^+$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$.

Let $\tilde{\gamma} := (1 + \varepsilon/d)\gamma_G^+$. By (1-6) again, we have for $u \in (0, t)$ that

$$\begin{aligned}
&E\biggl[(t-u)^{-d/2} \exp\biggl(-\frac{\tilde{\gamma} |X_u^x - y|^2}{t-u}\biggr)\biggr] \\
&\qquad \le C_G^+ u^{-d/2} (t-u)^{-d/2} \int_{\mathbb{R}^d} \exp\biggl(-\frac{\tilde{\gamma} |\xi - y|^2}{t-u}\biggr) \exp\biggl(-\frac{\gamma_G^+ |\xi - x|^2}{u}\biggr) d\xi \\
&\qquad = C_G^+ u^{-d/2} (t-u)^{-d/2} \exp\biggl(-\frac{\gamma_G^+ \tilde{\gamma}}{u\tilde{\gamma} + (t-u)\gamma_G^+} |x-y|^2\biggr) \\
&\qquad\qquad \times \int_{\mathbb{R}^d} \exp\biggl(-\frac{u\tilde{\gamma} + (t-u)\gamma_G^+}{u(t-u)} \biggl|\xi - \frac{\gamma_G^+ (t-u)x + \tilde{\gamma} u y}{u\tilde{\gamma} + (t-u)\gamma_G^+}\biggr|^2\biggr) d\xi \\
&\qquad = (2\pi)^{d/2} C_G^+ (u\tilde{\gamma} + (t-u)\gamma_G^+)^{-d/2} \exp\biggl(-\frac{\gamma_G^+ \tilde{\gamma}}{u\tilde{\gamma} + (t-u)\gamma_G^+} |x-y|^2\biggr) \\
&\qquad \le (2\pi)^{d/2} C_G^+ \gamma_G^{+\,-d/2} t^{-d/2} \exp\biggl(-\frac{\gamma_G^+ |x-y|^2}{t}\biggr).
\end{aligned}$$

Hence, there exists a positive constant $C$ depending on $d$, $\gamma_G^+$, $C_G^+$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$, such that

$$E\biggl[(t-u)^{-d/2} \exp\biggl(-\frac{\tilde{\gamma} |X_u^x - y|^2}{t-u}\biggr)\biggr] \le C t^{-d/2} \exp\biggl(-\frac{\gamma_G^+ |x-y|^2}{t}\biggr), \quad u \in (0, t).$$

Thus, we obtain (3-8). □

**Lemma 3.5.** *Let $t \in (0, \infty)$ and $\tau_1, \tau_2$ be stopping times such that $0 \le \tau_1 \le \tau_2 \le t$ almost surely. Then, it holds that*

$$p^X(0, x; t, y) E^{X_t^x = y}[\mathscr{E}(s \wedge \tau_1, s \wedge \tau_2; X^x)^q] \le C t^{-d/2} e^{C(1+q^2)t} \exp\biggl(-\frac{\gamma |x-y|^2}{t}\biggr)$$

*for* $t \in (0,\infty)$, $x, y \in \mathbb{R}^d$, $q \in \mathbb{R}$ *and* $s \in [0,t)$, *where* $C$ *and* $\gamma$ *are positive constants depending on* $d$, $\gamma_{\mathrm{G}}^-$, $\gamma_{\mathrm{G}}^+$, $C_{\mathrm{G}}^-$, $C_{\mathrm{G}}^+$, $m$, $M$, $\theta$, $\Lambda$, $\|b\|_\infty$ *and* $\|c\|_\infty$.

*Proof.* Let $s_1, s_2 \in (0,t)$ such that $s_1 \le s_2$. In view of (2-10), by (2-11) and Itō's formula we have

$$\begin{aligned}
&p^X(0,x;t,y)E^{X_t^x=y}[\mathscr{E}((s\wedge\tau_1)\vee s_1,(s\wedge\tau_2)\wedge s_2;X^x)^q]\\
&\quad= E[p^X(s_2,X_{s_2}^x;t,y)\mathscr{E}((s\wedge\tau_1)\vee s_1,(s\wedge\tau_2)\wedge s_2;X^x)^q]\\
&\quad= p^X(0,x;t,y)+E\biggl[\int_{(s\wedge\tau_1)\vee s_1}^{(s\wedge\tau_2)\wedge s_2}\Bigl(\frac{\partial}{\partial u}p^X(u,\xi;t,y)\Bigr)\Big|_{\xi=X_u^x}\mathscr{E}((s\wedge\tau_1)\vee s_1,u;X^x)^q\,du\biggr]\\
&\qquad+E\biggl[\int_{(s\wedge\tau_1)\vee s_1}^{(s\wedge\tau_2)\wedge s_2}(L_u^Xp^X(u,\cdot\,;t,y))(X_u^x)\mathscr{E}((s\wedge\tau_1)\vee s_1,u;X^x)^q\,du\biggr]\\
&\qquad+\frac{q}{2}E\biggl[\int_{(s\wedge\tau_1)\vee s_1}^{(s\wedge\tau_2)\wedge s_2}\mathscr{E}((s\wedge\tau_1)\vee s_1,u;X^x)^q\times\bigl\langle\sigma(u,X_u^x)^T\nabla_zp^X(u,z;t,y)|_{z=X_u^x},b_\sigma(u,X_u^x)\bigr\rangle\,du\biggr]\\
&\qquad+\frac{q^2}{2}E\biggl[\int_{(s\wedge\tau_1)\vee s_1}^{(s\wedge\tau_2)\wedge s_2}p^X(u,X_u^x;t,y)\mathscr{E}((s\wedge\tau_1)\vee s_1,u;X^x)^q|b_\sigma(u,X_u^x)|^2\,du\biggr]\\
&\qquad+qE\biggl[\int_{(s\wedge\tau_1)\vee s_1}^{(s\wedge\tau_2)\wedge s_2}p^X(u,X_u^x;t,y)\mathscr{E}((s\wedge\tau_1)\vee s_1,u;X^x)^qc(u,X_u^x)\,du\biggr].
\end{aligned}$$

Hence, by (2-6) we obtain

$$\begin{aligned}
&p^X(0,x;t,y)E^{X_t^x=y}[\mathscr{E}((s\wedge\tau_1)\vee s_1,(s\wedge\tau_2)\wedge s_2;X^x)^q]\\
&= p^X(0,x;t,y)\\
&\quad+\frac{q}{2}E\biggl[\int_{(s\wedge\tau_1)\vee s_1}^{(s\wedge\tau_2)\wedge s_2}\mathscr{E}((s\wedge\tau_1)\vee s_1,u;X^x)^q\times\bigl\langle\sigma(u,X_u^x)^T\nabla_zp^X(u,z;t,y)|_{z=X_u^x},b_\sigma(u,X_u^x)\bigr\rangle\,du\biggr]\\
&\quad+\frac{q^2}{2}E\biggl[\int_{(s\wedge\tau_1)\vee s_1}^{(s\wedge\tau_2)\wedge s_2}p^X(u,X_u^x;t,y)\mathscr{E}((s\wedge\tau_1)\vee s_1,u;X^x)^q|b_\sigma(u,X_u^x)|^2\,du\biggr]\\
&\quad+qE\biggl[\int_{(s\wedge\tau_1)\vee s_1}^{(s\wedge\tau_2)\wedge s_2}p^X(u,X_u^x;t,y)\mathscr{E}((s\wedge\tau_1)\vee s_1,u;X^x)^qc(u,X_u^x)\,du\biggr].
\end{aligned}$$

In view of the boundedness of $\det\sigma$, $b$ and $c$, the desired estimate is obtained, once we show the estimates

$$E\biggl[\int_{s_1}^{s_2}\mathscr{E}(u\wedge[(s\wedge\tau_1)\vee s_1],u;X^x)^q\,p^X(u,X_u^x;t,y)\,du\biggr]
\le Ct^{-d/2+1-\varepsilon}e^{C(1+q^2)t}\exp\Bigl(-\gamma\frac{|x-y|^2}{t}\Bigr), \tag{3-9}$$

$$E\biggl[\int_{s_1}^{s_2}\mathscr{E}(u\wedge[(s\wedge\tau_1)\vee s_1],u;X^x)^q\bigl|\nabla_\xi p^X(u,\xi;t,y)|_{\xi=X_u^x}\bigr|\,du\biggr]
\le Ct^{-d/2+(1-\varepsilon)/2}e^{C(1+q^2)t}\exp\Bigl(-\frac{\gamma|x-y|^2}{t}\Bigr) \tag{3-10}$$

for sufficiently small $\varepsilon > 0$, where $C$ and $\gamma$ are positive constants depending on $d$, $\varepsilon$, $\gamma_{\mathrm{G}}^{-}$, $\gamma_{\mathrm{G}}^{+}$, $C_{\mathrm{G}}^{-}$, $C_{\mathrm{G}}^{+}$, $m$, $M$, $\theta$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$. The first estimate (3-9) follows, because by (2-11) and Lemma 3.4 we have

$$E\biggl[\int_{s_1}^{s_2} \mathscr{E}(u \wedge [(s \wedge \tau_1) \vee s_1], u; X^x)^q p^X(u, X_u^x; t, y)\, du\biggr] \\ = p^X(0, x; t, y) E^{X_t^x = y}\biggl[\int_{s_1}^{s_2} \mathscr{E}(u \wedge [(s \wedge \tau_1) \vee s_1], u; X^x)^q\, du\biggr] \\ \le C t^{-d/2+1-\varepsilon} e^{C(1+q^2)t} \exp\biggl(-\gamma \frac{|x-y|^2}{t}\biggr),$$

where $C$ and $\gamma$ are positive constants depending on $d$, $\varepsilon$, $\gamma_{\mathrm{G}}^{+}$, $C_{\mathrm{G}}^{+}$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$. Now we show (3-10). By (2-11) and Hölder's inequality, we have

$$E\biggl[\int_{s_1}^{s_2} \mathscr{E}(u \wedge [(s \wedge \tau_1) \vee s_1], u; X^x)^q \bigl|\nabla_\xi p^X(u, \xi; t, y)|_{\xi = X_u^x}\bigr|\, du\biggr] \\ = p^X(0, x; t, y) E^{X_t^x = y}\biggl[\int_{s_1}^{s_2} \mathscr{E}(u \wedge [(s \wedge \tau_1) \vee s_1], u; X^x)^q \frac{|\nabla_\xi p^X(u, \xi; t, y)|_{\xi = X_u^x}|}{p^X(u, X_u^x; t, y)}\, du\biggr] \\ \le p^X(0, x; t, y) \biggl(\int_{s_1}^{s_2} E^{X_t^x = y}\Bigl[\mathscr{E}(u \wedge [(s \wedge \tau_1) \vee s_1], u; X^x)^{2q/(1-2\varepsilon)}\Bigr]\, du\biggr)^{1/2-\varepsilon} \\ \times \biggl(\int_{s_1}^{s_2} [u(t-u)]^{-1/2}\, du\biggr)^{\varepsilon} \biggl(\int_{s_1}^{s_2} [u(t-u)]^{\varepsilon} E^{X_t^x = y}\biggl[\biggl(\frac{|\nabla_\xi p^X(u, \xi; t, y)|_{\xi = X_u^x}|}{p^X(u, X_u^x; t, y)}\biggr)^2\biggr]\, du\biggr)^{1/2}.$$

Lemma 3.4 and (1-6) imply that

$$E\biggl[\int_{s_1}^{s_2} \mathscr{E}(u \wedge [(s \wedge \tau_1) \vee s_1], u; X^x)^q \bigl|\nabla_\xi p^X(u, \xi; t, y)|_{\xi = X_u^x}\bigr|\, du\biggr] \\ \le C t^{-d/4+1/2-\varepsilon} e^{C(1+q^2)t} \exp\biggl(-\gamma \frac{|x-y|^2}{t}\biggr) \\ \times \biggl(p^X(0, x; t, y) \int_{s_1}^{s_2} [u(t-u)]^{\varepsilon} E^{X_t^x = y}\biggl[\biggl(\frac{|\nabla_\xi p^X(u, \xi; t, y)|_{\xi = X_u^x}|}{p^X(u, X_u^x; t, y)}\biggr)^2\biggr]\, du\biggr)^{1/2},$$

where $C$ and $\gamma$ are positive constants depending on $d$, $\varepsilon$, $\gamma_{\mathrm{G}}^{+}$, $C_{\mathrm{G}}^{+}$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$. Hence, to show (3-10), it is sufficient to prove

$$p^X(0, x; t, y) \int_0^t [u(t-u)]^{\varepsilon/2} E^{X_t^x = y}\biggl[\biggl(\frac{|\nabla_\xi p^X(u, \xi; t, y)|_{\xi = X_u^x}|}{p^X(u, X_u^x; t, y)}\biggr)^2\biggr]\, du \\ \le C(t^{-d/2+\varepsilon} + t^{-d/2+\varepsilon}|\log t|), \quad \text{(3-11)}$$

where $C$ is a constant depending on $d$, $\varepsilon$, $\gamma_{\mathrm{G}}^-$, $\gamma_{\mathrm{G}}^+$, $C_{\mathrm{G}}^-$, $C_{\mathrm{G}}^+$, $m$, $M$, $\theta$ and $\Lambda$. Equation (2-11) implies that

$$\begin{aligned}
p^X(0,x;t,y)&\int_0^t [u(t-u)]^{\varepsilon/2} E^{X_t^x=y}\left[\left(\frac{|\nabla_\xi p^X(u,\xi;t,y)|_{\xi=X_u^x}|}{p^X(u,X_u^x;t,y)}\right)^2\right] du \\
&= \int_0^t [u(t-u)]^{\varepsilon/2} E\left[\left(\frac{|\nabla_\xi p^X(u,\xi;t,y)|_{\xi=X_u^x}|}{p^X(u,X_u^x;t,y)}\right)^2 p^X(u,X_u^x;t,y)\right] du \\
&= \int_0^t [u(t-u)]^{\varepsilon/2} \int_{\mathbb{R}^d} \left(\frac{\left|\nabla_\xi p^X(u,\xi;t,y)\right|}{p^X(u,\xi;t,y)}\right)^2 p^X(u,\xi;t,y) p^X(0,x;u,\xi)\, d\xi\, du.
\end{aligned}$$

By (1-6), we have

$$\begin{aligned}
p^X(0,x;t,y)&\int_0^t [u(t-u)]^{\varepsilon/2} E^{X_t^x=y}\left[\left(\frac{|\nabla_\xi p^X(u,\xi;t,y)|_{\xi=X_u^x}|}{p^X(u,X_u^x;t,y)}\right)^2\right] du \\
&\le (C_{\mathrm{G}}^+)^2 \int_0^t \int_{\mathbb{R}^d} \left(\frac{|\nabla_\xi p^X(u,\xi;t,y)|}{p^X(u,\xi;t,y)}\right)^2 [u(t-u)]^{-(d-\varepsilon)/2} \\
&\qquad \times \exp\left[-\gamma_{\mathrm{G}}^+\left(\frac{|\xi-x|^2}{u} + \frac{|y-\xi|^2}{t-u}\right)\right] d\xi\, du. \qquad (3\text{-}12)
\end{aligned}$$

For fixed $t$, $x$ and $y$, let

$$\phi(u,\xi) := [u(t-u)]^{-(d-\varepsilon)/2} \exp\left[-\gamma_{\mathrm{G}}^+\left(\frac{|\xi-x|^2}{u} + \frac{|y-\xi|^2}{t-u}\right)\right].$$

Denote the surface area of the unit sphere in $\mathbb{R}^d$ by $\omega_d$ for $d \ge 2$. In the case $d = 1$, let $\omega_d = 2$. Explicit calculation implies that

$$\begin{aligned}
&\int_0^t \left(\int_{\mathbb{R}^d} e^{2m|\xi|/(\theta-2)} \phi(u,\xi)^{\theta/(\theta-2)}\, d\xi\right)^{(\theta-2)/\theta} du \\
&\quad = \int_0^t [u(t-u)]^{-(d-\varepsilon)/2} \times \left(\int_{\mathbb{R}^d} e^{2m|\xi|/(\theta-2)} \exp\left[-\frac{\gamma_{\mathrm{G}}^+\theta}{\theta-2}\left(\frac{|\xi-x|^2}{u} + \frac{|y-\xi|^2}{t-u}\right)\right] d\xi\right)^{(\theta-2)/\theta} du \\
&\quad = \exp\left(-\frac{\gamma_{\mathrm{G}}^+|x-y|^2}{t}\right) \int_0^t [u(t-u)]^{-(d-\varepsilon)/2} \\
&\qquad \times \left(\int_{\mathbb{R}^d} e^{2m|\xi|/(\theta-2)} \exp\left[-\frac{\gamma_{\mathrm{G}}^+\theta t}{(\theta-2)u(t-u)}\left|\xi - \frac{(t-u)x+uy}{t}\right|^2\right] d\xi\right)^{(\theta-2)/\theta} du.
\end{aligned}$$

Hence, noting that for $\mu_1 \in [0,\infty)$, $\mu_2 \in (0,\infty)$ and $\nu \in \mathbb{R}^d$

$$\begin{aligned}
&\int_{\mathbb{R}^d} e^{\mu_1|\xi|} \exp(-\mu_2|\xi-\nu|^2)\,d\xi \\
&\quad = \int_{\mathbb{R}^d} e^{\mu_1|\xi+\nu|} \exp(-\mu_2|\xi|^2)\,d\xi \\
&\quad \le e^{\mu_1|\nu|} \int_{\mathbb{R}^d} \exp(\mu_1|\xi| - \mu_2|\xi|^2)\,d\xi \\
&\quad = \omega_d e^{\mu_1|\nu|} \int_{(0,\infty)} r^{d-1} \exp(\mu_1 r - \mu_2 r^2)\,dr \\
&\quad = \omega_d \mu_2^{-d/2} e^{\mu_1|\nu|} \int_{(0,\infty)} r^{d-1} \exp\biggl(\frac{\mu_1}{\sqrt{\mu_2}} r - r^2\biggr)\,dr \\
&\quad = \omega_d \mu_2^{-d/2} \int_0^{1+\mu_1/\sqrt{\mu_2}} r^{d-1} \exp\biggl(\frac{\mu_1}{\sqrt{\mu_2}} r - r^2\biggr)\,dr + \omega_d \mu_2^{-d/2} \int_{1+\mu_1/\sqrt{\mu_2}}^{\infty} r^{d-1} \exp\biggl(\frac{\mu_1}{\sqrt{\mu_2}} r - r^2\biggr)\,dr \\
&\quad \le \frac{\omega_d \mu_2^{-d/2}}{d} \biggl(1 + \frac{\mu_1}{\sqrt{\mu_2}}\biggr)^d \exp\biggl[\frac{\mu_1}{\sqrt{\mu_2}}\biggl(1 + \frac{\mu_1}{\sqrt{\mu_2}}\biggr)\biggr] + \omega_d \mu_2^{-d/2} \int_{1+\mu_1/\sqrt{\mu_2}}^{\infty} r^{d-1} e^{-r}\,dr \\
&\quad \le C \mu_2^{-d/2} \exp\biggl[C \frac{\mu_1}{\sqrt{\mu_2}}\biggl(1 + \frac{\mu_1}{\sqrt{\mu_2}}\biggr)\biggr],
\end{aligned}$$

where $C$ is a constant depending on $d$, we have

$$\begin{aligned}
&\int_0^t \biggl(\int_{\mathbb{R}^d} e^{2m|\xi|/(\theta-2)} \phi(u,\xi)^{\theta/(\theta-2)}\,d\xi\biggr)^{(\theta-2)/\theta} du \\
&\qquad \le C_1 t^{-d/2+d/\theta} \exp\biggl(-\frac{\gamma_{\mathrm{G}}^{+}|x-y|^2}{t}\biggr) \\
&\qquad\qquad \times \int_0^t [u(t-u)]^{-d/\theta+\varepsilon/2} \exp\biggl[C_1 \sqrt{u\biggl(1-\frac{u}{t}\biggr)}\biggl(1 + \sqrt{u\biggl(1-\frac{u}{t}\biggr)}\biggr)\biggr]\,du \\
&\qquad \le C_2 t^{-d/2+d/\theta} e^{C_2 t} \int_0^t [u(t-u)]^{-d/\theta+\varepsilon/2}\,du \\
&\qquad \le C_3 t^{-d/2+1-d/\theta+\varepsilon/2} e^{C_2 t},
\end{aligned}$$

where $C_1, C_2, C_3$ are constants depending on $d$, $\varepsilon$, $\gamma_{\mathrm{G}}^{+}$, $m$ and $\theta$. Hence, by Hölder's inequality and (1-3), we have

$$\begin{aligned}
&\sum_{i,j=1}^{d} \int_{s_1}^{s_2} \int_{\mathbb{R}^d} \biggl|\frac{\partial}{\partial \xi_j} a_{ij}(u,\xi)\biggr|^2 \phi(u,\xi)\,d\xi\,du \\
&\qquad \le \int_0^t \biggl(\int_{\mathbb{R}^d} e^{2m|\xi|/(\theta-2)} \phi(u,\xi)^{\theta/(\theta-2)}\,d\xi\biggr)^{(\theta-2)/\theta} du \sum_{i,j=1}^{d} \biggl(\sup_{u\in[0,t]} \int_{\mathbb{R}^d} \biggl|\frac{\partial}{\partial \xi_j} a_{ij}(u,\xi)\biggr|^{\theta} e^{-m|\xi|}\,d\xi\biggr)^{2/\theta} \\
&\qquad \le C t^{-d/2+1-d/\theta+\varepsilon/2} e^{Ct},
\end{aligned}$$

where $C$ is a constant depending on $d$, $\varepsilon$, $\gamma_{\mathrm{G}}^{+}$, $m$, $M$ and $\theta$. On the other hand, by explicit calculation, we have

$$\begin{gathered}
\lim_{s\downarrow 0}(1+|\log(t-s)|)\int_{\mathbb{R}^d}\phi(s,\xi)\,d\xi=0,\\
\lim_{s\downarrow 0}(t-s)^{-1}\int_{\mathbb{R}^d}|y-\xi|^2\phi(s,\xi)\,d\xi=0,\\
\lim_{s\uparrow t}(1+|\log(t-s)|)\int_{\mathbb{R}^d}\phi(s,\xi)\,d\xi=0,\\
\lim_{s\uparrow t}(t-s)^{-1}\int_{\mathbb{R}^d}|y-\xi|^2\phi(s,\xi)\,d\xi=0,\\
\int_0^t\int_{\mathbb{R}^d}\frac{|\nabla_\xi\phi(u,\xi)|^2}{\phi(u,\xi)}\,d\xi\,du\le Ct^{-d/2+\varepsilon},\\
\int_0^t(1+|\log(t-u)|)\int_{\mathbb{R}^d}\left|\frac{\partial}{\partial u}\phi(u,\xi)\right|d\xi\,du\le Ct^{-d/2+\varepsilon}|\log t|,\\
\int_0^t(t-u)^{-1}\int_{\mathbb{R}^d}|y-\xi|^2\left|\frac{\partial}{\partial u}\phi(u,\xi)\right|d\xi\,du\le Ct^{-d/2+\varepsilon},
\end{gathered}$$

where $C$ is a constant depending on $d$, $\varepsilon$ and $\gamma_{\mathrm{G}}^{+}$. In view of these results, applying Lemma 3.1 to (3-12), we obtain (3-11). □

From Lemma 3.5 we can easily show the Gaussian estimate for $p$ with constants depending on the suitable constants.

**Proposition 3.6.** *For $s, t \in [0,\infty)$ such that $s < t$, and $x, y \in \mathbb{R}^d$,*

$$\frac{C_1e^{-C_1(t-s)}}{(t-s)^{\frac{d}{2}}}\exp\left(-\frac{\gamma_1|x-y|^2}{t-s}\right)\le p(s,x;t,y)\le\frac{C_2e^{C_2(t-s)}}{(t-s)^{\frac{d}{2}}}\exp\left(-\frac{\gamma_2|x-y|^2}{t-s}\right),$$

*where $\gamma_1$, $\gamma_2$, $C_1$ and $C_2$ are positive constants depending on $d$, $\gamma_{\mathrm{G}}^{-}$, $\gamma_{\mathrm{G}}^{+}$, $C_{\mathrm{G}}^{-}$, $C_{\mathrm{G}}^{+}$, $m$, $M$, $\theta$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$.*

*Proof.* Since the argument follows even if $a$, $b$ and $c$ are replaced by $a(\cdot-s,\cdot)$, $b(\cdot-s,\cdot)$ and $c(\cdot-s,\cdot)$ respectively, it is sufficient to show that there exist positive constants $\gamma_1$, $\gamma_2$, $C_1$ and $C_2$ depending on $d$, $m$, $M$, $\theta$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$ such that

$$C_1t^{-d/2}e^{-C_1t}\exp\left(-\frac{\gamma_1|x-y|^2}{t}\right)\le p(0,x;t,y)\le C_2t^{-d/2}e^{C_2t}\exp\left(-\frac{\gamma_2|x-y|^2}{t}\right) \tag{3-13}$$

for $t\in(0,\infty)$ and $x, y\in\mathbb{R}^d$. The upper estimate in (3-13) follows immediately from (2-8) and Lemma 3.5.

Now we prove the lower estimate in (3-13). From Hölder's inequality, it follows that

$$1\le E^{X_t^x=y}[\mathscr{E}(s\wedge\tau_1,s\wedge\tau_2;X^x)^{-1}]E^{X_t^x=y}[\mathscr{E}(s\wedge\tau_1,s\wedge\tau_2;X^x)].$$

Hence, by Lemma 3.5, we have

$$\begin{aligned} p^X(0,x;t,y)E^{X_t^x=y}[\mathscr{E}(s\wedge\tau_1,s\wedge\tau_2;X^x)] &\geq \frac{p^X(0,x;t,y)^2}{p^X(0,x;t,y)E^{X_t^x=y}[\mathscr{E}(s\wedge\tau_1,s\wedge\tau_2;X^x)^{-1}]} \\ &\geq Ct^{d/2}e^{-C't}p^X(0,x;t,y)^2, \end{aligned}$$

where $C$ and $C'$ are positive constants depending on $d$, $m$, $M$, $\theta$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$. This inequality, (2-8) and (1-6) imply the lower bound in (3-13). □

## 4. The regularity of $p(0, x; t, y)$ in $x$

Assume that $a$ is smooth and set notation as in Section 2. In this section, we prove the Hölder continuity of $p(0, x; t, y)$ in $x$, with constant depending only on suitable ones. The precise statement is as follows:

**Proposition 4.1.** *For any $R > 0$ and sufficiently small $\varepsilon > 0$, there exists a constant $C$ depending on $d$, $\varepsilon$, $\gamma_G^-$, $\gamma_G^+$, $C_G^-$, $C_G^+$, $m$, $M$, $\theta$, $R$, $\rho_R$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$ such that*

$$|p(0,x;t,y)-p(0,z;t,y)| \leq Ct^{-d/2-1}e^{Ct}|x-z|^{1-\varepsilon}$$

*for $t \in (0,\infty)$, $x, z \in B(0; R/2)$ and $y \in \mathbb{R}^d$.*

We use the coupling method; see, for example, [Lindvall and Rogers 1986; Cranston 1991]. Let $x, z \in \mathbb{R}^d$. Given $(X^x, B)$ defined by (2-3), we consider the stochastic process $Z^z$ defined by

$$\begin{cases} Z_t^z = z + \displaystyle\int_0^{t\wedge\tau} \sigma(s, Z_s^z)\, d\widetilde{B}_s + \int_{t\wedge\tau}^t \sigma(s, Z_s^z)\, dB_s, \\ \widetilde{B}_t = \displaystyle\int_0^{t\wedge\tau} \left( I - \frac{2(\sigma(s,Z_s^z)^{-1}(X_s^x - Z_s^z)) \otimes (\sigma(s,Z_s^z)^{-1}(X_s^x - Z_s^z))}{|\sigma(s,Z_s^z)^{-1}(X_s^x - Z_s^z)|^2} \right) dB_s, \end{cases} \tag{4-1}$$

where $\tau$ is the stopping time defined by $\tau := \inf\{t \geq 0 : X_t^x = Z_t^z\}$.

To see the existence and uniqueness of $Z^z$, for each $n \in \mathbb{N}$ consider the following stochastic differential equation for $\check{Z}^{n,z}$:

$$\begin{cases} d\check{Z}_t^{n,z} = \sigma(t, \check{Z}_t^{n,z}) \times \left( I - \dfrac{2(\sigma(t,\check{Z}_t^{n,z})^{-1}(X_t^x - \check{Z}_t^{n,z})) \otimes (\sigma(t,\check{Z}_t^{n,z})^{-1}(X_t^x - \check{Z}_t^{n,z}))}{|\sigma(t,\check{Z}_t^{n,z})^{-1}(X_t^x - \check{Z}_t^{n,z})|^2} \right) \mathbb{1}_{\{t<\check{\tau}_n\}}\, dB_t \\ \check{Z}_0^{n,z} = z, \end{cases}$$

where $\check{\tau}_n := \inf\{t \geq 0 : |X_t^x - \check{Z}_t^{n,z}| < 1/n\}$. Note that the equations have random coefficients, since we are considering equations where $X^x$ is given. Now we see the existence and the uniqueness of $\check{Z}^{n,z}$. Let

$$G_n(t,\xi) := \begin{cases} \sigma(t,\xi)\left( I - \dfrac{2(\sigma(t,\xi)^{-1}(X_t^x - \xi)) \otimes (\sigma(t,\xi)^{-1}(X_t^x - \xi))}{|\sigma(t,\xi)^{-1}(X_t^x - \xi)|^2} \right) & \text{if } (t,\xi) \in [0,\infty) \times \mathbb{R}^d \\ & \quad \text{and } |X_t^x - \xi| \geq 1/n\}, \\ 0 & \text{otherwise.} \end{cases}$$

Then, there exists a constant $C_n$ such that $|G_n(t,\xi)-G_n(t,\eta)|\le C_n$ for $t\in[0,\infty)$ and $\xi,\eta\in\mathbb{R}^d$. Note that $C_n$ is nonrandom. Let $Y$, $W$ be stochastic processes satisfying

$$\begin{cases} dY_t = G_n(t,Y_t)\mathbb{1}_{\{t<\tau_n^Y\}}\,dB_t, \\ Y_0 = z, \end{cases} \qquad \begin{cases} dW_t = G_n(t,W_t)\mathbb{1}_{\{t<\tau_n^W\}}\,dB_t, \\ W_0 = z, \end{cases} \tag{4-2}$$

where $\tau_n^Y := \inf\{t\ge 0 : |X_t^x - Y_t| < 1/n\}$ and $\tau_n^W := \inf\{t\ge 0 : |X_t^x - W_t| < 1/n\}$. Then, by Proposition 1.1(iv), Chapter II and (6.16) of Theorem 6.10, Chapter I of [Ikeda and Watanabe 1989], we have

$$\begin{aligned} E\Big[\sup_{s\in[0,t]} |Y_{s\wedge\tau_n^Y\wedge\tau_n^W} - W_{s\wedge\tau_n^Y\wedge\tau_n^W}|^2\Big] &= E\Big[\sup_{s\in[0,t]}\Big|\int_0^{s\wedge\tau_n^Y\wedge\tau_n^W} (G_n(v,Y_v)\mathbb{1}_{\{v<\tau_n^Y\}} - G_n(v,W_v)\mathbb{1}_{\{v<\tau_n^W\}})\,dB_v\Big|^2\Big] \\ &= E\Big[\sup_{s\in[0,t]}\Big|\int_0^{s} (G_n(v,Y_v) - G_n(v,W_v))\mathbb{1}_{\{v<\tau_n^Y\}}\mathbb{1}_{\{v<\tau_n^W\}}\,dB_v\Big|^2\Big] \\ &\le 4E\Big[\int_0^t |G_n(v,Y_v) - G_n(v,W_v)|^2\mathbb{1}_{\{v<\tau_n^Y\}}\mathbb{1}_{\{v<\tau_n^W\}}\,dv\Big] \\ &\le 4C_n^2\int_0^t E\Big[|Y_v - W_v|^2\mathbb{1}_{\{v<\tau_n^Y\}}\mathbb{1}_{\{v<\tau_n^W\}}\Big]\,dv \\ &\le 4C_n^2\int_0^t E\Big[\sup_{s\in[0,v]} |Y_{s\wedge\tau_n^Y\wedge\tau_n^W} - W_{s\wedge\tau_n^Y\wedge\tau_n^W}\Big|^2\Big]\,dv. \end{aligned}$$

Hence, by Gronwall's inequality, we have

$$Y_{t\wedge\tau_n^Y\wedge\tau_n^W} = W_{t\wedge\tau_n^Y\wedge\tau_n^W}, \quad t\in[0,\infty) \tag{4-3}$$

almost surely. If $\tau_n^Y \le \tau_n^W$ and $\tau_n^Y < \infty$ for some events, then by letting $t\to\infty$ in (4-3) we have $Y_{\tau_n^Y} = W_{\tau_n^Y}$. Hence, $\tau_n^Y = \tau_n^W$ for these events. Similarly, if $\tau_n^Y \ge \tau_n^W$ and $\tau_n^W < \infty$ for some events, then we have $\tau_n^Y = \tau_n^W$ for these events. Therefore, we obtain

$$\tau_n^Y = \tau_n^W \tag{4-4}$$

almost surely. On the other hand, (4-2) implies that $Y_{\tau_n^Y} = Y_{t\vee\tau_n^Y}$ and $W_{\tau_n^W} = W_{t\vee\tau_n^W}$ for $t\in[0,\infty)$. Hence, by (4-3) and (4-4) we obtain that $Y_t = W_t$ for $t\in[0,\infty)$ almost surely. Thus, we have uniqueness. To see existence, let

$$\widehat{G}_n(t,\xi) := \sigma(t,\xi)\bigg(I - \frac{2(\sigma(t,\xi)^{-1}(X_t^x-\xi))\otimes(\sigma(t,\xi)^{-1}(X_t^x-\xi))}{(|\sigma(t,\xi)^{-1}(X_t^x-\xi)|\vee(\Lambda^{-(1/2)}n^{-1}))^2}\bigg)$$

for $t\in[0,\infty)$ and $\xi\in\mathbb{R}^d$. Then, there exists a constant $C_n$ such that $|\widehat{G}_n(t,\xi)-\widehat{G}_n(t,\eta)|\le C_n$ for $t\in[0,\infty)$ and $\xi,\eta\in\mathbb{R}^d$. Define a sequence of stochastic processes $\{Y^m : m\in\mathbb{N}\cup\{0\}\}$ by $Y_t^0 = z$ for $t\in[0,\infty)$ and

$$Y_t^m := z + \int_0^t \widehat{G}_n(s, Y_s^{m-1})\,dB_s \tag{4-5}$$

for $t \in [0, \infty)$ and $m \in \mathbb{N}$ by iteration. Then, by a similar calculation as above, we have for $m \in \mathbb{N}$ and $t \in [0, \infty)$ that

$$E\biggl[\sup_{s\in[0,t]} |Y_s^{m+1} - Y_s^m|^2\biggr] \le 4C_n^2 \int_0^t E\biggl[\sup_{s\in[0,v]} |Y_s^m - Y_s^{m-1}|^2\biggr]\,dv.$$

Applying this inequality iteratively, for $m \in \mathbb{N}$ and $t \in [0, \infty)$ we obtain

$$\begin{aligned} E\biggl[\sup_{s\in[0,t]} |Y_s^{m+1} - Y_s^m|^2\biggr] &\le (4C_n^2)^m \int_0^t \int_0^{v_m} \cdots \int_0^{v_2} E\biggl[\sup_{s\in[0,v_1]} |Y_s^1 - Y_s^0|^2\biggr]\,dv_1\,dv_2 \cdots dv_m \\ &= (4C_n^2)^m \int_0^t \int_0^{v_m} \cdots \int_0^{v_2} E\biggl[\sup_{s\in[0,v_1]} \biggl|\int_0^s \widehat{G}_n(w, z)\,dB_w\biggr|^2\biggr]\,dv_1\,dv_2 \cdots dv_m \\ &\le \frac{(4C_n^2)^m t^m}{(m+1)!} E\biggl[\sup_{s\in[0,t]} \biggl|\int_0^s \widehat{G}_n(w, z)\,dB_w\biggr|^2\biggr]. \end{aligned}$$

Since [Ikeda and Watanabe 1989, (6.16) of Theorem 6.10, Chapter I] implies

$$E\biggl[\sup_{s\in[0,t]} \biggl|\int_0^s \widehat{G}_n(w, z)\,dB_w\biggr|^2\biggr] \le 4E\biggl[\int_0^t |\widehat{G}_n(w, z)|^2\,dw\biggr] < \infty$$

for $t \in [0, \infty)$, we have

$$\sum_{m=1}^{\infty} E\biggl[\sup_{s\in[0,t]} |Y_s^{m+1} - Y_s^m|^2\biggr] < \infty$$

for $t \in [0, \infty)$. Hence, $\{Y^m\}$ is a Cauchy sequence in $L^2(\Omega; \mathscr{C})$, where $\mathscr{C}$ is the complete metric space $C([0, \infty); \mathbb{R}^d)$ with distance function given by

$$D(w, w') := \sum_{k=1}^{\infty} 2^{-k} \Bigl(\sup_{t\in[0,k]} |w(t) - w'(t)|\Bigr) \wedge 1, \qquad w, w' \in C([0, \infty); \mathbb{R}^d).$$

Therefore, there exists a stochastic process $Y$ in $L^2(\Omega; \mathscr{C})$ which satisfies

$$\lim_{m\to\infty} E\biggl[\sup_{s\in[0,t]} |Y_s - Y_s^m|^2\biggr] = 0$$

for $t \in [0, \infty)$. By taking the limit in (4-5) as $m \to \infty$, we have

$$Y_t = z + \int_0^t \widehat{G}_n(s, Y_s)\,dB_s, \qquad t \in [0, \infty) \tag{4-6}$$

almost surely. Let $\tau_n^Y := \inf\{t \ge 0 : |X_t^x - Y_t| < 1/n\}$. Note that (2-2) implies that

$$|\sigma(t, Y_t)^{-1}(X_t^x - Y_t)| \vee (\Lambda^{-(1/2)} n^{-1}) = |\sigma(t, Y_t)^{-1}(X_t^x - Y_t)|, \qquad t \in [0, \tau_n^Y)$$

almost surely. Applying [Ikeda and Watanabe 1989, Proposition 1.1(iv), Chapter II] to (4-6), we see that the $Y_{\cdot\wedge\tau_n^Y}$ satisfy the stochastic differential equation for $\check{Z}^{n,z}$. Thus, we obtain existence.

We remark that $\{\check{Z}^{n,z} : n \in \mathbb{N}\}$ is consistent; i.e., $\check{Z}^{m,z}_{t\wedge\check{\tau}_n} = \check{Z}^{n,z}_{t\wedge\check{\tau}_n}$ for $m > n$ almost surely. This fact is immediately obtained by [Ikeda and Watanabe 1989, Proposition 1.1(iv), Chapter II] and uniqueness. Define the stochastic processes $(Z^z_t, \widetilde{B}_t; t \in [0, \tau))$ by

$$
\begin{aligned}
Z^z_t &= \check{Z}^{n,z}_t, \\
\widetilde{B}_t &= \int_0^t \left( I - \frac{2(\sigma(s, \check{Z}^{n,z}_s)^{-1}(X^x_s - \check{Z}^{n,z}_s)) \otimes (\sigma(s, \check{Z}^{n,z}_s)^{-1}(X^x_s - \check{Z}^{n,z}_s))}{|\sigma(s, \check{Z}^{n,z}_s)^{-1}(X^x_s - \check{Z}^{n,z}_s)|^2} \right) dB_s
\end{aligned}
$$

for $t \in [0, \tau^Z_n)$ and $n \in \mathbb{N}$, where $\tau^Z_n := \inf\{t \geq 0 : |X^x_t - Z^z_t| < 1/n\}$. Then, (4-1) holds for $t \in [0, \tau)$. On the other hand, by applying [Ikeda and Watanabe 1989, Proposition 1.1(iv), Chapter II], we have that $Z^z_{\cdot\wedge\tau^Z_n}$ solves the stochastic differential equation of $\check{Z}^{n,z}$ for $n \in \mathbb{N}$. Hence, $(Z^z_t, \widetilde{B}_t; t \in [0, \tau))$ are determined almost surely and uniquely. Let

$$
H_t := I - \frac{2[\sigma(t, Z^z_t)^{-1}(X^x_t - Z^z_t)] \otimes [\sigma(t, Z^z_t)^{-1}(X^x_t - Z^z_t)]}{|\sigma(t, Z^z_t)^{-1}(X^x_t - Z^z_t)|^2}
$$

for $t \in [0, \tau)$. Then, $H_t$ is an orthogonal matrix for all $t \in [0, \tau)$, and hence $\widetilde{B}_t$ is a $d$-dimensional Brownian motion for $t \in [0, \tau)$. Hence, $(Z^z_t, \widetilde{B}_t; t \in [0, \tau))$ are extended to $(Z^z_t, \widetilde{B}_t; t \in [0, \tau])$ almost surely and uniquely. By the Lipschitz continuity of $\sigma$, (4-1) is solved almost surely and uniquely for $t \in [\tau, \infty)$, and thus we obtain $(Z^z_t; t \in [0, \infty))$ almost surely and uniquely; see [Stroock and Varadhan 1979, Section 6.6]. From this fact we have that $Z^z_t$ is $\mathscr{F}_t$-measurable for $t \in [0, \infty)$. Hence, if $x = z$, $X^x$ and $Z^z$ have the same law. Moreover, $X^x_t = Z^z_t$ for $t \in [\tau, \infty)$ almost surely.

**Lemma 4.2.** *For $R > 0$ and sufficiently small $\varepsilon > 0$, there exist positive constants $C$ and $c_0$ depending on $d, \varepsilon, R, \rho_R$ and $\Lambda$ such that*

$$
E[t \wedge \tau] \leq C(1 + t^2)|x - z|^{1-\varepsilon} \tag{4-7}
$$

*for $t \in [0, \infty)$ and $x, z \in B(0; R/2)$ such that $|x - z| \leq c_0$.*

*Proof.* Let $R > 0$ and $x, z \in B(0; R/2)$. Define

$$
\xi_t := X^x_t - Z^z_t \quad \text{and} \quad \alpha_t := \sigma(t, X^x_t) - \sigma(t, Z^z_t)H_t.
$$

Then, by Itō's formula we have, for $t \in [0, \tau)$,

$$
d(|\xi_t|) = \left\langle \frac{\xi_t}{|\xi_t|}, \alpha_t \, dB_t \right\rangle + \frac{1}{2|\xi_t|} \left( \operatorname{tr}(\alpha_t \alpha_t^T) - \frac{|\alpha_t^T \xi_t|^2}{|\xi_t|^2} \right) dt, \tag{4-8}
$$

where $\operatorname{tr}(A)$ is the trace of the matrix $A$. Now we follow the argument in [Lindvall and Rogers 1986, Section 3]. Since

$$
\begin{aligned}
\alpha_t &= \sigma(t, X^x_t) - \sigma(t, Z^z_t) + \frac{2\xi_t \otimes (\sigma(t, Z^z_t)^{-1}\xi_t)}{|\sigma(t, Z^z_t)^{-1}\xi_t|^2} \\
&= \sigma(t, X^x_t) - \sigma(t, Z^z_t) + \frac{2\xi_t \xi_t^T (\sigma(t, Z^z_t)^{-1})^T}{|\sigma(t, Z^z_t)^{-1}\xi_t|^2},
\end{aligned}
$$

it holds that

$$\operatorname{tr}(\alpha_t\alpha_t^T)-\frac{|\alpha_t^T\xi_t|^2}{|\xi_t|^2}=\operatorname{tr}\big([\sigma(t,X_t^x)-\sigma(t,Z_t^z)][\sigma(t,X_t^x)-\sigma(t,Z_t^z)]^T\big)-\frac{|[\sigma(t,X_t^x)-\sigma(t,Z_t^z)]^T\xi_t|^2}{|\xi_t|^2}.$$

Hence, in view of (2-1), there exists a positive constant $\gamma_1$ depending on $d$ and $\Lambda$ such that

$$\left|\operatorname{tr}(\alpha_t\alpha_t^T)-\frac{|\alpha_t^T\xi_t|^2}{|\xi_t|^2}\right|\le\gamma_1\rho_R(|\xi_t|)\quad\text{for } t\in[0,\tau) \text{ such that } X_t^x, Z_t^z\in B(0;R). \tag{4-9}$$

On the other hand, following the argument in [Lindvall and Rogers 1986, Section 3], we have a positive constant $\gamma_2$ depending on $d$ and $\Lambda$ such that

$$\frac{|\alpha_t^T\xi_t|}{|\xi_t|}\ge\gamma_2^{-1}\quad\text{for } t\in[0,\tau) \text{ such that } |\sigma(t,X_t^x)-\sigma(t,Z_t^z)|\le 2\Lambda^{-1}. \tag{4-10}$$

Note that if $\rho_R(|\xi_t|)\le 2\Lambda^{-1}$ and $X_t^x, Z_t^z\in B(0;R)$, then $|\sigma(t,X_t^x)-\sigma(t,Z_t^z)|\le 2\Lambda^{-1}$. Let $\gamma:=\gamma_1\vee\gamma_2$. Define stopping times $\tau_n$ by $\tau_n:=\inf\{t>0:|X_t^x-Z_t^z|\le 1/n\}$ for $n\in\mathbb{N}$. For a given $\varepsilon>0$, let

$$\begin{aligned}\tilde\tau &:= \tau\wedge\inf\left\{t\in[0,\infty):\ \rho_R(|\xi_t|)>\frac{\varepsilon}{2\gamma^3}\wedge 2\Lambda^{-1},\ X_t^x\notin B(0;R) \text{ or } Z_t^z\notin B(0;R)\right\},\\ \tilde\tau_n &:= \tau_n\wedge\inf\left\{t\in[0,\infty):\ \rho_R(|\xi_t|)>\frac{\varepsilon}{2\gamma^3}\wedge 2\Lambda^{-1},\ X_t^x\notin B(0;R) \text{ or } Z_t^z\notin B(0;R)\right\}\end{aligned}$$

for $n\in\mathbb{N}$. Then, it holds that $\tilde\tau_n\uparrow\tilde\tau$ almost surely as $n\to\infty$. By Itō's formula, (4-8), (4-9) and (4-10), we have for $t\in[0,\infty)$ that

$$\begin{aligned}E[|\xi_{t\wedge\tilde\tau_n}|^{1-\varepsilon}] &= |x-z|^{1-\varepsilon}+(1-\varepsilon)E\left[\int_0^{t\wedge\tilde\tau_n}|\xi_s|^{-\varepsilon}\frac{1}{2|\xi_s|}\left(\operatorname{tr}(\alpha_s\alpha_s^T)-\frac{|\alpha_s^T\xi_s|^2}{|\xi_s|^2}\right)ds\right]\\ &\qquad -\frac{\varepsilon(1-\varepsilon)}{2}E\left[\int_0^{t\wedge\tilde\tau_n}|\xi_s|^{-1-\varepsilon}\frac{|\alpha_s^T\xi_s|^2}{|\xi_s|^2}\,ds\right]\\ &\le |x-z|^{1-\varepsilon}+\frac{(1-\varepsilon)\gamma}{2}E\left[\int_0^{t\wedge\tilde\tau_n}|\xi_s|^{-1-\varepsilon}\rho_R(|\xi_s|)\,ds\right]-\frac{\varepsilon(1-\varepsilon)}{2\gamma^2}E\left[\int_0^{t\wedge\tilde\tau_n}|\xi_s|^{-1-\varepsilon}\,ds\right]\\ &\le |x-z|^{1-\varepsilon}+\frac{\varepsilon(1-\varepsilon)}{4\gamma^2}E\left[\int_0^{t\wedge\tilde\tau_n}|\xi_s|^{-1-\varepsilon}\,ds\right]-\frac{\varepsilon(1-\varepsilon)}{2\gamma^2}E\left[\int_0^{t\wedge\tilde\tau_n}|\xi_s|^{-1-\varepsilon}\,ds\right]\\ &\le |x-z|^{1-\varepsilon}-\frac{\varepsilon(1-\varepsilon)}{4\gamma^2}E\left[\int_0^{t\wedge\tilde\tau_n}|\xi_s|^{-1-\varepsilon}\,ds\right]\\ &\le |x-z|^{1-\varepsilon}-\frac{\varepsilon(1-\varepsilon)}{2^{3+\varepsilon}\gamma^2R^{1+\varepsilon}}E[t\wedge\tilde\tau_n].\end{aligned}$$

Hence, it holds that

$$E[t\wedge\tilde\tau]\le C|x-z|^{1-\varepsilon}\quad\text{for } t\in[0,\infty), \tag{4-11}$$

where $C$ is a constant depending on $d$, $\varepsilon$, $R$ and $\Lambda$.

Now we consider the estimate of the expectation of $\tau$ by using that of $\tilde{\tau}$. To simplify the notation, let

$$\delta_0 := \frac{1}{3}\rho_R^{-1}\left(\frac{\varepsilon}{2\gamma^3} \wedge 2\Lambda^{-1}\right).$$

Since

$$|\xi_t| > 3\delta_0 \text{ implies } |X_t^x - x| > \delta_0,\ |Z_t^z - z| > \delta_0,\ \text{or } |x-z| > \delta_0,$$
$$X_t^x \notin B(0; R) \text{ or } Z_t^z \notin B(0; R) \text{ implies } |X_t^x - x| > \frac{R}{2} \text{ or } |Z_t^z - z| > \frac{R}{2},$$

we have, for $x, z \in B(0; R/2)$ such that $|x-z| \le \delta_0$,

$$P(\tau \ge t) \le P(\tilde{\tau} \ge t) + P\left(\sup_{s\in[0,t]} |X_s^x - x| > \delta_0 \wedge \frac{R}{2}\right) + P\left(\sup_{s\in[0,t]} |Z_s^z - z| > \delta_0 \wedge \frac{R}{2}\right). \tag{4-12}$$

Let $\eta = x$ or $z$. By Chebyshev's inequality and Burkholder's inequality we have

$$\begin{aligned}
P\left(\sup_{s\in[0,t]} |X_s^\eta - \eta| > \delta_0 \wedge \frac{R}{2}\right) &\le \left(\delta_0 \wedge \frac{R}{2}\right)^{-2/\varepsilon} E\left[\sup_{s\in[0,t]} |X_s^\eta - \eta|^{2/\varepsilon}\right] \\
&\le \left(\delta_0 \wedge \frac{R}{2}\right)^{-2/\varepsilon} E\left[\sup_{s\in[0,t]} \left|\int_0^s \sigma(u, X_u^\eta) dB_u\right|^{2/\varepsilon}\right] \\
&\le \left(\delta_0 \wedge \frac{R}{2}\right)^{-2/\varepsilon} CE\left[\left(\sum_{i,j=1}^d \int_0^t \sigma_{ij}(u, X_u^\eta)\sigma_{ji}(u, X_u^\eta)\, du\right)^{1/\varepsilon}\right] \\
&\le d^{1/\varepsilon}\left(\delta_0 \wedge \frac{R}{2}\right)^{-2/\varepsilon} C\Lambda^{1/\varepsilon} t^{1/\varepsilon},
\end{aligned}$$

where $C$ is a constant depending on $\varepsilon$. Hence, there exists a constant $C$ depending on $d$, $\varepsilon$, $R$, $\rho_R$ and $\Lambda$ such that

$$P\left(\sup_{s\in[0,t]} |X_s^\eta - \eta| > \delta_0 \wedge \frac{R}{2}\right) \le C|x-z| \tag{4-13}$$

for $\eta = x, z$ and $t \in [0, |x-z|^\varepsilon]$. By (4-11), (4-12) and (4-13) we have, for $x, z \in B(0; R/2)$ such that $|x-z| \le \delta_0$, and $t \in [0, |x-z|^\varepsilon]$,

$$\begin{aligned}
E[t \wedge \tau] &\le \int_0^t P(\tau \ge s)\, ds \\
&\le \int_0^t P(\tilde{\tau} \ge s)\, ds + t\left[P\left(\sup_{s\in[0,t]} |X_s^x - x| > \delta_0 \wedge \frac{R}{2}\right) + P\left(\sup_{s\in[0,t]} |X_s^z - z| > \delta_0 \wedge \frac{R}{2}\right)\right] \\
&\le C(1+t)|x-z|^{1-\varepsilon},
\end{aligned}$$

where $C$ is a constant depending on $d$, $\varepsilon$, $R$, $\rho_R$ and $\Lambda$. Therefore, we obtain

$$E[t \wedge \tau] \le C(1+t)|x-z|^{1-\varepsilon} \tag{4-14}$$

for $x, z \in B(0; R/2)$ such that $|x-z| \le \delta_0$ and $t \in [0, |x-z|^\varepsilon]$. By using Chebyshev's inequality, we calculate $E[t \wedge \tau]$ as

$$\begin{aligned} E[t\wedge\tau] = \int_0^{|x-z|^\varepsilon} P(\tau \ge s)\,ds + \int_{|x-z|^\varepsilon}^t P(\tau \ge s)\,ds &\le E[|x-z|^\varepsilon \wedge \tau] + tP(\tau \ge |x-z|^\varepsilon) \\ &\le E[|x-z|^\varepsilon \wedge \tau] + \frac{t}{|x-z|^\varepsilon} E[\tau \wedge |x-z|^\varepsilon] \\ &\le (1+t|x-z|^{-\varepsilon}) E[|x-z|^\varepsilon \wedge \tau]. \end{aligned}$$

Thus, applying (4-14) with $t = |x-z|^\varepsilon$ and choosing another small $\varepsilon$, we obtain (4-7) for all $t \in [0, \infty)$. □

**Lemma 4.3.** *For $R > 0$ and sufficiently small $\varepsilon > 0$, there exist positive constants $C$ and $c_0$ depending on $d$, $\varepsilon$, $C_{\mathrm{G}}^+$, $R$, $\rho_R$ and $\Lambda$ such that*

$$\begin{aligned} p^X(0, x; t, y) E^{X_t^x = y}[t \wedge \tau] &\le C t^{-d/2}(1+t^2)|x-z|^{1-\varepsilon}, \\ p^X(0, z; t, y) E^{Z_t^z = y}[t \wedge \tau] &\le C t^{-d/2}(1+t^2)|x-z|^{1-\varepsilon} \end{aligned}$$

*for $t \in (0, \infty)$, $x, z \in B(0; R/2)$ such that $|x-z| \le c_0$, and $y \in \mathbb{R}^d$.*

*Proof.* It holds that

$$E^{X_t^x = y}[t \wedge \tau] = E^{X_t^x = y}[(t\wedge\tau)\mathbb{1}_{[0,t/2]}(\tau)] + E^{X_t^x = y}[(t\wedge\tau)\mathbb{1}_{[t/2,\infty)}(\tau)]. \tag{4-15}$$

By (2-11) and (1-6), we have

$$p^X(0, x; t, y) E^{X_t^x = y}[(t\wedge\tau)\mathbb{1}_{[0,t/2]}(\tau)] = E\left[(t\wedge\tau)\mathbb{1}_{[0,t/2]}(\tau)\; p^X\left(\frac{t}{2}, X_{t/2}^x; t, y\right)\right] \le 2^{d/2} C_{\mathrm{G}}^+ t^{-d/2} E[t\wedge\tau].$$

Hence, in view of Lemma 4.2, there exists positive constants $C$ and $c_0$ depending on $d$, $\varepsilon$, $C_{\mathrm{G}}^+$, $R$, $\rho_R$ and $\Lambda$ such that

$$p^X(0, x; t, y) E^{X_t^x = y}[(t\wedge\tau)\mathbb{1}_{[0,t/2]}(\tau)] \le C t^{-d/2}(1+t^2)|x-z|^{1-\varepsilon} \tag{4-16}$$

for $x, z \in B(0; R/2)$ such that $|x-z| \le c_0$ and $y \in \mathbb{R}^d$.

On the other hand, by (2-11) and (1-6), we have

$$\begin{aligned} p^X(0, x; t, y) E^{X_t^x = y}[(t\wedge\tau)\mathbb{1}_{[t/2,\infty)}(\tau)] &\le t p^X(0, x; t, y) P^{X_t^x = y}\left(\tau > \frac{t}{2}\right) \\ &= t \int_{\mathbb{R}^d} p^X\left(\frac{t}{2}, z; t, y\right) P\left(\tau > \frac{t}{2},\; X_{t/2}^x \in dz\right) \\ &\le 2^{d/2} C_{\mathrm{G}}^+ t^{-d/2+1} P\left(\tau > \frac{t}{2}\right). \end{aligned}$$

Hence, by applying Chebyshev's inequality we have

$$p^X(0,x;t,y)E^{X_t^x=y}[(t\wedge\tau)\mathbb{1}_{[t/2,\infty)}(\tau)]\le Ct^{-d/2}E[t\wedge\tau],$$

where $C$ is a constant depending on $d$ and $C_{\mathrm{G}}^+$. Thus, Lemma 4.2 implies that

$$p^X(0,x;t,y)E^{X_t^x=y}[(t\wedge\tau)\mathbb{1}_{[t/2,\infty)}(\tau)]\le Ct^{-d/2}(1+t^2)|x-z|^{1-\varepsilon} \tag{4-17}$$

for $x, z\in B(0;R/2)$ such that $|x-z|\le c_0$, where $C$ and $c_0$ are positive constants depending on $d$, $\varepsilon$, $C_{\mathrm{G}}^+$, $R$, $\rho_R$ and $\Lambda$. Therefore, we obtain the assertion for $x$ by (4-15), (4-16) and (4-17). Similar argument yields the assertion for $z$. □

**Lemma 4.4.** *For $q\ge 1$, $R>0$ and sufficiently small $\varepsilon>0$, there exist positive constants $C$ and $c_0$ depending on $q$, $d$, $\varepsilon$, $R$, $\rho_R$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$, such that*

$$E\Big[\sup_{s\in[0,t]}|\mathscr{E}(0,\tau\wedge s;X^x)-1|^q\Big]\le Ce^{Ct}|x-z|^{2/(q\vee 2)-\varepsilon},$$
$$E\Big[\sup_{s\in[0,t]}|\mathscr{E}(0,\tau\wedge s;Z^z)-1|^q\Big]\le Ce^{Ct}|x-z|^{2/(q\vee 2)-\varepsilon}$$

*for $t\in[0,\infty)$, $x,z\in B(0;R/2)$ such that $|x-z|\le c_0$, and $y\in\mathbb{R}^d$.*

*Proof.* By (2-10) we have

$$\begin{aligned}
&E\Big[\sup_{v\in[0,\tau\wedge t]}|\mathscr{E}(0,v;X^x)-1|^q\Big]\\
&\quad=E\Big[\sup_{v\in[0,\tau\wedge t]}\Big|\int_0^v\mathscr{E}(0,u;X^x)\langle b_\sigma(u,X_u^x),\,dB_u\rangle+\int_0^v\mathscr{E}(0,u;X^x)c(u,X_u^x)\,du\Big|^q\Big]\\
&\quad\le CE\Big[\sup_{v\in[0,\tau\wedge t]}\Big|\int_0^v\mathscr{E}(0,u;X^x)\langle b_\sigma(u,X_u^x),\,dB_u\rangle\Big|^q\Big]+CE\Big[\sup_{v\in[0,\tau\wedge t]}\Big|\int_0^v\mathscr{E}(0,u;X^x)c(u,X_u^x)\,du\Big|^q\Big],
\end{aligned}$$

where $C$ is a constant depending on $q$. The terms of the right-hand side of this inequality are dominated as follows: By Burkholder's inequality and Hölder's inequality we have

$$\begin{aligned}
&E\Big[\sup_{v\in[0,\tau\wedge t]}\Big|\int_0^v\mathscr{E}(0,u;X^x)\langle b_\sigma(u,X_u^x),\,dB_u\rangle\Big|^q\Big]\\
&\qquad\le CE\Big[\Big(\int_0^{\tau\wedge t}\mathscr{E}(0,u;X^x)^2|b_\sigma(u,X_u^x)|^2\,du\Big)^{q/2}\Big]\\
&\qquad\le C\Lambda^q\|b\|_\infty^q t^{1-1/[(q/2)\vee 1]}E\Big[\int_0^{\tau\wedge t}\mathscr{E}(0,u;X^x)^{2[(q/2)\vee 1]}\,du\Big]^{1/[(q/2)\vee 1]}\\
&\qquad\le C\Lambda^q\|b\|_\infty^q t^{1-2/[q\vee 2]}E[\tau\wedge t]^{2(1-\varepsilon)/(q\vee 2)}E\Big[\Big(\int_0^t\mathscr{E}(0,u;X^x)^{(q\vee 2)/\varepsilon}\,du\Big)\Big]^{2\varepsilon/(q\vee 2)},
\end{aligned}$$

where $C$ is a constant depending on $q$, and by Hölder's inequality we have

$$E\left[\sup_{v\in[0,\tau\wedge t]}\left|\int_0^v \mathscr{E}(0,u;X^x)c(u,X_u^x)\,du\right|^q\right]\le \|c\|_\infty^q t^{1-1/q}E\left[\int_0^{\tau\wedge t}\mathscr{E}(0,u;X^x)^q\,du\right]$$
$$\le \|c\|_\infty^q t^{1-1/q}E[\tau\wedge t]^{1-\varepsilon}E\left[\int_0^t \mathscr{E}(0,u;X^x)^{q/\varepsilon}\,du\right]^\varepsilon.$$

Thus, applying by Lemmas 3.3 and 4.2 to these inequalities and choosing another small $\varepsilon$, we obtain

$$E\left[\sup_{v\in[0,\tau\wedge t]}|\mathscr{E}(0,v;X^x)-1|^q\right]\le Ce^{Ct}|x-z|^{2/(q\vee 2)-\varepsilon},$$

where $C$ is a constant depending on $q$, $d$, $\varepsilon$, $R$, $\rho_R$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$. Similar argument yields the same estimate for $Z^z$. □

Now we start the proof of Proposition 4.1. Let $t\in(0,\infty)$, $x, z\in B(0;R/2)$ such that $x\neq z$, $y\in\mathbb{R}^d$ and $s\in(t/2,t)$. Recall that $X^z$ and $Z^z$ have the same law. By (2-11) and (2-9) we have

$$\left|p^X(0,x;t,y)E^{X_t^x=y}\left[\mathscr{E}(0,s;X^x);\ \tau\le\frac{t}{2}\right]-p^X(0,z;t,y)E^{Z_t^z=y}\left[\mathscr{E}(0,s;Z^z);\ \tau\le\frac{t}{2}\right]\right|$$
$$=\left|E\left[\mathscr{E}(0,s;X^x)p^X(s,X_s^x;t,y);\ \tau\le\frac{t}{2}\right]-E\left[\mathscr{E}(0,s;Z^z)p^X(s,Z_s^z;t,y);\ \tau\le\frac{t}{2}\right]\right|$$
$$\le E\left[\mathscr{E}(0,s;Z^z)|p^X(s,X_s^x;t,y)-p^X(s,Z_s^z;t,y)|;\ \tau\le\frac{t}{2}\right]$$
$$+E\left[|\mathscr{E}(0,\tau\wedge s;X^x)-\mathscr{E}(0,\tau\wedge s;Z^z)|\mathscr{E}(\tau\wedge s,s;Z^z)p^X(s,X_s^x;t,y);\ \tau\le\frac{t}{2}\right]$$
$$+E\left[\mathscr{E}(0,\tau\wedge s;X^x)|\mathscr{E}(\tau\wedge s,s;X^x)-\mathscr{E}(\tau\wedge s,s;Z^z)|p^X(s,X_s^x;t,y);\ \tau\le\frac{t}{2}\right]$$

Noting that

$$X_s^x=Z_s^z\quad\text{for } s\ge\tau,$$

we obtain

$$\left|p^X(0,x;t,y)E^{X_t^x=y}\left[\mathscr{E}(0,s;X^x);\ \tau\le\frac{t}{2}\right]-p^X(0,z;t,y)E^{Z_t^z=y}\left[\mathscr{E}(0,s;Z^z);\ \tau\le\frac{t}{2}\right]\right|$$
$$\le E\left[|\mathscr{E}(0,\tau\wedge s;X^x)-\mathscr{E}(0,\tau\wedge s;Z^z)|\mathscr{E}(\tau\wedge s,s;Z^z)p^X(s,X_s^x;t,y);\ \tau\le\frac{t}{2}\right].\quad(4\text{-}18)$$

By the triangle inequality and Hölder's inequality, we obtain

$$
\begin{aligned}
&E\Big[|\mathscr{E}(0,\tau\wedge s;X^x)-\mathscr{E}(0,\tau\wedge s;Z^z)|\mathscr{E}(\tau\wedge s,s;Z^z)p^X(s,X^x_s;t,y);\ \tau\le\frac{t}{2}\Big]\\
&\qquad\le E\Big[|\mathscr{E}(0,\tau\wedge s;X^x)-1|\mathscr{E}(\tau\wedge s,s;Z^z)p^X(s,X^x_s;t,y);\ \tau\le\frac{t}{2}\Big]\\
&\qquad\quad+E\Big[|\mathscr{E}(0,\tau\wedge s;Z^z)-1|\mathscr{E}(\tau\wedge s,s;Z^z)p^X(s,X^x_s;t,y);\ \tau\le\frac{t}{2}\Big]\\
&\qquad\le\bigg(E\Big[|\mathscr{E}(0,\tau\wedge s;X^x)-1|^{2/(2-\varepsilon)}p^X(s,X^x_s;t,y);\ \tau\le\frac{t}{2}\Big]^{1-\varepsilon/2}\\
&\qquad\qquad+E\Big[|\mathscr{E}(0,\tau\wedge s;Z^z)-1|^{2/(2-\varepsilon)}p^X(s,X^x_s;t,y);\ \tau\le\frac{t}{2}\Big]^{1-\varepsilon/2}\bigg)\\
&\qquad\quad\times E\Big[\mathscr{E}(\tau\wedge s,s;Z^z)^{2/\varepsilon}p^X(s,X^x_s;t,y);\ \tau\le\frac{t}{2}\Big]^{\varepsilon/2}.
\end{aligned}
$$

Hence, by (2-11) and (1-6), we have

$$
\begin{aligned}
&E\Big[\mathscr{E}(0,\tau\wedge s;X^x)-\mathscr{E}(0,\tau\wedge s;Z^z)|\mathscr{E}(\tau\wedge s,s;Z^z)p^X(s,X^x_s;t,y);\ \tau\le\frac{t}{2}\Big]\\
&\qquad\le\bigg(E^{X^x_t=y}\Big[|\mathscr{E}(0,\tau\wedge s;X^x)-1|^{2/(2-\varepsilon)};\ \tau\le\frac{t}{2}\Big]^{1-\varepsilon/2}\\
&\qquad\qquad+E^{X^x_t=y}\Big[|\mathscr{E}(0,\tau\wedge s;Z^z)-1|^{2/(2-\varepsilon)};\ \tau\le\frac{t}{2}\Big]^{1-\varepsilon/2}\bigg)\\
&\qquad\quad\times p^X(0,x;t,y)E^{X^x_t=y}\Big[\mathscr{E}(\tau\wedge s,s;Z^z)^{2/\varepsilon};\ \tau\le\frac{t}{2}\Big]^{\varepsilon/2}\\
&\qquad\le\bigg(E\Big[|\mathscr{E}(0,\tau\wedge s;X^x)-1|^{2/(2-\varepsilon)}p^X(t/2,X^x_{t/2};t,y);\ \tau\le\frac{t}{2}\Big]^{1-\varepsilon/2}\\
&\qquad\qquad+E\Big[|\mathscr{E}(0,\tau\wedge s;Z^z)-1|^{2/(2-\varepsilon)}p^X(t/2,X^x_{t/2};t,y);\ \tau\le\frac{t}{2}\Big]^{1-\varepsilon/2}\bigg)\\
&\qquad\quad\times\bigg(p^X(0,x;t,y)E^{X^x_t=y}\Big[\mathscr{E}(\tau\wedge s,s;Z^z)^{2/\varepsilon};\ \tau\le\frac{t}{2}\Big]\bigg)^{\varepsilon/2}\\
&\qquad\le(C^+_{\mathrm{G}})^{1-\varepsilon/2}t^{-d/2+d\varepsilon/4}(p^X(0,x;t,y)E^{X^x_t=y}[\mathscr{E}(\tau\wedge s,s;Z^z)^{2/\varepsilon}])^{\varepsilon/2}\\
&\qquad\quad\times(E[|\mathscr{E}(0,\tau\wedge s;X^x)-1|^{2/(2-\varepsilon)}]^{1-\varepsilon/2}+E[|\mathscr{E}(0,\tau\wedge s;Z^z)-1|^{2/(2-\varepsilon)}]^{1-\varepsilon/2}).
\end{aligned}
$$

Applying Lemmas 3.3 and 4.4 to this inequality, we obtain

$$
E\Big[|\mathscr{E}(0,\tau\wedge s;X^x)-\mathscr{E}(0,\tau\wedge s;Z^z)|\mathscr{E}(\tau\wedge s,s;Z^z)p^X(s,X^x_s;t,y);\ \tau\le\frac{t}{2}\Big]\le Ct^{-d/2}e^{Ct}|x-z|^{1-\varepsilon}\quad(4\text{-}19)
$$

for $x,z\in B(0;R/2)$ such that $|x-z|\le c_0$, where $C$ and $c_0$ are constants depending on $d$, $\varepsilon$, $C^+_{\mathrm{G}}$, $R$, $\rho_R$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$.

Hölder's inequality and Chebyshev's inequality imply

$$E^{X_t^x=y}\left[\mathscr{E}(0,s;X^x);\ \tau\ge\frac{t}{2}\right]\le P^{X_t^x=y}\left(\tau\ge\frac{t}{2}\right)^{1-\varepsilon/2}E^{X_t^x=y}[\mathscr{E}(0,s;X^x)^{2/\varepsilon}]^{\varepsilon/2}$$
$$\le\frac{2^{1-\varepsilon/2}}{t^{1-\varepsilon/2}}E^{X_t^x=y}[\tau\wedge t]^{1-\varepsilon/2}E^{X_t^x=y}[\mathscr{E}(0,s;X^x)^{2/\varepsilon}]^{\varepsilon/2}.$$

Hence, by Lemmas 3.5 and 4.3, we obtain

$$p^X(0,x;t,y)E^{X_t^x=y}\left[\mathscr{E}(0,s;X^x);\ \tau\ge\frac{t}{2}\right]\le Ct^{-d/2-1}e^{Ct}|x-z|^{1-\varepsilon}\tag{4-20}$$

for $x, z\in B(0;R/2)$ such that $|x-z|\le c_0$, where $C$ and $c_0$ are constants depending on $d$, $\varepsilon$, $m$, $M$, $\theta$, $R$, $\rho_R$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$. Similarly we have

$$p^X(0,z;t,y)E^{Z_t^z=y}\left[\mathscr{E}(0,s;Z^z);\ \tau\ge\frac{t}{2}\right]\le Ct^{-d/2-1}e^{Ct}|x-z|^{1-\varepsilon}\tag{4-21}$$

for $x, z\in B(0;R/2)$ such that $|x-z|\le c_0$, where $C$ and $c_0$ are constants depending on $d$, $\varepsilon$,$\gamma_G^-$, $\gamma_G^+$, $C_G^-$, $C_G^+$, $m$, $M$, $\theta$, $R$, $\rho_R$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$. Thus, (2-8), (4-18), (4-19), (4-20) and (4-21) imply

$$|p(0,x;t,y)-p(0,z;t,y)|\le Ct^{-d/2-1+\varepsilon/2}e^{Ct}|x-z|^{1-\varepsilon}$$

for $t\in(0,\infty)$, $x, z\in B(0;R/2)$ such that $|x-z|\le c_0$, and $y\in\mathbb{R}^d$, with constants $C$ and $c_0$ depending on $d$, $\varepsilon$, $\gamma_G^-$, $\gamma_G^+$, $C_G^-$, $C_G^+$, $m$, $M$, $\theta$, $R$, $\rho_R$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$. By (1-6) we can remove the restriction on $|x-z|$, and, therefore, we obtain Proposition 4.1.

## 5. The case of general $a$ (proof of the main theorem)

Let $a^{(n)}(t,x)=(a_{ij}^{(n)}(t,x))$ be symmetric $d\times d$-matrix-valued bounded measurable functions on $[0,\infty)\times\mathbb{R}^d$ which converge to $a(t,x)$ for each $(t,x)\in[0,\infty)\times\mathbb{R}^d$ and satisfy (1-1), (1-3) and (1-4). Consider the parabolic partial differential equation

$$\begin{cases}\dfrac{\partial}{\partial t}u(t,x)=\dfrac{1}{2}\displaystyle\sum_{i,j=1}^d a_{ij}^{(n)}(t,x)\frac{\partial^2}{\partial x_i\partial x_j}u(t,x)+\sum_{i=1}^d b_i(t,x)\frac{\partial}{\partial x_i}u(t,x)+c(t,x)u(t,x),\\ u(0,x)=f(x).\end{cases}\tag{5-1}$$

Denote the fundamental solution to (5-1) by $p^{(n)}(s,x;t,y)$. From (1-6) and Proposition 3.6 we have positive constants $\gamma_1$, $\gamma_2$, $C_1$ and $C_2$ depending on $d$, $\gamma_G^-$, $\gamma_G^+$, $C_G^-$, $C_G^+$, $m$, $M$, $\theta$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$ such that

$$\frac{C_1e^{-C_1(t-s)}}{(t-s)^{\frac{d}{2}}}\exp\left(-\frac{\gamma_1|x-y|^2}{t-s}\right)\le p^{(n)}(s,x;t,y)\le\frac{C_2e^{C_2(t-s)}}{(t-s)^{\frac{d}{2}}}\exp\left(-\frac{\gamma_2|x-y|^2}{t-s}\right)\tag{5-2}$$

for $s, t\in[0,\infty)$ such that $s<t$, $x, y\in\mathbb{R}^d$ and $n\in\mathbb{N}$.

It is known that local Hölder continuity of the fundamental solution follows, with index and constant depending only on the constants appearing in the Gaussian estimate; see [Stroock 1988]. This fact and (5-2) imply that the Arzelà–Ascoli theorem is applicable to $p^{(n)}$. Moreover, in view of Proposition 4.1, there exists a constant $C$ depending on $d$, $\varepsilon$, $\gamma_G^-$, $\gamma_G^+$, $C_G^-$, $C_G^+$, $m$, $M$, $\theta$, $R$, $\rho_R$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$ such that

$$|p^{(n)}(0,x;t,y)-p^{(n)}(0,z;t,y)|\le Ct^{-d/2-1}e^{Ct}|x-z|^{1-\varepsilon}$$

for $t\in(0,\infty)$, $y\in\mathbb{R}^d$ and $x,z\in B(0;R/2)$. Hence, there exists a continuous function $p^{(\infty)}(0,\cdot\,;\,\cdot,\cdot)$ on $\mathbb{R}^d\times(0,\infty)\times\mathbb{R}^d$ such that

$$\lim_{n\to\infty}\sup_{|x|\le R/2}|p^{(n)}(0,x;t,y)-p^{(\infty)}(0,x;t,y)|=0, \tag{5-3}$$

$$|p^{(\infty)}(0,x;t,y)-p^{(\infty)}(0,z;t,y)|\le Ct^{-d/2-1}e^{Ct}|x-z|^{1-\varepsilon},\quad x,z\in B(0;R/2), \tag{5-4}$$

for $t\in(0,\infty)$ and $y\in\mathbb{R}^d$, where $C$ is a constant depending on $d$, $\varepsilon$, $\gamma_G^-$, $\gamma_G^+$, $C_G^-$, $C_G^+$, $m$, $M$, $\theta$, $R$, $\rho_R$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$. Moreover, we have positive constants $C_1$, $C_2$, $\gamma_1$ and $\gamma_2$ depending on $d$, $\gamma_G^-$, $\gamma_G^+$, $C_G^-$, $C_G^+$, $m$, $M$, $\theta$, $\Lambda$, $\|b\|_\infty$ and $\|c\|_\infty$ such that

$$\frac{C_1e^{-C_1(t-s)}}{(t-s)^{\frac d2}}\exp\left(-\frac{\gamma_1|x-y|^2}{t-s}\right)\le p^{(\infty)}(s,x;t,y)\le\frac{C_2e^{C_2(t-s)}}{(t-s)^{\frac d2}}\exp\left(-\frac{\gamma_2|x-y|^2}{t-s}\right)$$

for $s,t\in[0,\infty)$ such that $s<t$, and $x,y\in\mathbb{R}^d$. To prove Theorem 1.1, we show that $p^{(\infty)}(0,\cdot\,;\,\cdot,\cdot)$ coincides with the fundamental solution $p(0,\cdot\,;\,\cdot,\cdot)$ of the original parabolic partial differential equation (1-2). Let $\phi,\psi\in C_0^\infty(\mathbb{R}^d)$, and set

$$P_t^{(n)}g(x):=\int_{\mathbb{R}^d}g(y)p^{(n)}(0,x;t,y)\,dy\quad\text{for } g\in C_b(\mathbb{R}^d),$$

$$L_t^{(n)}:=\frac12\sum_{i,j=1}^d a_{ij}^{(n)}(t,x)\frac{\partial^2}{\partial x_i\partial x_j}+\sum_{i=1}^d b_i(t,x)\frac{\partial}{\partial x_i}+c(t,x).$$

Noting that $p^{(n)}(s,x;t,y)$ is smooth in $(s,x,t,y)$, we have $P_t^{(n)}L_t^{(n)}\phi(x)=(\partial/\partial t)P_t^{(n)}\phi(x)$. Hence,

$$\begin{aligned}
&\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}\phi(y)p^{(n)}(0,x;t,y)\,dy\right)\psi(x)\,dx-\int_{\mathbb{R}^d}\phi(x)\psi(x)\,dx\\
&=\int_{\mathbb{R}^d}[P_t^{(n)}\phi(x)]\psi(x)\,dx-\int_{\mathbb{R}^d}\phi(x)\psi(x)\,dx\\
&=\int_0^t\int_{\mathbb{R}^d}\left[\frac{\partial}{\partial s}P_s^{(n)}\phi(x)\right]\psi(x)\,dx\,ds\\
&=\int_0^t\int_{\mathbb{R}^d}[P_s^{(n)}L_s^{(n)}\phi(x)]\psi(x)\,dx\,ds\\
&=\int_0^t\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}\left[\frac12\sum_{i,j=1}^d a_{ij}^{(n)}(s,y)\frac{\partial^2}{\partial y_i\partial y_j}+\sum_{i=1}^d b_i(s,y)\frac{\partial}{\partial y_i}+c(s,y)\right]\phi(y)p^{(n)}(0,x;s,y)\,dy\right)\psi(x)\,dx\,ds.
\end{aligned}$$

Taking the limit as $n \to \infty$, we obtain

$$\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}\phi(y)p^{(\infty)}(0,x;t,y)\,dy\right)\psi(x)\,dx-\int_{\mathbb{R}^d}\phi(x)\psi(x)\,dx$$
$$=\int_0^t\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}\left[\frac{1}{2}\sum_{i,j=1}^d a_{ij}(s,y)\frac{\partial^2}{\partial y_i\partial y_j}+\sum_{i=1}^d b_i(s,y)\frac{\partial}{\partial y_i}+c(s,y)\right]\phi(y)p^{(\infty)}(0,x;s,y)\,dy\right)\psi(x)\,dx\,ds.$$

This equality implies that $p^{(\infty)}(0,x;t,y)$ is also the fundamental solution to the parabolic partial differential equation (1-2). Since the weak solution to (1-2) is unique, $p^{(\infty)}(0,x;t,y)$ coincides with $p(0,x;t,y)$. Therefore, we obtain Theorem 1.1.

## Acknowledgments

The author is grateful to Professor Felix Otto for finding a mistake and giving the author the information on the Schauder estimate. The author also thank to the anonymous referee for careful reading, indicating mistakes in the previous version and giving the author information on references. This work was supported by JSPS KAKENHI grant number 25800054.

## References

[Aronson 1967] D. G. Aronson, “Bounds for the fundamental solution of a parabolic equation”, *Bull. Amer. Math. Soc.* **73** (1967), 890–896. MR 36 #534 Zbl 0153.42002

[Bogachev et al. 2005] V. I. Bogachev, M. Röckner, and S. V. Shaposhnikov, “Глобальная регулярность и оценки решений параболических уравнений”, *Teor. Veroyatn. Primen.* **50**:4 (2005), 652–674. Translated as “Global regularity and bounds for solutions of parabolic equations for probability measures” in *Theory Probab. Appl.* **50**:4 (2006), 561–581. MR 2008f:60073 Zbl 1203.60095

[Bogachev et al. 2009] V. I. Bogachev, N. V. Krylov, and M. Röckner, “Эллиптические и параболические уравнения для мер”, *Uspekhi Mat. Nauk* **64**:6(390) (2009), 5–116. Translated as “Elliptic and parabolic equations for measures” in *Russian Math. Surveys* **64**:6 (2009), 973–1078. MR 2011c:35592 Zbl 1194.35481

[Chen and Kumagai 2003] Z.-Q. Chen and T. Kumagai, “Heat kernel estimates for stable-like processes on $d$-sets”, *Stochastic Process. Appl.* **108**:1 (2003), 27–62. MR 2005d:60135 Zbl 1075.60556

[Cranston 1991] M. Cranston, “Gradient estimates on manifolds using coupling”, *J. Funct. Anal.* **99**:1 (1991), 110–124. MR 93a:58175 Zbl 0770.58038

[Escauriaza 2000] L. Escauriaza, “Bounds for the fundamental solution of elliptic and parabolic equations in nondivergence form”, *Comm. Partial Differential Equations* **25**:5-6 (2000), 821–845. MR 2001i:35009 Zbl 0946.35004

[Fabes and Kenig 1981] E. B. Fabes and C. E. Kenig, “Examples of singular parabolic measures and singular transition probability densities”, *Duke Math. J.* **48**:4 (1981), 845–856. MR 86j:35081 Zbl 0482.35021

[Friedman 1964] A. Friedman, *Partial differential equations of parabolic type*, Prentice-Hall, Englewood Cliffs, NJ, 1964. MR 31 #6062 Zbl 0144.34903

[Giorgi 1957] E. De Giorgi, “Sulla differenziabilità e l’analiticità delle estremali degli integrali multipli regolari”, *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat.* (3) **3** (1957), 25–43. Reprinted, pp. 167–184 in *Selected papers*, edited by G. Dal Maso et al., Springer, New York, 2006; translated as “On the differentiability and the analyticity of extremals of regular multiple integrals”, ibid., 149–166. MR 20 #172 Zbl 0084.31901

[Ikeda and Watanabe 1989] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, 2nd ed., North-Holland Mathematical Library **24**, North-Holland, Amsterdam, 1989. MR 90m:60069 Zbl 0684.60040

[Karatzas and Shreve 1991] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, 2nd ed., Graduate Texts in Mathematics **113**, Springer, New York, 1991. MR 92h:60127 Zbl 0734.60060

[Karrmann 2001] S. Karrmann, "Gaussian estimates for second-order operators with unbounded coefficients", *J. Math. Anal. Appl.* **258**:1 (2001), 320–348. MR 2002d:35085 Zbl 0983.35053

[Krylov 1996] N. V. Krylov, *Lectures on elliptic and parabolic equations in Hölder spaces*, Graduate Studies in Mathematics **12**, American Mathematical Society, Providence, RI, 1996. MR 97i:35001 Zbl 0865.35001

[Kusuoka and Stroock 1985] S. Kusuoka and D. W. Stroock, "Applications of the Malliavin calculus, II", *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **32**:1 (1985), 1–76. MR 86k:60100b Zbl 0568.60059

[Ladyženskaja et al. 1967] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'ceva, *Линейные и квазилинейные уравнения параболического типа*, Nauka, Moscow, 1967. Translated as *Linear and quasilinear equations of parabolic type*, Transl. Math. Monogr. **23**, American Mathematical Society, Providence, RI, 1968. MR 39 #3159b Zbl 0164.12302

[Lax and Milgram 1954] P. D. Lax and A. N. Milgram, "Parabolic equations", pp. 167–190 in *Contributions to the theory of partial differential equations*, edited by L. Bers et al., Annals of Mathematics Studies **33**, Princeton University Press, 1954. MR 16,709b Zbl 0058.08703

[Lindvall and Rogers 1986] T. Lindvall and L. C. G. Rogers, "Coupling of multidimensional diffusions by reflection", *Ann. Probab.* **14**:3 (1986), 860–872. MR 88b:60179 Zbl 0593.60076

[Metafune et al. 2009] G. Metafune, D. Pallara, and A. Rhandi, "Global properties of transition probabilities of singular diffusions", *Teor. Veroyatn. Primen.* **54**:1 (2009), 116–148. MR 2011m:35144 Zbl 1206.60072

[Nash 1958] J. Nash, "Continuity of solutions of parabolic and elliptic equations", *Amer. J. Math.* **80** (1958), 931–954. MR 20 #6592 Zbl 0096.06902

[Porper and Eidelman 1984] F. O. Porper and S. D. Eidelman, "Двусторонние оценки фундаментальных решений параболических уравнений второго порядка и некоторые их приложения", *Uspekhi Mat. Nauk* **39**:3(237) (1984), 107–156. Translated as "Two-sided estimates of fundamental solutions of second-order parabolic equations, and some applications" in *Russian Math. Surveys* **39**:3 (1984), 119–178. MR 86b:35078 Zbl 0582.35052

[Porper and Eidelman 1992] F. O. Porper and S. D. Eidelman, "Свойства решений параболических уравнений второго порядка с младшими членами", *Trudy Moskov. Mat. Obshch.* **54** (1992), 118–159. Translated as "Properties of solutions of second-order parabolic equations with lower-order terms" in *Trans. Moscow Math. Soc.* **54** (1993), 101–137. MR 95b:35084 Zbl 0783.35023

[Revuz and Yor 1999] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, 3rd ed., Grundlehren der Mathematischen Wissenschaften **293**, Springer, Berlin, 1999. MR 2000h:60050 Zbl 0917.60006

[Stroock 1988] D. W. Stroock, "Diffusion semigroups corresponding to uniformly elliptic divergence form operators", pp. 316–347 in *Séminaire de Probabilités XXII*, edited by J. Azéma et al., Lecture Notes in Math. **1321**, Springer, Berlin, 1988. MR 90b:35071 Zbl 0651.47031

[Stroock and Varadhan 1979] D. W. Stroock and S. R. S. Varadhan, *Multidimensional diffusion processes*, Grundlehren der Mathematischen Wissenschaften **233**, Springer, Berlin, 1979. MR 81f:60108 Zbl 0426.60069

Received 16 Oct 2013. Revised 6 Nov 2014. Accepted 21 Dec 2014.

SEIICHIRO KUSUOKA: kusuoka@math.tohoku.ac.jp
*Graduate School of Science, Tohoku University, 6-3 Aramaki Aza-Aoba, Aoba-ku, Sendai 980-8578, Japan*

ANALYSIS AND PDE
Vol. 8, No. 1, 2015
dx.doi.org/10.2140/apde.2015.8.33

# EIGENVALUE DISTRIBUTION OF OPTIMAL TRANSPORTATION

BO'AZ B. KLARTAG AND ALEXANDER V. KOLESNIKOV

We investigate the Brenier map $\nabla\Phi$ between the uniform measures on two convex domains in $\mathbb{R}^n$, or, more generally, between two log-concave probability measures on $\mathbb{R}^n$. We show that the eigenvalues of the Hessian matrix $D^2\Phi$ exhibit concentration properties on a multiplicative scale, regardless of the choice of the two measures or the dimension $n$.

## 1. Introduction

Let $\mu$ and $\nu$ be two absolutely continuous probability measures on $\mathbb{R}^n$. It was discovered by Brenier [1991] and McCann [1995] that there exists a convex function $\Phi$ on $\mathbb{R}^n$ with $(\nabla\Phi)_*\mu = \nu$, i.e.,

$$\int_{\mathbb{R}^n} b(\nabla\Phi(x))\, d\mu(x) = \int_{\mathbb{R}^n} b(x)\, d\nu(x) \tag{1}$$

for any $\nu$-integrable function $b : \mathbb{R}^n \to \mathbb{R}$. Moreover, the Brenier map $x \mapsto \nabla\Phi(x)$ is uniquely determined $\mu$-almost everywhere. In this paper we consider the case where $\mu$ and $\nu$ are log-concave probability measures. An absolutely continuous probability measure on $\mathbb{R}^n$ is called log-concave if it has a density $\rho$ which satisfies

$$\rho(\lambda x + (1-\lambda)y) \geq \rho(x)^\lambda \rho(y)^{1-\lambda} \quad (x, y \in \mathbb{R}^n,\ 0 < \lambda < 1).$$

The uniform measure on any convex domain is log-concave, as is the Gaussian measure. Write $\mathrm{Supp}(\mu)$ for the interior of the support of $\mu$, which is an open, convex set in $\mathbb{R}^n$. We make the assumption:

($\star$) The function $\Phi$ is $C^2$-smooth in $\mathrm{Supp}(\mu)$.

It follows from work of Caffarelli [1990; 1992; 1999] that ($\star$) holds true when each of the measures $\mu$ and $\nu$ satisfies the following additional condition: either the support of the measure is the entire $\mathbb{R}^n$ or else the support is a bounded, convex domain and the density of the measure is bounded away from zero and from infinity in this convex domain. It is fair to say that Caffarelli's regularity theory covers most cases of interest, yet it is very plausible that ($\star$) is in fact always correct, without any additional conditions. For related results on the regularity of optimal transportation, see Delanoë [1991] and Urbas [1997].

As it turns out, the positive-definite Hessian matrix $D^2\Phi(x)$ exhibits remarkable regularity in the behavior of its eigenvalues. We write $\mathrm{Var}[X]$ for the variance of the random variable $X$.

---

*MSC2010:* 35J96.
*Keywords:* transportation of measure, log-concave measures.

**Theorem 1.1.** *Let $\mu$, $\nu$ be absolutely continuous, log-concave probability measures on $\mathbb{R}^n$. Let $\nabla\Phi$ be the Brenier map between $\mu$ and $\nu$, and assume ($\star$). Write $0 < \lambda_1(x) \le \cdots \le \lambda_n(x)$ for the eigenvalues of the matrix $D^2\Phi(x)$, repeated according to their multiplicity. Let $X$ be a random vector in $\mathbb{R}^n$ that is distributed according to $\mu$. Then, for $i = 1, \dots, n$,*

$$\mathrm{Var}[\log\lambda_i(X)] \le 4.$$

Thus, on a multiplicative scale, the eigenvalues of $D^2\Phi$ are quite stable. Note that the multiplicative scale is indeed the natural scale in the generality of Theorem 1.1: by applying appropriate linear transformations to $\mu$ and $\nu$, one may effectively multiply all eigenvalues by an arbitrary positive constant. The variance bound in Theorem 1.1 follows from a Poincaré inequality which we now formulate. For $x \in \mathrm{Supp}(\mu)$ set

$$\Lambda(x) = (\log\lambda_1(x), \dots, \log\lambda_n(x)).$$

We write $|\cdot|$ for the standard Euclidean norm in $\mathbb{R}^n$.

**Theorem 1.2.** *Under the notation and assumptions of Theorem 1.1, for any locally Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$ with $\mathbb{E}|f(\Lambda(X))| < \infty$,*

$$\mathrm{Var}[f(\Lambda(X))] \le 4\mathbb{E}|\nabla f|^2(\Lambda(X))$$

*whenever the right-hand side is finite. At the points at which $f$ is not continuously differentiable, we define $|\nabla f|$ via (36) below.*

Set $\pi = \Lambda_*(\mu)$, the push-forward of the measure $\mu$ under the map $\Lambda$. Theorem 1.2 is a spectral gap estimate for the metric-measure space $(\mathbb{R}^n, |\cdot|, \pi)$. Gromov and Milman [1983] proved that a spectral gap estimate implies exponential concentration of Lipschitz functions. Therefore, Theorem 1.2 admits the following immediate corollary:

**Corollary 1.3.** *We work under the notation and assumptions of Theorem 1.1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a 1-Lipschitz function (i.e., $|f(x) - f(y)| \le |x - y|$).*

*Write $A = \mathbb{E} f(\Lambda(X))$. Then $A$ is finite and*

$$\mathbb{E}\exp\big(c|f(\Lambda(X)) - A|\big) \le 2,$$

*where $c > 0$ is a universal constant.*

**Remark 1.4.** Corollary 1.3 implies that $\mathbb{E}e^{c|\Lambda(X)|} < \infty$. Consequently, one may replace the condition $\mathbb{E}|f(\Lambda(X))| < \infty$ in Theorem 1.2 by the requirement that $e^{-c|x|}|f(x)|$ is bounded in $\mathbb{R}^n$ for a certain universal constant $c > 0$.

Our next result is that the diagonal elements of the matrix $D^2\Phi(x)$ are also concentrated on a logarithmic scale, pretty much like the eigenvalues.

**Theorem 1.5.** *We work under the notation and assumptions of Theorem 1.1. Fix $v \in \mathbb{R}^n$, let $H(x) = \log(D^2\Phi(x)v \cdot v)$ and let $Y = H(X)$. Then:*

(i) $\mathrm{Var}[Y] \le 4$.

(ii) *For any locally Lipschitz function* $f : \mathbb{R} \to \mathbb{R}$ *with* $\mathbb{E}|f(Y)| < \infty$,

$$\mathrm{Var}[f(Y)] \le 4\mathbb{E}|f'|^2(Y).$$

(iii) *For any* 1*-Lipschitz function* $f : \mathbb{R} \to \mathbb{R}$, *denoting* $A = \mathbb{E} f(Y)$ *we have that* $A \in \mathbb{R}$ *and*

$$\mathbb{E}\exp(c|f(Y) - A|) \le 2,$$

*where* $c > 0$ *is a universal constant.*

All of the assertions made so far follow from Theorem 5.1 below, which is in fact a sound reformulation of [Klartag 2013, Theorem 1.4]. The results in [Klartag 2013] were obtained under a technical assumption dubbed "regularity at infinity", which we shall address in this paper. Our argument is based on analysis of the transportation metric; this means that we use the positive-definite Hessian $D^2\Phi$ in order to define a Riemannian metric in $\mathrm{Supp}(\mu)$. The weighted Riemannian manifold

$$M_{\mu,\nu} = (\mathrm{Supp}(\mu),\, D^2\Phi,\, \mu)$$

was studied in [Kolesnikov 2014], where it was shown that the associated Ricci–Bakry–Émery tensor is nonnegative when $\mu$ and $\nu$ are log-concave. We will also consider the map

$$x \mapsto D^2\Phi(x)$$

from $\mathrm{Supp}(\mu) \subseteq \mathbb{R}^n$ into the space of positive-definite matrices. The space of positive-definite matrices is endowed with a natural Riemannian metric, which fits very nicely with computations related to the weighted Riemannian manifold $M_{\mu,\nu}$. This leads to a certain Poincaré inequality with respect to the standard Riemannian metric on the space of positive-definite matrices, formulated in Theorem 5.1 below.

We have tried to make the exposition self-contained, apart from the regularity theory of mass-transport. The rest of this paper is organized as follows: In Section 2 we recall some well-known constructions related to positive-definite matrices. In Section 3 and Section 4 we prove the main results under regularity assumptions by employing the Bakry–Émery $\Gamma_2$-calculus. Section 5 is devoted to the elimination of these regularity assumptions. In Section 6 we complete the proofs of the theorems formulated above. We denote derivatives by $\partial_k f = f_k = \partial f/\partial x_k$ and $f_{ij} = \partial^2 f/(\partial x_i \partial x_j)$. By a smooth function we mean a $C^\infty$-smooth one. We write log for the natural logarithm, $x \cdot y$ stands for the standard scalar product of $x, y \in \mathbb{R}^n$, and $\mathrm{Tr}(A)$ stands for the trace of the matrix $A$.

## 2. Positive-definite quadratic forms

This section surveys standard material on positive-definite matrices. Denote by $M_n^+(\mathbb{R})$ the collection of all symmetric, positive-definite $n \times n$ matrices. For a function $f : (0, \infty) \to \mathbb{R}$ and $A \in M_n^+(\mathbb{R})$ we may define the symmetric matrix $f(A)$ via the spectral theorem. In other words,

$$f\Bigl(\sum_{i=1}^n \lambda_i v_i \otimes v_i\Bigr) = \sum_{i=1}^n f(\lambda_i) v_i \otimes v_i$$

for any orthonormal basis $v_1, \dots, v_n \in \mathbb{R}^n$ and $\lambda_1, \dots, \lambda_n > 0$, where we write $x \otimes x = (x_i x_j)_{i,j=1,\dots,n}$ for $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

**Lemma 2.1.** *For any $A, B \in M_n^+(\mathbb{R})$,*

$$\|\log(A^{1/2} B A^{1/2})\|_{HS} \le \|\log(A)\|_{HS} + \|\log(B)\|_{HS}, \tag{2}$$

*where $\|\cdot\|_{HS}$ stands for the Hilbert–Schmidt norm.*

*Proof.* For an $n \times n$ matrix $T$ and $k = 1, \dots, n$ we define

$$D_k(T) = \sup_{\substack{E \subseteq \mathbb{R}^n \\ \dim(E)=k}} \frac{\mathrm{Vol}_k(T(B^n \cap E))}{\mathrm{Vol}_k(B^n \cap E)}, \tag{3}$$

where $B^n = \{x \in \mathbb{R}^n \mid |x| < 1\}$ and the supremum in (3) runs over all $k$-dimensional subspaces in $\mathbb{R}^n$. Thus, an application of the linear transformation $T$ may increase $k$-dimensional volumes by a factor of at most $D_k(T)$. It follows that, for any $n \times n$ matrices $A$ and $B$,

$$D_k(AB) \le D_k(A) D_k(B) \quad (k = 1, \dots, n). \tag{4}$$

In the case where $A \in M_n^+(\mathbb{R})$, we have $D_k(A) = \prod_{i=1}^k \lambda_i$, where $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n > 0$ are the eigenvalues of $A$. Assume that $A, B \in M_n^+(\mathbb{R})$. Denote the eigenvalues of the symmetric, positive-definite matrix $A^{1/2} B A^{1/2}$ by $e^{\gamma_1} \ge \cdots \ge e^{\gamma_n} > 0$. Then, for $k = 1, \dots, n$,

$$\prod_{i=1}^k e^{\gamma_i} = D_k(A^{1/2} B A^{1/2}) \le D_k(A^{1/2}) D_k(B) D_k(A^{1/2}) = D_k(A) D_k(B) = \prod_{i=1}^k (e^{\alpha_i} e^{\beta_i}), \tag{5}$$

where $e^{\alpha_1} \ge \cdots \ge e^{\alpha_n} > 0$ are the eigenvalues of $A$ and $e^{\beta_1} \ge \cdots \ge e^{\beta_n} > 0$ are the eigenvalues of $B$. We will next apply a lemma of Weyl [1949]; see also [Polya 1950]. According to the inequality of Weyl and Polya, the inequalities (5) entail that

$$\sum_{i=1}^n h(\gamma_i) \le \sum_{i=1}^n h(\alpha_i + \beta_i) \tag{6}$$

for any convex, nondecreasing function $h : \mathbb{R} \to \mathbb{R}$. For $t \in \mathbb{R}$ let $t_+ = \max\{t, 0\}$. The function $t \mapsto (t_+)^2$ is convex and nondecreasing; hence, from (6),

$$\sum_{i=1}^n ((\gamma_i)_+)^2 \le \sum_{i=1}^n ((\alpha_i + \beta_i)_+)^2. \tag{7}$$

By using (4) for the inverse matrices, we conclude that, for $k = 1, \dots, n$,

$$\prod_{i=n-k+1}^n e^{-\gamma_i} = D_k(A^{-1/2} B^{-1} A^{-1/2}) \le D_k(A^{-1}) D_k(B^{-1}) = \prod_{i=n-k+1}^n (e^{-\alpha_i} e^{-\beta_i}).$$

The inequality of Weyl and Polya now implies that $\sum_{i=1}^n h(-\gamma_i) \le \sum_{i=1}^n h(-\alpha_i - \beta_i)$ for any convex, nondecreasing function $h$. By again using $h(t) = (t_+)^2$, we get

$$\sum_{i=1}^n ((-\gamma_i)_+)^2 \le \sum_{i=1}^n ((-\alpha_i - \beta_i)_+)^2. \tag{8}$$

Adding (7) and (8), we finally obtain

$$\sum_{i=1}^n \gamma_i^2 \le \sum_{i=1}^n (\alpha_i + \beta_i)^2 \le \Big(\sqrt{\textstyle\sum_{i=1}^n \alpha_i^2} + \sqrt{\textstyle\sum_{i=1}^n \beta_i^2}\Big)^2, \tag{9}$$

where we used the Cauchy–Schwartz inequality in the last step. By taking the square root of (9) we deduce (2). □

For two matrices $A, B \in M_n^+(\mathbb{R})$, set

$$\mathrm{dist}(A, B) = \|\log(A^{-1/2} B A^{-1/2})\|_{HS}. \tag{10}$$

Equivalently, $\mathrm{dist}(A, B)$ equals

$$\sqrt{\textstyle\sum_i \log^2 \lambda_i},$$

where $\lambda_1, \dots, \lambda_n > 0$ are the eigenvalues of the matrix $A^{-1}B$ which is conjugate to $A^{-1/2}BA^{-1/2}$. The latter equivalent definition of dist shows that, for any invertible $n \times n$ matrix $T$,

$$\mathrm{dist}(A, B) = \mathrm{dist}(T^t A T, T^t B T) \quad (A, B \in M_n^+(\mathbb{R}^n)), \tag{11}$$

where $T^t$ is the transpose of the matrix $T$. Observe too that $\mathrm{dist}(A, B) = \mathrm{dist}(A^{-1}, B^{-1})$ for any $A, B \in M_n^+(\mathbb{R})$. Lemma 2.1 states that, for $A, B \in M_n^+(\mathbb{R})$,

$$\mathrm{dist}(A, B) \le \mathrm{dist}(A, \mathrm{Id}) + \mathrm{dist}(\mathrm{Id}, B), \tag{12}$$

where Id is the identity matrix. From (11) and (12) one realizes that dist satisfies the triangle inequality in $M_n^+(\mathbb{R})$, so it is a metric. For $A \in M_n^+(\mathbb{R}^n)$ and a symmetric $n \times n$ matrix $B$, we define

$$\|B\|_A = \|A^{-1/2} B A^{-1/2}\|_{HS} = \sqrt{\mathrm{Tr}[(A^{-1}B)^2]}.$$

For a smooth curve $\gamma : [a, b] \to M_n^+(\mathbb{R})$, set

$$\mathrm{Length}(\gamma) = \int_a^b \|\dot\gamma(s)\|_{\gamma(s)}\, ds, \tag{13}$$

where $\dot\gamma(s) = d\gamma(s)/ds$ is a symmetric $n \times n$ matrix. Then Length is invariant under conjugations. That is, the length of the curve $\gamma(s)$ equals that of the curve $T^t\gamma(s)T$ for any invertible $n \times n$ matrix $T$.

**Lemma 2.2.** (i) *For any $A \in M_n^+(\mathbb{R}^n)$ and a symmetric $n \times n$ matrix $B$,*

$$\lim_{\varepsilon \to 0} \frac{\mathrm{dist}^2(A + \varepsilon B, A)}{\varepsilon^2} = \|B\|_A^2 = \mathrm{Tr}[(A^{-1}B)^2]. \tag{14}$$

(ii) *Let* $A, B \in M_n^+(\mathbb{R}^n)$ *and consider the curve*

$$\gamma_{A,B}(s) = A^{1/2}(A^{-1/2}BA^{-1/2})^s A^{1/2} \quad (0 \le s \le 1).$$

*Then* $\gamma_{A,B}$ *is a curve connecting* $A$ *and* $B$ *with* $\mathrm{Length}(\gamma_{A,B}) = \mathrm{dist}(A, B)$.

*Proof.* The invariance property (11) implies that

$$\mathrm{dist}(A + \varepsilon B, A) = \mathrm{dist}(\mathrm{Id} + \varepsilon A^{-1/2}BA^{-1/2}, \mathrm{Id}).$$

It therefore suffices to prove (i) under the additional assumption that $A = \mathrm{Id}$. Let $\lambda_1, \dots, \lambda_n > 0$ be the eigenvalues of $B$. It follows from (10) that

$$\lim_{\varepsilon \to 0} \frac{\mathrm{dist}^2(\mathrm{Id} + \varepsilon B, \mathrm{Id})}{\varepsilon^2} = \lim_{\varepsilon \to 0} \frac{\sum_{i=1}^n \log^2(1 + \varepsilon \lambda_i)}{\varepsilon} = \sum_{i=1}^n \lambda_i^2,$$

and (i) follows from the fact that $\|B\|_A^2 = \sum_i \lambda_i^2$. We now turn to the proof of (ii). Again, we may reduce matters to the case where $A = \mathrm{Id}$ by noting that

$$\gamma_{A,B}(s) = A^{1/2}\gamma_{\mathrm{Id}, A^{-1/2}BA^{-1/2}}(s)A^{1/2} \quad (0 \le s \le 1).$$

Abbreviate $\gamma(s) = \gamma_{A,B}(s) = \gamma_{\mathrm{Id},B}(s)$. Since $\gamma(s) = B^s$, we have $\dot{\gamma}(s) = B^s \log(B)$ and hence, for any $0 \le s \le 1$,

$$\|\dot{\gamma}(s)\|_{\gamma(s)} = \|B^{-s/2}(B^s \log(B))B^{-s/2}\|_{HS} = \|\log(B)\|_{HS} = \mathrm{dist}(\mathrm{Id}, B).$$

From the definition (13) it follows that $\mathrm{Length}(\gamma) = \mathrm{dist}(\mathrm{Id}, B)$, and (ii) is proven. □

The right-hand side of (14) depends quadratically on $B$, and therefore Lemma 2.2 tells us that our distance function dist on $M_n^+(\mathbb{R})$ is induced by a Riemannian metric. We refer to this Riemannian metric as the *standard Riemannian metric* on $M_n^+(\mathbb{R})$. The next two lemmas describe certain Lipschitz functions on $M_n^+(\mathbb{R})$.

**Lemma 2.3.** *Fix* $v \in \mathbb{R}^n$ *and set* $f(A) = \log(Av \cdot v)$ *for* $A \in M_n^+(\mathbb{R})$. *Then* $f$ *is a* 1*-Lipschitz function with respect to the standard Riemannian metric on* $M_n^+(\mathbb{R})$.

*Proof.* The map $f$ is clearly smooth. Fix $A \in M_n^+(\mathbb{R})$ and let us show that the norm of the Riemannian gradient of $f$ at the point $A$ is bounded by one. For any symmetric $n \times n$ matrix $B$, we have

$$\frac{d}{dt} f(A + tB)\Big|_{t=0} = \frac{Bv \cdot v}{Av \cdot v}.$$

Thus, in order to prove the lemma, it suffices to show that

$$\frac{Bv \cdot v}{Av \cdot v} \le \|B\|_A = \|A^{-1/2}BA^{-1/2}\|_{HS}. \tag{15}$$

By switching to another orthonormal basis if necessary, we may assume that $A$ is a diagonal matrix. Denote by $\lambda_1, \dots, \lambda_n > 0$ the numbers on the diagonal of $A$. Let $B = (b_{ij})_{i,j=1,\dots,n}$ and $v = (v_1, \dots, v_n) \in \mathbb{R}^n$.

From the Cauchy–Schwartz inequality,

$$\sum_{i,j=1}^{n} b_{ij} v_i v_j \le \sqrt{\textstyle\sum_{i,j=1}^{n} b_{ij}^2/(\lambda_i\lambda_j)}\sqrt{\textstyle\sum_{i,j=1}^{n} \lambda_i\lambda_j v_i^2 v_j^2} = \sqrt{\textstyle\sum_{i,j=1}^{n} b_{ij}^2/(\lambda_i\lambda_j)}\Big(\sum_{i=1}^{n}\lambda_i v_i^2\Big),$$

which is equivalent to the desired inequality (15). □

**Lemma 2.4.** *For $A \in M_n^+(\mathbb{R})$, let its eigenvalues be $\lambda_1(A) \ge \cdots \ge \lambda_n(A) > 0$. The map $\Lambda : M_n^+(\mathbb{R}) \to \mathbb{R}^n$ defined via*

$$\Lambda(A) = \big(\log(\lambda_1(A)), \dots, \log(\lambda_n(A))\big). \tag{16}$$

*is a 1-Lipschitz map with respect to the standard Riemannian metric on $M_n^+(\mathbb{R})$ and the standard Euclidean metric on $\mathbb{R}^n$.*

*Proof.* Let $\mathcal{F} \subseteq M_n^+(\mathbb{R})$ be the collection of all positive-definite, symmetric matrices with $n$ distinct eigenvalues. Then $\mathcal{F}$ is an open, dense set. The function $\Lambda$ is continuous, since the eigenvalues vary continuously with the matrix. It therefore suffices to prove that

$$|\Lambda(A_1) - \Lambda(A_2)| \le \operatorname{dist}(A_1, A_2) \quad \text{for } A_1,\ A_2 \in \mathcal{F}.$$

Fix $A_1, A_2 \in \mathcal{F}$ with $A_1 \ne A_2$. Consider the curve $\gamma(s) = \gamma_{A_1,A_2}(s/\operatorname{dist}(A_1, A_2))$, where $\gamma_{A_1,A_2}(s)$ is as in Lemma 2.2. Then $\gamma$ is a length-minimizing curve between $A_1$ and $A_2$ parametrized by Riemannian arclength. We claim that $\gamma(s) \in \mathcal{F}$ for all but finitely many values of $s$. Indeed, the resultant of the matrix $\gamma(s)$ is a real-analytic function of $s$ which is not identically zero; hence its zeros are isolated. Since $\Lambda \circ \gamma$ is continuous, in order to prove the lemma it suffices to show that

$$\left|\frac{d\Lambda(\gamma(s))}{ds}\right| \le 1 \tag{17}$$

for all $s$ with $\gamma(s) \in \mathcal{F}$. Let us fix $s_0$ with $\gamma(s_0) \in \mathcal{F}$. Let $A = \gamma(s_0)$ and $B = \dot\gamma(s_0)$. Since $\gamma$ is parameterized by arclength,

$$\|B\|_A = \|A^{-1/2} B A^{-1/2}\|_{HS} = 1. \tag{18}$$

Let $v_1, \dots, v_n \in \mathbb{R}^n$ be the orthonormal basis of eigenvectors that corresponds to the eigenvalues $\lambda_1(A), \dots, \lambda_n(A)$ of the matrix $A$. Then,

$$\left.\frac{d\lambda_i(\gamma(s))}{ds}\right|_{s=s_0} = Bv_i \cdot v_i \quad (i = 1, \dots, n). \tag{19}$$

The relation (19) is standard; see, e.g., [Reed and Simon 1978, Section XII.1]. Consequently,

$$\left.\frac{d\Lambda(\gamma(s))}{ds}\right|_{s=s_0} = \left(\frac{Bv_1 \cdot v_1}{\lambda_1(A)}, \dots, \frac{Bv_n \cdot v_n}{\lambda_n(A)}\right). \tag{20}$$

However, by (18),

$$\sum_{i=1}^{n}\left(\frac{Bv_i \cdot v_i}{\lambda_i(A)}\right)^2 = \sum_{i=1}^{n}(A^{-1/2}BA^{-1/2}v_i \cdot v_i)^2 \le \|A^{-1/2}BA^{-1/2}\|_{HS}^2 = 1. \tag{21}$$

Now (17) follows from (20) and (21). □

**Corollary 2.5.** *Whenever $A$ and $B$ are positive-definite $n \times n$ matrices,*

$$\sum_{i=1}^{n} \log^2 \frac{\lambda_i}{\mu_i} \leq \|\log(A^{-1/2} B A^{-1/2})\|_{HS}^2,$$

*where $\lambda_1 \geq \cdots \geq \lambda_n > 0$ are the eigenvalues of $A$ and $\mu_1 \geq \cdots \geq \mu_n > 0$ are the eigenvalues of $B$.*

## 3. Bakry–Émery $\Gamma_2$-calculus

Let $\mu$ and $\nu$ be two absolutely continuous, log-concave probability measures on $\mathbb{R}^n$. Assume that $d\mu = e^{-V(x)}dx$ and $d\nu = e^{-W(x)}dx$ for certain smooth, convex functions $V, W : \mathbb{R}^n \to \mathbb{R}$. Let $\nabla\Phi$ be the Brenier map between $\mu$ and $\nu$. Caffarelli's regularity theory states that $\Phi : \mathbb{R}^n \to \mathbb{R}$ is a smooth, convex function. Therefore (1) implies that the transport equation

$$-V(x) = \log\det D^2\Phi(x) - W(\nabla\Phi(x)) \tag{22}$$

holds everywhere in $\mathbb{R}^n$. In particular, the matrix $D^2\Phi(x) = (\Phi_{ij}(x))_{i,j=1,\dots,n}$ is invertible and hence positive-definite for any $x \in \mathbb{R}^n$. The inverse to $D^2\Phi(x)$ is denoted by $(D^2\Phi(x))^{-1} = (\Phi^{ij}(x))_{i,j=1,\dots,n}$. We use the Einstein summation convention; thus an index that appears twice in an expression, once as a subscript and once as a superscript, is being summed upon. We also use abbreviations such as $\Phi^i_{jk} = \Phi^{i\ell}\Phi_{jk\ell}$ and $\Phi^{ij}_k = \Phi^{i\ell}\Phi^{jm}\Phi_{km\ell}$. Differentiating (22), we obtain

$$V_j(x) = -\Phi^i_{ji}(x) + \sum_{i=1}^{n} \Phi_{ij}(x) W_i(\nabla\Phi(x)) \quad (j = 1, \dots, n,\ x \in \mathbb{R}^n). \tag{23}$$

Following [Kolesnikov 2014], we use the positive-definite matrices $D^2\Phi(x)$ in order to induce a Riemannian metric on $\mathbb{R}^n$ and consider the weighted Riemannian manifold

$$M = M_{\mu,\nu} = (\mathbb{R}^n, D^2\Phi, \mu).$$

See [Grigor'yan 2009] and Bakry, Gentil and Ledoux [Bakry et al. 2014] for background on weighted Riemannian manifolds and the $\Gamma_2$-calculus. For a smooth function $u : \mathbb{R}^n \to \mathbb{R}$ we have $|\nabla_M u|_M^2 = \Phi^{ij} u_i u_j$, where $|\nabla_M u|_M^2$ stands for the square of the Riemannian norm of the Riemannian gradient of $u$. The Dirichlet form associated with the weighted Riemannian manifold $M_{\mu,\nu}$ is defined, for smooth functions $u, v : \mathbb{R}^n \to \mathbb{R}$, via

$$\Gamma(u, v) = \int_{\mathbb{R}^n} \langle \nabla_M u, \nabla_M v\rangle_M \, d\mu = \int_{\mathbb{R}^n} (\Phi^{ij} u_i v_j) \, d\mu$$

whenever the integral converges. The Laplacian associated with the weighted Riemannian manifold $M_{\mu,\nu}$ is defined, for a smooth function $u : \mathbb{R}^n \to \mathbb{R}$, by

$$Lu = \Phi^{ij} u_{ij} - \sum_{j=1}^{n} W_j(\nabla\Phi(x)) u_j = \Phi^{ij} u_{ij} - (\Phi^{ij}_i + \Phi^{ij} V_i) u_j, \tag{24}$$

where the last equality holds in view of (23). Integrating by parts, we verify that

$$-\int_{\mathbb{R}^n}(Lu)v\,d\mu = -\int_{\mathbb{R}^n}(\Phi^{ij}u_{ij} - [\Phi^{ij}_i + \Phi^{ij}V_i]u_j)ve^{-V} = \int_{\mathbb{R}^n}(\Phi^{ij}u_i v_j)\,d\mu = \Gamma(u,v)$$

for any smooth functions $u, v : \mathbb{R}^n \to \mathbb{R}$, one of which is compactly supported. The next step is to consider the carré du champ of $M_{\mu,\nu}$: As in [Bakry and Émery 1985], for a smooth function $u : K \to \mathbb{R}$ we define

$$\Gamma_2(u) = \tfrac{1}{2}L(|\nabla_M u|^2_M) - \langle \nabla_M u, \nabla_M(Lu)\rangle_M = \tfrac{1}{2}L(\Phi^{ij}u_i u_j) - \Phi^{ij}(Lu)_i u_j. \tag{25}$$

**Lemma 3.1.** *For any smooth function $u : \mathbb{R}^n \to \mathbb{R}$, we have the pointwise inequality*

$$\Gamma_2(u) \geq \tfrac{1}{4}\Phi^{ik}_\ell \Phi^{j\ell}_k u_i u_j.$$

Lemma 3.1 is proven in [Klartag 2013] by introducing a Kähler structure and interpreting the left-hand side of (26) below as the Hilbert–Schmidt norm of a certain Hessian operator restricted to a subspace. There are several additional ways to prove Lemma 3.1. The brute force way involves a tedious but straightforward computation which shows that

$$\Gamma_2(u) = \Phi^{kl}\Phi^{ij}u_{ik}u_{j\ell} - \Phi^{ijk}u_{ij}u_k + \tfrac{1}{2}(\Phi^{ik}_\ell\Phi^{j\ell}_k + \Phi^{ik}\Phi^{j\ell}V_{k\ell})u_i u_j + \tfrac{1}{2}\sum_{i,j=1}^n (W_{ij}\circ\nabla\Phi)u_i u_j.$$

This computation is more or less equivalent to reproving Bochner's formula. Then, one proves the pointwise inequality

$$\Phi^{kl}\Phi^{ij}u_{ik}u_{j\ell} - \Phi^{ijk}u_{ij}u_k + \tfrac{1}{4}\Phi^{ik}_\ell\Phi^{j\ell}_k u_i u_j \geq 0 \tag{26}$$

by representing the left-hand side of (26) as the trace of the square of the matrix $B = (b^j_i)_{i,j=1,\dots,n}$, where $b^j_i = \Phi^{jk}u_{ki} - \frac{1}{2}\Phi^{jk}_i u_k$. The product $A = (D^2\Phi)B$ is a symmetric matrix; hence

$$\mathrm{Tr}(B^2) = \mathrm{Tr}\big[((D^2\Phi)^{-1/2}A(D^2\Phi)^{-1/2})^2\big] \geq 0.$$

Lemma 3.1 follows from (26) and from the fact that $D^2V$ and $D^2W$ are positive semidefinite matrices.

Another approach to Lemma 3.1 is to use the notation of Riemannian geometry as in [Kolesnikov 2014] and use the Bochner formula. We first observe that identity (23) in the case $j = 1$ has the simple form

$$L\Phi_1 = -V_1. \tag{27}$$

Differentiating (27) and using $\partial_k(\Phi^{ij}) = -\Phi^{ij}_k$, we obtain

$$L(\Phi_{11}) - \Phi^{jk}_1\Phi_{1jk} - \sum_{j,k=1}^n \Phi_{j1}\Phi_{1k}(W_{jk}\circ\nabla\Phi) = -V_{11}. \tag{28}$$

The Bochner–Lichnerowicz–Weitzenböck formula states that, for any smooth $u : \mathbb{R}^n \to \mathbb{R}$,

$$\Gamma_2(u) = \|D^2_M u\|^2_M + \mathrm{Ric}_M(\nabla_M u, \nabla_M u), \tag{29}$$

where $\|D^2_M u\|^2_M$ is the Hilbert–Schmidt norm of the Riemannian Hessian of $u$ and $\mathrm{Ric}_M$ is the Bakry–Émery–Ricci tensor of the weighted Riemannian manifold $M = M_{\mu,\nu}$. Let us analyze the term in (29) involving the Hessian of $u$. The Christoffel symbols of our Riemannian metric are $\Gamma^k_{ij} = \frac{1}{2}\Phi^k_{ij}$, and therefore $(D^2_M u)_{ij} = u_{ij} - \frac{1}{2}\Phi^k_{ij}u_k$ and

$$\|D^2_M u\|^2_M = \Phi^{ik}\Phi^{jm}\big(u_{ij} - \tfrac{1}{2}\Phi^\ell_{ij}u_\ell\big)\big(u_{mk} - \tfrac{1}{2}\Phi^s_{mk}u_s\big).$$

In the particular case where $u = \Phi_1$, we obtain $(D^2_M\Phi_1)_{jk} = \frac{1}{2}\Phi_{1jk}$ and hence $\|D^2_M\Phi_1\|^2_M = \frac{1}{4}\Phi^k_{1j}\Phi^j_{1k}$. Furthermore, the vector field $\nabla_M\Phi_1$ satisfies $\nabla_M\Phi_1 = \partial/\partial x_1$ and $|\nabla_M\Phi_1|^2_M = \Phi_{11}$. Since $L\Phi_1 = -V_1$, the Bochner formula (29) for $u = \Phi_1$ takes the form

$$\begin{aligned}\tfrac{1}{2}L(\Phi_{11}) &= -\langle\nabla_M\Phi_1, \nabla_M V_1\rangle_M + \tfrac{1}{4}\Phi^k_{1j}\Phi^j_{1k} + \mathrm{Ric}_M(\nabla_M u, \nabla_M u)\\ &= -V_{11} + \tfrac{1}{4}\Phi^k_{1j}\Phi^j_{1k} + (\mathrm{Ric}_M)_{11}.\end{aligned} \tag{30}$$

From (28) and (30), we obtain a formula for the Bakry–Émery–Ricci tensor:

$$(\mathrm{Ric}_M)_{11} = \tfrac{1}{4}\Phi^k_{1j}\Phi^j_{1k} + \tfrac{1}{2}V_{11} + \tfrac{1}{2}\sum_{j,k=1}^n \Phi_{j1}\Phi_{1k}(W_{jk}\circ\nabla\Phi).$$

It is clear that there is nothing special about the derivative $u = \Phi_1$, and that we could have repeated the argument with $u = \nabla\Phi\cdot\theta$ for any $\theta\in\mathbb{R}^n$. We thus obtain the formula

$$(\mathrm{Ric}_M)_{i\ell} = \tfrac{1}{4}\Phi^k_{ij}\Phi^j_{\ell k} + \tfrac{1}{2}V_{i\ell} + \tfrac{1}{2}\sum_{j,k=1}^n \Phi_{ji}\Phi_{\ell k}(W_{jk}\circ\nabla\Phi). \tag{31}$$

Since $D^2V$ and $D^2W$ are positive semidefinite, for any smooth $u:\mathbb{R}^n\to\mathbb{R}$ we have

$$\Gamma_2(u) \ge \mathrm{Ric}_M(\nabla_M u, \nabla_M u) \ge \tfrac{1}{4}\Phi^{ik}_j\Phi^{j\ell}_k u_i u_\ell,$$

and the third proof of Lemma 3.1 is complete.

Having finished with Lemma 3.1, let us introduce one of the main ideas in this paper, which was absent from [Klartag 2013]. The idea is to consider the map

$$\mathbb{R}^n \ni x \mapsto D^2\Phi(x) \in M^+_n(\mathbb{R}). \tag{32}$$

Denote by $(g_{ij}(x))_{i,j=1,\dots,n}$ the pull-back of the standard Riemannian metric on $M^+_n(\mathbb{R})$ via the map (32). It follows from Lemma 2.2 that $g_{ij}$ is given by the formula

$$g_{ij} = \mathrm{Tr}[(D^2\Phi)^{-1}\cdot\partial_i(D^2\Phi)\cdot(D^2\Phi)^{-1}\cdot\partial_j(D^2\Phi)] = \Phi^\ell_{ik}\Phi^k_{j\ell}. \tag{33}$$

Note that the positive semidefinite matrix $(g_{ij}(x))_{i,j=1,\dots,n}$ is not necessarily invertible, and it could happen that distinct points of $\mathbb{R}^n$ have zero Riemannian distance with respect to the Riemannian metric $(g_{ij})$. The metric $g_{ij}$ resembles an expression appearing in Lemma 3.1, a fact that will be exploited in the next section.

## 4. Dualizing the Bochner inequality

It is by now well known that, in the presence of convexity assumptions, Poincaré-type inequalities may be deduced from Bochner's formula via a dualization procedure. In this section we investigate the Poincaré inequality that is dual to Lemma 3.1. This Poincaré inequality was also obtained in [Klartag 2013], but in a cumbersome formulation and under an undesired assumption called "regularity at infinity", which we eliminate here.

We begin with an easy case. Throughout this section we assume, in addition to the smoothness assumptions made at the beginning of Section 3, that there exists $\varepsilon_0 > 0$ for which

$$D^2\Phi(x) \geq \varepsilon_0 \cdot \mathrm{Id} \quad (x \in \mathbb{R}^n) \tag{34}$$

in the sense of symmetric matrices. Write $C_c^\infty(\mathbb{R}^n)$ for the space of all compactly supported, smooth functions on $\mathbb{R}^n$. The following lemma is a variant of a well-known fact (see, e.g., [Strichartz 1983]), that compactly supported functions are dense in Sobolev spaces when the Riemannian manifold is complete. Our assumption (34) implies the completeness of the Riemannian manifold $M = M_{\mu,\nu}$.

**Lemma 4.1.** *Let $f \in L^2(\mu)$ satisfy $\int f\,d\mu = 0$. Then there exists a sequence $u_k \in C_c^\infty(\mathbb{R}^n)$ with*

$$\|Lu_k - f\|_{L^2(\mu)} \longrightarrow 0 \quad \text{as } k \to \infty.$$

*Proof.* Recall that $\int (Lu)\,d\mu = 0$ for all $u \in C_c^\infty(\mathbb{R}^n)$. To show that the linear space $\{Lu \mid u \in C_c^\infty(\mathbb{R}^n)\}$ is dense, we analyze its orthogonal complement. Let $f \in L^2(\mu)$ be in the orthogonal complement, i.e., for any $u \in C_c^\infty(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} f(Lu)\,d\mu = 0. \tag{35}$$

Our goal is to show that $f \equiv \mathrm{Const}$. Note that (35) means that $f$ is a weak solution of $Lf \equiv 0$. Since $L$ is elliptic, $f$ is smooth and $Lf \equiv 0$ in the classical sense. Thus,

$$L(f^2) = 2fLf + 2|\nabla_M f|^2 = 2|\nabla_M f|^2.$$

Therefore, for any $\eta \in C_c^\infty(\mathbb{R}^n)$,

$$\begin{aligned}\int_{\mathbb{R}^n} |\nabla_M(\eta f)|^2\,d\mu &= \int_{\mathbb{R}^n} \left[\eta^2|\nabla_M f|^2 + \tfrac{1}{2}\nabla_M(f^2)\cdot\nabla_M(\eta^2) + f^2|\nabla_M\eta|^2\right]d\mu \\ &= \int_{\mathbb{R}^n}\left[\eta^2|\nabla_M f|^2 - \tfrac{1}{2}\eta^2 L(f^2) + f^2|\nabla_M\eta|^2\right]d\mu = \int_{\mathbb{R}^n}|\nabla_M\eta|^2 f^2\,d\mu.\end{aligned}$$

However, according to our assumption (34), we have $|\nabla_M\eta|^2 = \Phi^{ij}\eta_i\eta_j \leq \varepsilon_0^{-1}|\nabla\eta|^2$. Let $\eta_R$ be a smooth cutoff function in $\mathbb{R}^n$ that equals one on a Euclidean ball of radius $R$ centered at the origin, equals zero outside a Euclidean ball of radius $2R$, and satisfies $|\nabla\eta_R| \leq 2/R$ throughout $\mathbb{R}^n$. Then,

$$\int_K |\nabla_M(\eta_R f)|^2\,d\mu \leq \int_{\mathbb{R}^n}|\nabla_M\eta_R|^2 f^2\,d\mu \leq \varepsilon_0^{-1}\int_{\mathbb{R}^n}|\nabla\eta_R|^2 f^2\,d\mu \leq \frac{2}{R\varepsilon_0}\int_{\mathbb{R}^n} f^2\,d\mu \longrightarrow 0 \quad \text{as } R\to\infty,$$

since $f \in L^2(\mu)$. Therefore, $\nabla f \equiv 0$ and $f$ is constant. □

Suppose that $F$ is a locally Lipschitz function on a Riemannian manifold such as $M_n^+(\mathbb{R})$. By the Rademacher theorem, the gradient $\nabla F$ is well defined almost everywhere with respect to the Riemannian volume measure. In order to have a function $|\nabla F|$ that is defined everywhere, in this note we set

$$|\nabla F|(x) = \limsup_{\substack{y\to x\\ z\to x}} \frac{|F(y)-F(z)|}{\mathrm{dist}(y,z)} = \lim_{\varepsilon\to 0^+} \sup_{\substack{y,z\in B(x,\varepsilon)\\ y\neq z}} \frac{|F(y)-F(z)|}{\mathrm{dist}(y,z)}, \tag{36}$$

where dist is the Riemannian distance and $B(x,\varepsilon) = \{y \mid \mathrm{dist}(x,y) < \varepsilon\}$. Since $F$ is locally Lipschitz, the function $|\nabla F|$ is locally bounded and upper semicontinuous. Clearly, at any point $x$ where $F$ is continuously differentiable, $|\nabla F|(x)$ equals the Riemannian length of $\nabla F(x)$.

**Proposition 4.2.** *Denote by $\theta$ the push-forward of the measure $\mu$ under the map* (32)*. Then, for any locally Lipschitz function $F : M_n^+(\mathbb{R}) \to \mathbb{R}$ that belongs to $L^2(\theta)$ with $\int_{M_n^+(\mathbb{R})} F\, d\theta = 0$,*

$$\int_{M_n^+(\mathbb{R})} F^2\, d\theta \le 4 \int_{M_n^+(\mathbb{R})} |\nabla F|^2\, d\theta$$

*whenever the right-hand side is finite.*

*Proof.* Since $F$ is locally Lipschitz in $L^2(\theta)$, the function $f$ defined via

$$f(x) = F(D^2\Phi(x)) \quad (x \in \mathbb{R}^n)$$

is locally Lipschitz in $\mathbb{R}^n$ and belongs to $L^2(\mu)$. Abbreviate $H = |\nabla F|^2$ and $h(x) = H(D^2\Phi(x))$. From the definition (36) of $|\nabla F|$, for any $x \in \mathbb{R}^n$ at which $f$ is differentiable,

$$h(x) \ge \sup\Big\{ \textstyle\sum_{i=1}^n V^i f_i \;\Big|\; \sum_{i,j=1}^n g_{ij} V^i V^j \le 1,\ V^1,\dots,V^n \in \mathbb{R}\Big\}, \tag{37}$$

where $f_i$ and $g_{ij}$ are evaluated at the point $x$. In the case where the matrix $(g_{ij}(x))_{i,j=1,\dots,n}$ is invertible, we may express the supremum in (37) in terms of the inverse matrix, yet it is the formula (37) which is valid in the general case. Setting $U_i = \Phi_{ij}V^j$, we reformulate (37) as

$$h(x) \ge \sup\Big\{ \Phi^{ij} U_j f_i \;\Big|\; g_{ij}\Phi^{ki}\Phi^{\ell j} U_k U_\ell \le 1,\ U_1,\dots,U_n \in \mathbb{R}\Big\}. \tag{38}$$

The formula (38) is valid for almost any $x \in \mathbb{R}^n$, since $f$ is differentiable almost everywhere in $\mathbb{R}^n$ by the Rademacher theorem. We would like to show that, for any $u \in C_c^\infty(\mathbb{R}^n)$,

$$-\int_{\mathbb{R}^n} f(Lu)\, d\mu \le 2\sqrt{\int_{\mathbb{R}^n} h^2\, d\mu}\sqrt{\int_{\mathbb{R}^n} (Lu)^2\, d\mu}. \tag{39}$$

To this end we observe that, since $u$ is compactly supported,

$$\int_{\mathbb{R}^n} \Gamma_2(u)\, d\mu = \frac{1}{2}\int_{\mathbb{R}^n} L(\Phi^{ij}u_iu_j)\, d\mu - \int_{\mathbb{R}^n} \Phi^{ij}(Lu)_i u_j\, d\mu = -\int_{\mathbb{R}^n} \Phi^{ij}(Lu)_i u_j\, d\mu = \int_{\mathbb{R}^n} (Lu)^2\, d\mu.$$

Therefore, Lemma 3.1 and (33) imply that, for any $u \in C_c^\infty(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} (Lu)^2\, d\mu \ge \frac{1}{4}\int_{\mathbb{R}^n} \Phi^{ik}\Phi^{j\ell} g_{k\ell} u_i u_j\, d\mu.$$

Since $f$ is locally Lipschitz, we may safely integrate by parts and obtain that, for any $u \in C_c^\infty(\mathbb{R}^n)$,

$$
\begin{aligned}
-\int_{\mathbb{R}^n} f(Lu)\,d\mu = \int_{\mathbb{R}^n} \Phi^{ij} f_i u_j \,d\mu &\le \int_{\mathbb{R}^n} h(x)\sqrt{g_{ij}\Phi^{ki}\Phi^{\ell j} u_k u_\ell}\,d\mu(x) \\
&\le \sqrt{\int_{\mathbb{R}^n} h^2\,d\mu}\sqrt{\int_{\mathbb{R}^n} g_{ij}\Phi^{ki}\Phi^{\ell j} u_k u_\ell\,d\mu} \le 2\sqrt{\int_{\mathbb{R}^n} h^2\,d\mu}\sqrt{\int_{\mathbb{R}^n} (Lu)^2\,d\mu}
\end{aligned}
$$

and (39) is proven. Since $\int_{M_n^+(\mathbb{R})} F\,d\theta = 0$, we also have $\int_{\mathbb{R}^n} f\,d\mu = 0$. From Lemma 4.1, there exists a sequence $u_k \in C_c^\infty(\mathbb{R}^n)$ with $Lu_k \longrightarrow -f$ in $L^2(\mu)$. We substitute $u = u_k$ in (39), and take the limit $k \to \infty$. This yields

$$
\int_{\mathbb{R}^n} f^2\,d\mu \le 2\sqrt{\int_{\mathbb{R}^n} h^2\,d\mu}\sqrt{\int_{\mathbb{R}^n} f^2\,d\mu}.
$$

Hence,

$$
\int_{\mathbb{R}^n} f^2\,d\mu \le 4\int_{\mathbb{R}^n} h^2\,d\mu.
$$

Since $h(x) = H(D^2\Phi)$ with $H = |\nabla F|^2$, the proposition is proven. □

## 5. Regularity issues

This section explains how to eliminate assumption (34) and also the smoothness assumptions of the previous two sections.

**Theorem 5.1.** *Assume that $\mu$ and $\nu$ are absolutely continuous, log-concave probability measures on $\mathbb{R}^n$. Let $\nabla\Phi$ be the Brenier map between $\mu$ and $\nu$ and assume condition ($\star$) from Section 1. Denote by $\theta$ the push-forward of the measure $\mu$ under the map $x \mapsto D^2\Phi(x)$.*

*Then, for any $\theta$-integrable, locally Lipschitz function $F : M_n^+(\mathbb{R}) \to \mathbb{R}$,*

$$
\int_{M_n^+(\mathbb{R})} F^2\,d\theta - \left(\int_{M_n^+(\mathbb{R})} F\,d\theta\right)^2 \le 4\int_{M_n^+(\mathbb{R})} |\nabla F|^2\,d\theta \tag{40}
$$

*whenever the right-hand side is finite and $|\nabla F|$ is interpreted as in* (36).

The strategy for proving Theorem 5.1 is to approximate $\Phi$ by a sequence of functions $\Phi_N$ that satisfy assumption (34), and to prove the pointwise (and even local uniform) convergence $D^2\Phi_N(x) \longrightarrow D^2\Phi(x)$ as $N \to \infty$. Below we discuss two possible justifications of this convergence, as we believe that both of them may be useful. The first proof occupies Section 5A, and is based on various results from the regularity theory of the Monge–Ampère equation. The log-concavity of the measures is not really required for the first proof, and it suffices to assume that the densities are locally Hölder.

The second proof, in Section 5B, is in fact an alternative approach to Caffareli's $C^{1,\alpha}$-regularity results in the log-concave case. The argument in Section 5B is more self-contained, and is based on integration-by-parts arguments. The log-concavity of the target measure plays an important role here, and we further assume a certain integrability condition on the logarithmic derivative of the density of $\mu$. This integrability condition is rather mild in our opinion, and it is satisfied in many cases of interest.

**5A.** ***First proof of Theorem 5.1.*** As before, we write $e^{-V}$ and $e^{-W}$ for the densities of $\mu$ and $\nu$ respectively. By log-concavity, the functions $V$ and $W$ are locally Lipschitz in the open sets $\mathrm{Supp}(\mu)$ and $\mathrm{Supp}(\nu)$ respectively. From condition ($\star$), the function $\Phi$ is $C^2$-smooth, and the push-forward equation (1) implies that

$$\det D^2\Phi(x) = e^{-V(x)+W(\nabla\Phi(x))} \tag{41}$$

for any $x \in \mathrm{Supp}(\mu)$. In particular, $D^2\Phi(x)$ is invertible, and hence positive-definite for all $x \in \mathrm{Supp}(\mu)$. Thus $\Phi$ is strictly convex. The modulus of convexity of $\Phi$ at the point $x$ is defined to be

$$\omega_\Phi(x;\delta) = \inf\bigl\{\Phi(y) - (\Phi(x) + \nabla\Phi(x)\cdot(y-x)) \bigm| y \in \mathbb{R}^n,\ |y-x| = \delta\bigr\}.$$

Then $\omega_\Phi(x;\delta)$ is a positive, continuous function of $x \in \mathrm{Supp}(\mu)$ and $\delta > 0$ when we restrict to $x$ and $\delta$ for which $\overline{B(x,\delta)} \subseteq \mathrm{Supp}(\mu)$. Here, $B(x,\delta) = \{y \in \mathbb{R}^n \mid |y-x| < \delta\}$. Next, the Legendre transform

$$\Phi^*(x) = \sup_{\substack{y\in\mathbb{R}^n \\ \Phi(y)<\infty}} [x\cdot y - \Phi(y)]$$

is also $C^2$-smooth and strictly convex in $\mathrm{Supp}(\nu)$, with $y \mapsto \nabla\Phi^*(y)$ being the inverse map to $x \mapsto \nabla\Phi(x)$. Thus, $\nabla\Phi$ is a $C^1$-diffeomorphism of $\mathrm{Supp}(\mu)$ and $\mathrm{Supp}(\nu)$. The reader is referred to [Rockafellar 1970] for the basic properties of the Legendre transform.

We will approximate $\mu$ and $\nu$ by sequences of probability measures $\mu_N$ and $\nu_N$ with the following properties:

(i) The probability measures $\mu_N$ and $\nu_N$ have densities in $\mathbb{R}^n$ of the form $e^{-V_N}$ and $e^{-W_N}$ respectively.

(ii) The functions $V_N, W_N : \mathbb{R}^n \to \mathbb{R}$ are smooth and, for any $x \in \mathbb{R}^n$,

$$D^2V_N(x) \geq \frac{1}{N}\cdot\mathrm{Id}, \qquad D^2W_N(x) \leq N\cdot\mathrm{Id}\,.$$

(iii) $V_N \longrightarrow V$ locally uniformly in $\mathrm{Supp}(\mu)$ and, similarly, $W_N \longrightarrow W$ locally uniformly in $\mathrm{Supp}(\nu)$.

It is quite standard to approximate $\mu$ and $\nu$ in this manner. For instance, in order to obtain $\mu_N$ (or $\nu_N$), we may convolve $\mu$ (or $\nu$) with a Gaussian of tiny variance, then multiply the resulting density by a Gaussian of huge variance, and then normalize to obtain a probability density. Denote by $\nabla\Phi_N$ the Brenier map between $\mu_N$ and $\nu_N$. Again, we use Caffarelli's regularity theory to conclude that $\Phi_N : \mathbb{R}^n \to \mathbb{R}$ is a smooth, strictly convex function, with

$$\det D^2\Phi_N(x) = e^{-V_N(x)+W_N(\nabla\Phi_N(x))} \quad (x \in \mathbb{R}^n). \tag{42}$$

The following lemma should be known to experts on the Monge–Ampère equation, yet we could not find it in the literature.

**Lemma 5.2.** *There exists an increasing sequence* $\{N_j\}$ *such that*

$$D^2\Phi_{N_j}(x) \longrightarrow D^2\Phi(x) \quad \text{as } j\to\infty$$

*locally uniformly in* $x \in \mathrm{Supp}(\mu)$.

*Proof.* Fix $x_0 \in \mathrm{Supp}(\mu)$. It suffices to find $\{N_j\}$ such that $D^2\Phi_{N_j} \longrightarrow D^2\Phi$ uniformly in a neighborhood of $x_0$. A standard convexity argument (e.g., [Klartag 2014, Section 2]) based on (iii) and the fact that $\int e^{-V} = \int e^{-W} = 1$ shows that there exist $A, B > 0$ with

$$\min\Bigl\{\inf_N V_N(x), \inf_N W_N(x), V(x), W(x)\Bigr\} \geq A|x| - B \quad (x \in \mathbb{R}^n). \tag{43}$$

Therefore,

$$\sup_N \int_{\mathbb{R}^n} |\nabla\Phi_N|^2 e^{-V_N(x)}\,dx = \sup_N \int_{\mathbb{R}^n} |x|^2 e^{-W_N(x)}\,dx \leq \int_{\mathbb{R}^n} |x|^2 e^{B - A|x|}\,dx < \infty. \tag{44}$$

Recall that $V_N \longrightarrow V$ locally uniformly in $\mathrm{Supp}(\mu)$, according to (iii). From (44) we learn that $\sup_N \|\Phi_N\|_{\dot{H}^1(K)} < \infty$ for any compact $K \subset \mathrm{Supp}(\mu)$. Here,

$$\|u\|^2_{\dot{H}^1(K)} = \int_K |\nabla u(x)|^2\,dx.$$

From the Rellich–Kondrachov compactness theorem (e.g., [Evans and Gariepy 1992, Section 4.6]), we conclude that there exist a subsequence $\Phi_{N_j}$, numbers $C_j \in \mathbb{R}$ and a certain function $F : \mathrm{Supp}(\mu) \to \mathbb{R}$ such that, for any compact $K \subset \mathrm{Supp}(\mu)$, the sequence $\Phi_{N_j} + C_j$ converges to $F$ in $L^2(K)$. Passing to another subsequence, which we conveniently denote again by $\{\Phi_N\}$, and using [Rockafellar 1970, Theorem 10.9], we may assume that $F$ is convex and that the convergence is locally uniform in $\mathrm{Supp}(\mu)$. Thus, from [ibid., Theorem 24.5],

$$\nabla\Phi_N(x) \longrightarrow \nabla F(x) \quad \text{as } N \to \infty \tag{45}$$

for almost any $x \in \mathrm{Supp}(\mu)$. However, $(\nabla\Phi_N)_*\mu_N = \nu_N$. From (iii), (43) and (45) we conclude that $(\nabla F)_*\mu = \nu$. From the uniqueness of the Brenier map, we deduce that $\nabla F = \nabla\Phi$ almost everywhere in $\mathrm{Supp}(\mu)$. Since $\Phi$ is $C^2$-smooth, we may then apply [ibid., Theorem 25.7] and upgrade (45) to

$$\nabla\Phi_N(x) \longrightarrow \nabla\Phi(x) \quad \text{as } N \to \infty \tag{46}$$

locally uniformly in $\mathrm{Supp}(\mu)$. The convexity arguments in [ibid., Section 25] also give that $\nabla\Phi_N^* \longrightarrow \nabla\Phi^*$ locally uniformly in $\mathrm{Supp}(\nu)$. As for the modulus of convexity, we have

$$\omega_{\Phi_N}(x;\delta) \longrightarrow \omega_{\Phi}(x;\delta) \quad \text{as } N \to \infty \qquad \text{and} \qquad \omega_{\Phi_N^*}(y;\delta) \longrightarrow \omega_{\Phi^*}(y;\delta) \quad \text{as } N \to \infty \tag{47}$$

locally uniformly in the sets $\{(x,\delta) \in \mathrm{Supp}(\mu) \times (0,\infty) \mid \overline{B(x,\delta)} \subset \mathrm{Supp}(\mu)\}$ and, respectively, $\{(y,\delta) \in \mathrm{Supp}(\nu) \times (0,\infty) \mid \overline{B(y,\delta)} \subset \mathrm{Supp}(\nu)\}$.

We will now invoke the estimates of Gutierrez and Huang [2000] and Forzani and Maldonado [2004; 2005], which are constructive versions of Caffarelli's $C^{1,\alpha}$-regularity theory. Thanks to (iii), (42), (46) and (47), we are allowed to apply [Gutiérrez and Huang 2000, Theorem 2.1] and [Forzani and Maldonado 2004, Theorem 15] locally near $x_0$. From the latter result, we learn that there exist $\alpha, \delta, C > 0$ such that, for any $x, y \in B(x_0,\delta)$ and $N \geq 1$,

$$|\nabla\Phi_N(x) - \nabla\Phi_N(y)| \leq C|x - y|^{\alpha}. \tag{48}$$

The function $V$ is locally Lipschitz. From (iii) and [Rockafellar 1970, Theorem 24.5], the sequence $\{V_N\}$ is uniformly locally Lipschitz. This means that, for any compact subset $K \subset \mathrm{Supp}(\mu)$, the Lipschitz constant of $V_N$ is bounded by some finite number $C_K$, independent of $N$. Similarly, the sequence $\{W_N\}$ is also uniformly locally Lipschitz. Together with (46) and (48), we deduce that there exists $\widehat{C} > 0$ such that $u_N(x) = -V_N(x) + W_N(\nabla\Phi_N(x))$ satisfies

$$|u_N(x) - u_N(y)| \le \widehat{C}|x-y|^\alpha \quad (x,\ y \in B(x_0,\delta),\ N \ge 1).$$

Recalling the Monge–Ampère equation (42), we learn that there exists $\widetilde{C} > 0$ such that

$$|\det D^2\Phi_N(x) - \det D^2\Phi_N(y)| \le \widetilde{C}|x-y|^\alpha \quad (x,\ y \in B(x_0,\delta),\ N \ge 1).$$

We are finally in good shape for applying the $C^{2,\alpha}$-estimates from [Trudinger and Wang 2008, Theorem 3.2]. These estimates yield the existence of $\overline{C} > 0$ such that, for any $x,\ y \in B(x_0, \delta/2)$ and $N \ge 1$,

$$\|D^2\Phi_N(x) - D^2\Phi_N(y)\|_{HS} \le \overline{C}|x-y|^\alpha. \tag{49}$$

The uniform $C^{2,\alpha}$-estimate in (49) allows us to apply the Arzella–Ascoli theorem. All we need is to let $K = B(x_0, \delta/2)$ and observe that

$$\int_K (\Delta\Phi_N)\xi = -\int_K \nabla\Phi_N \cdot \nabla\xi \longrightarrow -\int_K \nabla\Phi\cdot\nabla\xi = \int_K (\Delta\Phi)\xi \quad \text{as } N \to \infty,$$

where $\xi$ is any smooth, compactly supported function in $K$. Hence, the sequence $\left\{\int_K \Delta\Phi_N\right\}_{N\ge1}$ is bounded and, since $D^2\Phi_N$ is positive-definite, the sequence $\left\{\int_K \|D^2\Phi_N\|_{HS}\right\}_{N\ge1}$ is also bounded. From (49) and the Arzella–Ascoli theorem, there exists a subsequence, still denoted by $\{\Phi_N\}$, such that $D^2\Phi_N \longrightarrow D^2\Phi$ uniformly on $K = B(x_0, \delta/2)$. □

**Remark 5.3.** Our proof of Lemma 5.2 does not make any use of the log-concavity of $\mu$ and $\nu$. By inspecting the proof above, we see that Lemma 5.2 holds true as long as $V$ and $W$ are locally Hölder, and $V_N$, $W_N$ are uniformly locally Hölder.

In order to simplify the notation, we denote the sequence $\{\Phi_{N_j}\}$ from Lemma 5.2 by $\{\Phi_N\}$. Properties (i), (ii) and (iii) above are still satisfied.

**Corollary 5.4.** *Denote by $\theta_N$ the push-forward of the measure $\mu_N$ under the map $x \mapsto D^2\Phi_N(x)$. Then, for any bounded, continuous function $b : M_n^+(\mathbb{R}) \to \mathbb{R}$,*

$$\int_{M_n^+(\mathbb{R})} b\, d\theta_N \longrightarrow \int_{M_n^+(\mathbb{R})} b\, d\theta \quad \textit{as } N \to \infty. \tag{50}$$

*Furthermore, if $b : M_n^+(\mathbb{R}) \to \mathbb{R}$ is bounded and upper semicontinuous, then*

$$\limsup_{N\to\infty} \int_{M_n^+(\mathbb{R})} b\, d\theta_N \le \int_{M_n^+(\mathbb{R})} b\, d\theta. \tag{51}$$

*Proof.* In order to prove (50), we need to show that

$$\int_{\mathbb{R}^n} b(D^2\Phi_N(x))e^{-V_N(x)}\, dx \longrightarrow \int_{\mathbb{R}^n} b(D^2\Phi(x))e^{-V(x)}\, dx \quad \text{as } N \to \infty.$$

This follows from Lemma 5.2 and from the dominated convergence theorem, since (43) provides an integrable majorant. Next, assume that $b$ is bounded and upper semicontinuous. Then, for any $x \in \mathrm{Supp}(\mu)$,

$$\limsup_{N\to\infty} b(D^2\Phi_N(x))e^{-V_N(x)} \le b(D^2\Phi(x))e^{-V(x)}.$$

Now (51) follows from Fatou's lemma, since we have an integrable majorant by (43). □

*Proof of Theorem 5.1.* Assume first that the locally Lipschitz function $F$ is compactly supported. We observe that, for any fixed $N$, assumption (34) holds true. Indeed, we may apply a refinement of Caffarelli's contraction theorem [2000] which appears in [Kolesnikov 2010], and thus obtain from (ii) that, for any $x \in \mathbb{R}^n$,

$$D^2\Phi_N(x) \ge \frac{1}{N^2} \cdot \mathrm{Id}\,.$$

We may therefore apply Proposition 4.2 and conclude that, for any $N \ge 1$,

$$\int_{M_n^+(\mathbb{R})} F^2\, d\theta_N - \left(\int_{M_n^+(\mathbb{R})} F\, d\theta_N\right)^2 \le 4\int_{M_n^+(\mathbb{R})} |\nabla F|^2\, d\theta_N.$$

Recall that $|\nabla F|^2$ is upper semicontinuous and bounded, while $F$ is continuous and bounded. By taking the limit as $N \to \infty$ and using Corollary 5.4, we obtain that

$$\int_{M_n^+(\mathbb{R})} F^2\, d\theta - \left(\int_{M_n^+(\mathbb{R})} F\, d\theta\right)^2 \le 4\int_{M_n^+(\mathbb{R})} |\nabla F|^2\, d\theta,$$

and (40) is proven in the case where $F$ is a compactly supported function.

The next step is to prove (40) under the additional assumption that $F \in L^2(\theta)$. To that end, we select a smooth function $\theta_R : M_n^+(\mathbb{R}) \to [0, 1]$ such that $\theta_R$ equals one on $B(\mathrm{Id}, R)$ and vanishes outside $B(\mathrm{Id}, 2R)$, with $|\nabla\theta_R| \le 2/R$. Set $F_R = \theta_R F$. We have just proven that (40) holds true when $F$ is replaced by $F_R$. Clearly, $F_R \longrightarrow F$ in $L^2(\theta)$ as $R \to \infty$. All that remains is to show that

$$\limsup_{R\to\infty} \int_{M_n^+(\mathbb{R})} |\nabla F_R|^2\, d\theta \le \int_{M_n^+(\mathbb{R})} |\nabla F|^2\, d\theta. \tag{52}$$

The functions $\theta_R$ and $F$ are continuous, and we may therefore use the Leibnitz rule

$$|\nabla F_R| \le |F|\,|\nabla\theta_R| + \theta_R|\nabla F| \le |\nabla F| + 2|F|/R,$$

where we interpret $|\nabla F|$ and $|\nabla F_R|$ in the sense of definition (36). Since $F, |\nabla F| \in L^2(\theta)$, (52) follows in the case where $F \in L^2(\theta)$.

Finally, to eliminate the assumption that $F$ is in $L^2(\theta)$, we replace $F$ by $F_R = \max\{-R, \min\{F, R\}\}$, apply the inequality for $F_R$, and let $R$ tend to infinity. For all but countably many values of $R$, the level set $\{A \in M_n^+(\mathbb{R}) \mid F(A) = R\}$ has zero $\theta$-measure. Consequently, we have the inequality $\int |\nabla F_R|^2\, d\theta \le \int |\nabla F|^2\, d\theta$ for all but countably many values of $R$, and (40) follows. □

**5B.** ***Second proof: log-concave target measure.*** In our second proof we will exploit the fact that $\nu$ is log-concave, but we will not require the log-concavity of $\mu$. Throughout this subsection we make the following additional assumption:

**(A)** For some $p > n$,
$$\int_{\mathbb{R}^n} |\nabla V|^p e^{-V}\, dx < \infty,$$
where the derivatives $V_i$ are understood in the logarithmic derivative sense, i.e.,
$$\int_{\mathbb{R}^n} \xi V_i \, d\mu = -\int_{\mathbb{R}^n} \xi_i \, d\mu, \quad \xi \in C_c^\infty(\mathbb{R}^n),\ i = 1, \dots, n.$$

By the Morrey embedding theorem (see, e.g., [Evans and Gariepy 1992, Section 4.5]), the function $V$ is locally Hölder. We will approximate $\mu$ and $\nu$ by sequences of probability measures $\mu_N$ and $\nu_N$ having properties (i), (ii) and (iii) from Section 5A. We also require a fourth property:

(i) There exists $p > n$ such that
$$\sup_N \int_{\mathbb{R}^n} |\nabla V_N|^p e^{-V_N}\, dx < \infty.$$

The approach outlined in Section 5A — to convolve with a tiny Gaussian and then multiply by the density of a huge Gaussian — also yields property (iv). Recall that the Brenier map $\nabla\Phi_N$ between $\mu_N$ and $\nu_N$ is smooth and that it satisfies (42). The central ingredient of this subsection is the following a priori estimate:

**Proposition 5.5.** *Assume that functions $V$, $W$ and $\Phi$ are smooth on the entire $\mathbb{R}^n$ and that $\nu$ is a log-concave measure. Then, for every $q \geq 2$, $0 < \tau < 1$, $i = 1, \dots, n$, there exists $C(q, \tau) > 0$ such that*
$$\int_{\mathbb{R}^n} \Phi_{ii}^q \, d\mu \leq C(q,\tau)\biggl(\int_{\mathbb{R}^n} |V_i|^{2q/(2-\tau)}\, d\mu + \int_{\mathbb{R}^n} |x_i|^{2q/\tau}\, d\nu\biggr). \tag{53}$$

*Proof.* Assume in addition that $D^2W \geq (1/C)\cdot \mathrm{Id}$, $D^2V \leq C\cdot\mathrm{Id}$. In this case, $D^2\Phi \leq C^2\cdot\mathrm{Id}$. Recall formula (28):
$$L(\Phi_{ii}) - \Phi_i^{jk}\Phi_{ijk} - \sum_{j,k=1}^n \Phi_{ji}\Phi_{ik}W_{jk}\circ\nabla\Phi = -V_{ii},$$

which is obtained by differentiating the change of variables formula (22) along $x_i$. Let us multiply this formula by $\Phi_{ii}^p$, $p \geq 0$, and formally integrate by parts with respect to $\mu$. Using the convexity of $W$ we obtain
$$\int V_{ii}\Phi_{ii}^p\, d\mu \geq p\int \Phi_{ii}^{p-1}\langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle\, d\mu + \int \Phi_{ii}^p\Phi_i^{jk}\Phi_{ijk}\, d\mu. \tag{54}$$

Let us justify this formula. To this end, we fix a compactly supported function $\eta \geq 0$ and integrate with respect to $\eta\cdot\mu$:
$$\int V_{ii}\Phi_{ii}^p\eta\, d\mu$$
$$\geq \int \langle (D^2\Phi)^{-1}\nabla\eta, \nabla\Phi_{ii}\rangle \Phi_{ii}^p\, d\mu + p\int \Phi_{ii}^{p-1}\langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle\eta\, d\mu + \int \Phi_{ii}^p\Phi_i^{jk}\Phi_{ijk}\eta\, d\mu.$$

Applying the Cauchy inequality yields

$$-\int \langle (D^2\Phi)^{-1}\nabla\eta, \nabla\Phi_{ii}\rangle \Phi_{ii}^p \, d\mu \le \frac{4}{\varepsilon}\int \frac{\langle (D^2\Phi)^{-1}\nabla\eta, \nabla\eta\rangle}{\eta}\Phi_{ii}^{p+1}\, d\mu + \varepsilon \int \langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle \Phi_{ii}^{p-1}\eta\, d\mu.$$

Finally,

$$\int V_{ii}\Phi_{ii}^p \eta\, d\mu + \frac{4}{\varepsilon}\int \frac{\langle (D^2\Phi)^{-1}\nabla\eta, \nabla\eta\rangle}{\eta}\Phi_{ii}^{p+1}\, d\mu \ge (p-\varepsilon)\int \Phi_{ii}^{p-1}\langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle \eta\, d\mu + \int \Phi_{ii}^p \Phi_i^{jk}\Phi_{ijk}\eta\, d\mu.$$

Assume that $\eta$ has the form $\eta = \xi(\nabla\Phi)$, where $\xi$ is compactly supported. We obtain

$$\int V_{ii}\Phi_{ii}^p \eta\, d\mu + \frac{4C^{p+2}}{\varepsilon}\int \frac{|\nabla\xi|^2}{\xi}\, d\nu \ge (p-\varepsilon)\int \Phi_{ii}^{p-1}\langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle \eta\, d\mu + \int \Phi_{ii}^p \Phi_i^{jk}\Phi_{ijk}\eta\, d\mu.$$

It remains to construct a sequence of functions $1 \ge \xi_N \ge 0$ satisfying $\lim_N \xi_N(x) = 1$ for $\nu$-a.e. $x$ and $\lim_N \int |\nabla\xi_N|^2/\xi_N \, d\nu = 0$. Then, by applying the Fatou lemma we justify (54).

It is helpful to keep in mind that $\Phi_i^{jk}\Phi_{ijk} = \mathrm{Tr}[(D^2\Phi)^{-1/2}D^2\Phi_i(D^2\Phi)^{-1/2}]^2 \ge 0$. From (54),

$$\int V_{ii}\Phi_{ii}^p\, d\mu \ge p\int \Phi_{ii}^{p-1}\langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle\, d\mu.$$

Let us integrate by parts the left-hand side: $\int V_{ii}\Phi_{ii}^p\, d\mu = \int V_i^2\Phi_{ii}^p\, d\mu - p\int V_i\Phi_{ii}^{p-1}\Phi_{iii}\, d\mu$. The justification of this integration by parts is much easier, since $D^2\Phi$ and $D^2V$ are bounded. Applying

$$2|\Phi_{iii}V_i| \le 2|V_i|\sqrt{\Phi_{ii}\cdot\langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle} \le V_i^2\Phi_{ii} + \langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle,$$

one obtains

$$\int V_i^2\Phi_{ii}^p\, d\mu \ge \int \Phi_{ii}^{p-1}\langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle\, d\mu. \tag{55}$$

Let us show that the right-hand side controls powers of the second derivative $\Phi_{ii}$. Indeed, for every $q \ge 2$ and $\varepsilon > 0$, $0 \le \tau \le 1$, the following estimate holds:

$$\begin{aligned}\int \Phi_{ii}^q\, d\mu &= -(q-1)\int \Phi_i\Phi_{iii}\Phi_{ii}^{q-2}\, d\mu + \int \Phi_i V_i \Phi_{ii}^{q-1}\, d\mu \\ &\le \varepsilon\int \Phi_i^2\Phi_{ii}^{q-\tau}\, d\mu + \frac{(q-1)^2}{4\varepsilon}\int \Phi_{ii}^{q-3+\tau}\langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle\, d\mu \\ &\quad + \frac{q-1}{q}\int \Phi_{ii}^q\, d\mu + \frac{1}{q}\int |\Phi_i V_i|^q\, d\mu.\end{aligned}$$

Finally,

$$\int \Phi_{ii}^q \, d\mu \le \int |\Phi_i V_i|^q \, d\mu + q\varepsilon \int \Phi_i^2 \Phi_{ii}^{q-\tau} \, d\mu + \frac{q(q-1)^2}{4\varepsilon} \int \Phi_{ii}^{q-3+\tau} \langle (D^2\Phi)^{-1} \nabla \Phi_{ii}, \nabla \Phi_{ii} \rangle \, d\mu$$
$$\le \int |\Phi_i V_i|^q \, d\mu + q\varepsilon \int \Phi_i^2 \Phi_{ii}^{q-\tau} \, d\mu + \frac{q(q-1)^2}{4\varepsilon} \int \Phi_{ii}^{q-2+\tau} V_i^2 \, d\mu.$$

Applying the Hölder inequalities

$$\Phi_i^2 \Phi_{ii}^{q-\tau} \le \frac{q-\tau}{q} \Phi_{ii}^q + \frac{\tau}{q} |\Phi_i|^{2q/\tau},$$
$$\Phi_{ii}^{q-2+\tau} V_i^2 \le \varepsilon \Phi_{ii}^q + C(\varepsilon, q, \tau) |V_i|^{2q/(2-\tau)},$$
$$|\Phi_i V_i|^q \le \tfrac{1}{2}(2-\tau)|V_i|^{2q/(2-\tau)} + \tfrac{1}{2}\tau |\Phi_i|^{2q/\tau},$$

choosing a sufficiently small $\varepsilon$, and applying the change of variables formula $\int |\Phi_i|^q \, d\mu = \int |x_i|^q \, d\nu$, we easily obtain the claim.

Finally, let us get rid of the assumption that $D^2 W \ge (1/C) \cdot \mathrm{Id}$, $D^2 V \le C \cdot \mathrm{Id}$. To this end we approximate $\mu$ and $\nu$ by measures with smooth potentials satisfying $D^2 W_N \ge (1/C_N) \cdot \mathrm{Id}$, $D^2 V_N \le C_N \cdot \mathrm{Id}$ satisfying $\lim_N \int |(V_N)_i|^{2q} \, d\mu_N = \int |V_i|^{2q} \, d\mu$ and $\lim_N \int |x_i|^{2q} \, d\nu_N = \int |x_i|^{2q} \, d\nu$. It remains to show that the weak $L^q(\mu)$-limit of $(\Phi_N)_{ii}$ coincides with $\Phi_{ii}$. The latter can be easily shown with the help of integration by parts and identifications of the pointwise limit $\lim_N \nabla \Phi_N$ with $\nabla \Phi$ (see the proof of Lemma 5.2). □

**Remark 5.6.** The conclusion of Proposition 5.5 holds without any additional smoothness assumptions. This can be verified by smooth approximations (again, see [Kolesnikov 2013] for details). Finally, we see that (53) holds for every log-concave measure $\nu$ and measure $\mu$ satisfying $\int |V_i|^{2q/(2-\tau)} \, d\mu < \infty$, where $V_i$ is the logarithmic derivative of $\mu$ along $x_i$.

*Second proof of Lemma 5.2.* Let us demonstrate how Proposition 5.5 implies (48) above without appealing to the works by Forzani and Maldonado [2004; 2005] and Gutierrez and Huang [2000] related to Caffarelli's $C^{1,\alpha}$-regularity theory. We know that $\sup_N \int |\nabla V_N|^p e^{-V_N} \, dx < \infty$, $p > n$. Since $\nu$ is log-concave, all the moments of $\nu$ are finite. Thus, Proposition 5.5 implies

$$\sup_N \int \|D^2 \Phi_N\|_{HS}^{p'} e^{-V_N} \, dx < \infty$$

for any $n < p' < p$. Applying the fact that the $V_N$ are uniformly locally bounded from below, we see that $\sup_N \int_{B_R} \|D^2 \Phi_N\|_{HS}^{p'} \, dx < \infty$ for every $R$. The result then follows from the Morrey embedding theorem. □

## 6. Corollaries to Theorem 5.1

*Proof of Theorem 1.2.* For $A \in M_n^+(\mathbb{R})$, define

$$F(A) = f(\log \lambda_1(A), \dots, \log \lambda_n(A)),$$

where $0 < \lambda_1(A) \leq \cdots \leq \lambda_n(A)$ are the eigenvalues of $A$. According to Lemma 2.4, for any $A \in M_n^+(\mathbb{R})$,

$$|\nabla F|(A) \leq |\nabla f|(\log \lambda_1(A), \ldots, \log \lambda_n(A)). \tag{56}$$

Since $f$ is locally Lipschitz and the eigenvalues vary continuously with the matrix $A$, (56) implies that $F$ is locally Lipschitz. Denote by $\theta$ the push-forward of the probability measure $\mu$ under the map $x \mapsto D^2\Phi(x)$. Since $\mathbb{E}|f(\Lambda(X))| < \infty$, $F \in L^1(\theta)$. Since $\mathbb{E}|\nabla f|^2(\Lambda(X)) < \infty$, $\int |\nabla F|^2 \, d\theta < \infty$. We may apply Theorem 5.1 and conclude that

$$\int_{M_n^+(\mathbb{R})} F^2 \, d\theta - \left( \int_{M_n^+(\mathbb{R})} F \, d\theta \right)^2 \leq 4 \int_{M_n^+(\mathbb{R})} |\nabla F|^2 \, d\theta.$$

The left-hand side equals $\mathrm{Var}[f(\Lambda(X))]$. Glancing at (56), we thus obtain

$$\mathrm{Var}[f(\Lambda(X))] \leq 4\mathbb{E}|\nabla f|^2(\Lambda(X)),$$

and the proof is complete. □

*Proof of Theorem 1.1.* Substitute $f(x) = x_i$ in Theorem 1.2. Then $f$ is a 1-Lipschitz function, and by Remark 1.4 we have $\mathbb{E}|f(\Lambda(X))| < \infty$. Thus, the application of Theorem 1.2 is legitimate, and Theorem 1.1 follows. □

*Proof of Theorem 1.5.* The argument is almost identical to the proof of Theorem 1.1, with Lemma 2.3 replacing Lemma 2.4. □

Let us end this paper with a few remarks concerning future research. If we make further assumptions regarding the log-concave measures in question, it should be possible to prove concentration inequalities for the eigenvalues of $D^2\Phi$ themselves and not only for their logarithms. For example, there is a soft argument which shows that, when $\nabla\Phi$ is the Brenier map between the uniform measure on $K$ and the uniform measure on $T$,

$$\int_K \Delta\Phi \leq nV(K, \ldots, K, T),$$

where $V$ stands for mixed volume. The details will be discussed elsewhere. Another possible research direction is to investigate whether phenomena similar to Theorem 1.1 occur also in a nonlinear setting, when transporting measures with convexity properties supported on Riemannian manifolds.

## Acknowledgements

We would like to thank Emanuel Milman for interesting discussions. The research made in Sections 1–2, 4, 6 was supported by a grant from the European Research Council (ERC). The research made in Sections 3, 5 was supported by the Russian Science Foundation grant 14-11-00196.

## References

[Alesker et al. 1999] S. Alesker, S. Dar, and V. Milman, "A remarkable measure preserving diffeomorphism between two convex bodies in $\mathbf{R}^n$", *Geom. Dedicata* **74**:2 (1999), 201–212. MR 2000a:52004 Zbl 0927.52007

[Bakry and Émery 1985] D. Bakry and M. Émery, “Diffusions hypercontractives”, pp. 177–206 in *Séminaire de probabilités, XIX, 1983/84*, edited by J. Azéma and M. Yor, Lecture Notes in Math. **1123**, Springer, Berlin, 1985. MR 88j:60131 Zbl 0561.60080

[Bakry et al. 2014] D. Bakry, I. Gentil, and M. Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundlehren der Math. Wiss. **348**, Springer, Cham, 2014. MR 3155209 Zbl 06175511

[Brenier 1991] Y. Brenier, “Polar factorization and monotone rearrangement of vector-valued functions”, *Comm. Pure Appl. Math.* **44**:4 (1991), 375–417. MR 92d:46088 Zbl 0738.46011

[Caffarelli 1990] L. A. Caffarelli, “A localization property of viscosity solutions to the Monge–Ampère equation and their strict convexity”, *Ann. of Math.* (2) **131**:1 (1990), 129–134. MR 91f:35058 Zbl 0704.35045

[Caffarelli 1992] L. A. Caffarelli, “The regularity of mappings with a convex potential”, *J. Amer. Math. Soc.* **5**:1 (1992), 99–104. MR 92j:35018 Zbl 0753.35031

[Caffarelli 2000] L. A. Caffarelli, “Monotonicity properties of optimal transportation and the FKG and related inequalities”, *Comm. Math. Phys.* **214**:3 (2000), 547–563. Erratum in **225**:2 (2002), 449–450. MR 2002c:60029 Zbl 0978.60107

[Delanoë 1991] P. Delanoë, “Classical solvability in dimension two of the second boundary-value problem associated with the Monge–Ampère operator”, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **8**:5 (1991), 443–457. MR 92g:35070 Zbl 0778.35037

[Evans and Gariepy 1992] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, FL, 1992. MR 93f:28001 Zbl 0804.28001

[Forzani and Maldonado 2004] L. Forzani and D. Maldonado, “Properties of the solutions to the Monge–Ampère equation”, *Nonlinear Anal.* **57**:5-6 (2004), 815–829. MR 2005c:35095 Zbl 1137.35361

[Forzani and Maldonado 2005] L. Forzani and D. Maldonado, “Recent progress on the Monge–Ampère equation”, pp. 189–198 in *The $p$-harmonic equation and recent advances in analysis*, edited by P. Poggi-Corradini, Contemp. Math. **370**, Amer. Math. Soc., Providence, RI, 2005. MR 2005k:35115 Zbl 1134.35334

[Grigor’yan 2009] A. Grigor’yan, *Heat kernel and analysis on manifolds*, AMS/IP Studies in Advanced Mathematics **47**, American Mathematical Society, Providence, RI, 2009. MR 2011e:58041 Zbl 1206.58008

[Gromov and Milman 1983] M. Gromov and V. D. Milman, “A topological application of the isoperimetric inequality”, *Amer. J. Math.* **105**:4 (1983), 843–854. MR 84k:28012 Zbl 0522.53039

[Gutiérrez and Huang 2000] C. E. Gutiérrez and Q. Huang, “Geometric properties of the sections of solutions to the Monge–Ampère equation”, *Trans. Amer. Math. Soc.* **352**:9 (2000), 4381–4396. MR 2000m:35060 Zbl 0958.35043

[Klartag 2013] B. Klartag, “Poincaré inequalities and moment maps”, *Ann. Fac. Sci. Toulouse Math.* (6) **22**:1 (2013), 1–41. MR 3247770 Zbl 1279.60036

[Klartag 2014] B. Klartag, “Logarithmically-concave moment measures I”, pp. 231–260 in *Geometric aspects of functional analysis*, edited by B. Klartag and E. Milman, Lecture Notes in Math. **2116**, Springer International Publishing, 2014.

[Kolesnikov 2010] A. V. Kolesnikov, “Mass transportation and contractions”, *MIPT Proc.* **2**:4 (2010), 90–99. In Russian; translated at arXiv 1103.1479.

[Kolesnikov 2013] A. V. Kolesnikov, “Sobolev regularity of mass transportation and transportation inequalities”, *Teor. Veroyatnost. i Primenen.* **57**:2 (2013), 296–321. In Russian; translated in *Theory Probab. Appl.* **57**:2 (2013), 243–264. MR 3201654 Zbl 1279.60032

[Kolesnikov 2014] A. V. Kolesnikov, “Hessian metrics, $CD(K, N)$-spaces, and optimal transportation of log-concave measures”, *Discrete Contin. Dyn. Syst.* **34**:4 (2014), 1511–1532. MR 3121630 Zbl 1279.35045

[McCann 1995] R. J. McCann, “Existence and uniqueness of monotone measure-preserving maps”, *Duke Math. J.* **80**:2 (1995), 309–323. MR 97d:49045 Zbl 0873.28009

[Polya 1950] G. Polya, “Remark on Weyl’s note “Inequalities between the two kinds of eigenvalues of a linear transformation.””, *Proc. Nat. Acad. Sci. U. S. A.* **36** (1950), 49–51. MR 11,526b Zbl 0041.15402

[Reed and Simon 1978] M. Reed and B. Simon, *Methods of modern mathematical physics, IV: Analysis of operators*, Academic Press, New York, 1978. MR 58 #12429c Zbl 0401.47001

[Rockafellar 1970] R. T. Rockafellar, *Convex analysis*, Princeton Mathematical Series **28**, Princeton University Press, Princeton, N.J., 1970. MR 43 #445 Zbl 0193.18401

[Strichartz 1983] R. S. Strichartz, "Analysis of the Laplacian on the complete Riemannian manifold", *J. Funct. Anal.* **52**:1 (1983), 48–79. MR 84m:58138 Zbl 0515.58037

[Trudinger and Wang 2008] N. S. Trudinger and X.-J. Wang, "The Monge–Ampère equation and its geometric applications", pp. 467–524 in *Handbook of geometric analysis*, vol. 1, edited by L. Ji et al., Adv. Lect. Math. (ALM) **7**, International Press, 2008. MR 2010g:53065 Zbl 1156.35033

[Urbas 1997] J. Urbas, "On the second boundary value problem for equations of Monge–Ampère type", *J. Reine Angew. Math.* **487** (1997), 115–124. MR 98f:35057 Zbl 0880.35031

[Weyl 1949] H. Weyl, "Inequalities between the two kinds of eigenvalues of a linear transformation", *Proc. Nat. Acad. Sci. U. S. A.* **35** (1949), 408–411. MR 11,37d Zbl 0032.38701

Received 13 Feb 2014. Accepted 2 Dec 2014.

BO'AZ B. KLARTAG: klartagb@post.tau.ac.il
*School of Mathematical Sciences, Tel Aviv University, 69978 Tel Aviv, Israel*

ALEXANDER V. KOLESNIKOV: sascha77@mail.ru
*Faculty of Mathematics, National Research University Higher School of Economics, 117312, Moscow, Vavilova St., 7, Russia*

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 8, No. 1, 2015
dx.doi.org/10.2140/apde.2015.8.57

# NONLOCAL SELF-IMPROVING PROPERTIES

TUOMO KUUSI, GIUSEPPE MINGIONE AND YANNICK SIRE

Solutions to nonlocal equations with measurable coefficients are higher differentiable.

Specifically, we consider nonlocal integrodifferential equations with measurable coefficients whose model is given by

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}[u(x)-u(y)][\eta(x)-\eta(y)]K(x,y)\,dx\,dy=\int_{\mathbb{R}^n} f\eta\,dx \quad \text{for all } \eta\in C_c^\infty(\mathbb{R}^n),$$

where the kernel $K(\cdot)$ is a measurable function and satisfies the bounds

$$\frac{1}{\Lambda|x-y|^{n+2\alpha}}\le K(x,y)\le\frac{\Lambda}{|x-y|^{n+2\alpha}}$$

with $0<\alpha<1$, $\Lambda>1$, while $f\in L^q_{\mathrm{loc}}(\mathbb{R}^n)$ for some $q>2n/(n+2\alpha)$. The main result states that there exists a positive, universal exponent $\delta\equiv\delta(n,\alpha,\Lambda,q)$ such that for every weak solution $u$ the self-improving property

$$u\in W^{\alpha,2}(\mathbb{R}^n) \quad\Longrightarrow\quad u\in W^{\alpha+\delta,2+\delta}_{\mathrm{loc}}(\mathbb{R}^n)$$

holds. This differentiability improvement is a genuinely nonlocal phenomenon and does not appear in the local case, where solutions to linear equations in divergence form with measurable coefficients are known to be higher integrable but are not, in general, higher differentiable.

The result is achieved by proving a new version of the Gehring lemma involving certain families of lifted reverse Hölder-type inequalities in $\mathbb{R}^{2n}$ and which is implied by delicate covering and exit-time arguments. In turn, such reverse Hölder inequalities are based on the concept of dual pairs, that is, pairs $(\mu,U)$ of measures and functions in $\mathbb{R}^{2n}$ which are canonically associated to solutions. We also allow for more general equations involving as a source term an integrodifferential operator whose kernel does not necessarily have to be of order $\alpha$.



*MSC2010:* primary 35D10; secondary 35R11.
*Keywords:* elliptic equations, fractional differentiability, nonlocal operators.

## 1. Introduction

A basic and fundamental result in the theory of linear and nonlinear elliptic equations is given by the higher integrability of solutions. This falls in the realm of so-called self-improving properties. The result was first pioneered by Meyers [1963] and Elcrat and Meyers [1975], and then extended in various directions and in several different contexts; see for instance [Bojarski and Iwaniec 1983; Fusco and Sbordone 1990; Giusti 2003; Kinnunen and Lewis 2000]. Modern proofs of this property in the nonlinear case rely on the so-called Gehring lemma [Gehring 1973; Iwaniec 1998]. In the simplest possible instance the result in question asserts that distributional $W^{1,2}(\Omega)$-solutions $u$ to linear elliptic equations

$$-\operatorname{div}(A(x)Du) = f \in L_{\mathrm{loc}}^{\frac{2n}{n+2}+\delta_0}(\Omega), \quad \delta_0 > 0,$$

actually belong to a better Sobolev space:

$$u \in W_{\mathrm{loc}}^{1,2+\delta}(\Omega), \tag{1-1}$$

for some positive $\delta \le \delta_0$. Here $\Omega \subset \mathbb{R}^n$ is an open subset and $n \ge 2$. The matrix $A(x)$ is supposed to be elliptic and with bounded and measurable entries, that is,

$$\Lambda^{-1}|\xi|^2 \le \langle A(x)\xi, \xi\rangle \quad \text{and} \quad |A(x)| \le \Lambda \tag{1-2}$$

hold whenever $\xi \in \mathbb{R}^n$, $x \in \Omega$, where $\Lambda > 1$. The number $\delta > 0$ appearing in (1-1) is universal in the sense that, essentially, it depends neither on the solution $u$ nor the specific equation considered. It rather depends only on $n$, $\Lambda$, that is, on the ellipticity rate of the equation considered. The key point here is the measurability of the coefficients; when $A(\cdot)$ has more regular entries, higher regularity of solutions follows from the corresponding result for equations with constant coefficients, via perturbation. This is the reason why the result in (1-1) lies deep in the core of regularity theory, and allows for a proof of several other regularity results; see for instance [Giusti 2003].

We are interested in studying self-improving properties of solutions to nonlocal problems. To outline the results in a special yet meaningful model case, let us consider weak solutions $u \in W^{\alpha,2}(\mathbb{R}^n)$ of the nonlocal equation

$$\mathcal{E}_K(u, \eta) = \langle f, \eta\rangle \quad \text{for every test function } \eta \in C_c^\infty(\mathbb{R}^n), \tag{1-3}$$

where $f \in L_{\mathrm{loc}}^{2+\delta_0}(\mathbb{R}^n)$ and

$$\mathcal{E}_K(u, \eta) := \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} [u(x) - u(y)][\eta(x) - \eta(y)]K(x, y)\, dx\, dy.$$

The measurable kernel is assumed to satisfy the uniform ellipticity assumptions

$$\frac{1}{\Lambda|x-y|^{n+2\alpha}} \le K(x, y) \le \frac{\Lambda}{|x-y|^{n+2\alpha}} \tag{1-4}$$

for every $x, y \in \mathbb{R}^n$, where $\alpha \in (0, 1)$ and $\Lambda \geq 1$. We recall that the fractional Sobolev space $W^{s,\gamma}$, for $\gamma \geq 1$ and $s \in (0, 1)$, is given by the subspace of $L^\gamma(\mathbb{R}^n)$-functions $u$ for which the Gagliardo seminorm

$$[u]_{s,\gamma}^\gamma := \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|u(x) - u(y)|^\gamma}{|x - y|^{n+\gamma s}} \, dx \, dy \tag{1-5}$$

is finite (see for instance [Di Nezza et al. 2012; Maz'ya 2011]).

In view of (1-1), a natural question to begin with is whether or not the inclusion

$$u \in W^{\alpha,2+\delta}_{\mathrm{loc}}(\mathbb{R}^n) \tag{1-6}$$

holds for some $\delta > 0$, possibly depending only on the ellipticity parameters of the equation and not on the solution itself. For the definition of local fractional Sobolev spaces, see Section 2. This has been answered in a very interesting paper of Bass and Ren [2013], who consider the function

$$\Gamma(x) := \left( \int_{\mathbb{R}^n} \frac{|u(x) - u(y)|^2}{|x - y|^{n+2\alpha}} \, dy \right)^{1/2}, \tag{1-7}$$

and prove that $\Gamma \in L^{2(1+\delta)}(\mathbb{R}^n)$ for some positive $\delta$ depending only on $n$, $\alpha$, $\Lambda$ and $\delta_0$. Then (1-6) follows by characterizations of Bessel potential spaces [Dorronsoro 1985; Stein 1961]. In this paper we provide a stronger and surprising result. Indeed, we see that *for nonlocal problems the self-improvement property extends to the differentiability scale*. This means that there exists some positive $\delta \in (0, 1 - \alpha)$, depending only on $n$, $\alpha$, $\Lambda$, such that

$$u \in W^{\alpha+\delta,2+\delta}_{\mathrm{loc}}(\mathbb{R}^n). \tag{1-8}$$

This phenomenon is purely nonlocal, and has no parallel in the regularity theory of local equations, where, in order to get fractional Sobolev differentiability of $Du$, a similar fractional regularity must be assumed on the coefficient matrix $A(x)$, as for instance established in [Kuusi and Mingione 2012; Mingione 2003].

In the classical local case, measurability is, in general, not sufficient to get any gradient differentiability. To see this already in the one-dimensional case, $n = 1$, it is sufficient to consider the equation

$$\frac{d}{dx}\left( a(x) \frac{du}{dx} \right) = 0, \quad \frac{1}{\Lambda} \leq a(x) \leq \Lambda, \tag{1-9}$$

and to note that

$$x \mapsto \int_0^x \frac{dt}{a(t)}$$

is a solution with $a(\cdot)$ being any measurable function satisfying nothing but the inequalities in (1-9). It is then easy to build similar multidimensional examples.

We remark that the differentiability gain is in fact the main information in (1-8), since a standard application of the fractional Sobolev embedding theorem gives that, if $u \in W^{\alpha+\delta,2}$ for some $\delta > 0$, then (1-8) holds for some other number $\delta$. Our results actually cover a more general class of equations than the one in (1-3) and provide a full nonlocal analog of the classical higher integrability results valid in the local case. The precise statements are in the next section. Our results are a consequence of a new, fractional version of the Gehring lemma for fractional Sobolev functions that replaces the classical one valid in the local case.

We finally remark that, in recent times, there has been much attention to the regularity of solutions to nonlocal problems, especially in the basic case of kernels with measurable coefficients; see for instance [Bass and Kassmann 2005; Bjorland et al. 2012; Cabré and Cinti 2014; Cabré and Roquejoffre 2013; Caffarelli et al. 2011; Caffarelli and Silvestre 2011; Felsinger and Kassmann 2013].

**1A.** ***Higher differentiability results.*** A rather general statement concerning higher integrability for weak solutions to local problems involves nonhomogeneous equations such as

$$-\operatorname{div}(A(x)Du) = -\operatorname{div} g + f \quad \text{in } \Omega, \tag{1-10}$$

where the matrix $A(\cdot)$ has measurable coefficients and satisfies (1-2). Indeed, assuming that $g \in L^{2+\delta_0}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$ and $f \in L^{2n/(n+2)+\delta_0}_{\mathrm{loc}}(\Omega)$ hold for some $\delta_0 > 0$, it follows that there exists another positive number $\delta < \delta_0$ such that (1-1) holds. The exponent $2n/(n+2)$ is nothing but the conjugate of the Sobolev embedding exponent of $W^{1,2}$, that is, $2n/(n-2)$.

A first nonlocal analog of (1-10) is given by

$$\mathcal{E}_K(u, \eta) = \mathcal{E}_K(g, \eta) + \int_{\mathbb{R}^n} f\eta\,dx \quad \text{for all } \eta \in C^\infty_c(\mathbb{R}^n), \tag{1-11}$$

considering weak solutions $u \in W^{\alpha,2}(\mathbb{R}^n)$. The assumptions are the natural counterpart of the local ones; indeed, we take $g \in W^{\alpha+\delta_0,2}(\mathbb{R}^n)$ and

$$f \in L^{2_*+\delta_0}_{\mathrm{loc}}(\mathbb{R}^n) \tag{1-12}$$

for some $\delta_0 > 0$. The exponent $2_*$ is the conjugate of the relevant fractional Sobolev embedding exponent, that is,

$$2_* := \frac{2n}{n+2\alpha}, \quad 2^* := \frac{2n}{n-2\alpha}, \quad \frac{1}{2^*} + \frac{1}{2_*} = 1. \tag{1-13}$$

The terminology is motivated by the fractional version of the classical Sobolev embedding theorem, that is, $W^{\alpha,2} \subset L^{2^*}$. On the other hand, we recall that the essence of the structure of (1-10) lies in the fact that the right-hand side contains terms of all possible integer orders. A full extension to the fractional case then leads us to consider right-hand sides of *arbitrary fractional order*, not necessarily equal to the order of the considered nonlocal elliptic operator on the left-hand side. Moreover, since higher integrability of solutions still holds when considering monotone quasilinear equations, we will also examine nonlinear integrodifferential equations. Specifically, we will consider general equations of the type

$$\mathcal{E}^\varphi_K(u, \eta) = \mathcal{E}_H(g, \eta) + \int_{\mathbb{R}^n} f\eta\,dx \quad \text{for all } \eta \in C^\infty_c(\mathbb{R}^n). \tag{1-14}$$

The form $\mathcal{E}^\varphi_K(\cdot)$ is then defined by

$$\mathcal{E}^\varphi_K(u, \eta) := \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \varphi(u(x)-u(y))[\eta(x)-\eta(y)]K(x, y)\,dx\,dy,$$

where the Borel function $\varphi\colon \mathbb{R} \to \mathbb{R}$ satisfies

$$|\varphi(t)| \leq \Lambda |t| \quad \text{and} \quad \varphi(t)t \geq t^2 \quad \text{for all } t \in \mathbb{R}, \tag{1-15}$$

making in fact $\mathcal{E}_K^{\varphi}$ a coercive form in $W^{\alpha,2}$, and thereby (1-14) an elliptic equation. While we assume (1-4) for $K(\,\cdot\,)$, the measurable kernel $H(\,\cdot\,)$ is now assumed to satisfy

$$|H(x,y)| \leq \frac{\Lambda}{|x-y|^{n+2\beta}} \tag{1-16}$$

for $\beta \in (0,1)$. In particular, $\beta$ is also allowed to be larger than $\alpha$. Here the function $f$ is still assumed to satisfy (1-12), while the assumptions on $g$ sharply match the structure in (1-14). We actually consider two different cases; the first one is when $2\beta \geq \alpha$. In this situation we assume the existence of a positive number $\delta_0 > 0$ such that

$$g \in W^{2\beta-\alpha+\delta_0,2}(\mathbb{R}^n). \tag{1-17}$$

Needless to say, we also assume that $2\beta - \alpha + \delta_0 \in (0,1)$ to give (1-17) a sense in terms of the seminorm (1-5); this in particular implies that $\beta < \frac{1}{2}(1+\alpha)$. In the case $0 < 2\beta < \alpha$ we instead do not consider any differentiability on $g$, but only integrability:

$$g \in L^{p_0+\delta_0}(\mathbb{R}^n), \quad p_0 := \frac{2n}{n+2(\alpha-2\beta)}. \tag{1-18}$$

We then have the following *main result* of the paper:

**Theorem 1.1.** *Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ be a solution to* (1-14) *under the assumptions* (1-4) *and* (1-12)–(1-18). *Then there exists a positive number $\delta \in (0, 1-\alpha)$, depending only on $n, \alpha, \Lambda, \beta, \delta_0$, but otherwise independent of the solution $u$ and of the kernels $K(\,\cdot\,)$, $H(\,\cdot\,)$, such that $u \in W^{\alpha+\delta,2+\delta}_{\mathrm{loc}}(\mathbb{R}^n)$.*

Equation (1-11) is covered by taking $\alpha = \beta$. The optimality of the assumptions on $f$ and $g$ can be checked by considering the model equation $(-\triangle)^{\alpha}u = (-\triangle)^{\beta}g + f$, and using Fourier analysis. They sharply relate to the fractional Sobolev embedding theorem. As in the case of the classical, local Gehring lemma, explicit estimates on the exponent $\delta$ for Theorem 1.1 can be given by tracing back the dependence of the constants in the proof.

**1B.** ***Dual pairs $(\mu, U)$ and sketch of the proof.*** In order to get (1-8) we introduce here a new approach and develop a method aimed at exploiting the *hidden cancellation properties* which are intrinsic in the definition of the nonlocal seminorm (1-5). To this aim, we introduce dual pairs of measures and functions $(\mu, U)$ in $\mathbb{R}^{2n}$, proving that a version of the Gehring lemma applies to them; see Section 1C below. A natural choice would be to consider the measure generated by the density $|x-y|^{-n}$, but this would not yield a finite measure. We therefore consider a perturbation of it, i.e., the measure defined by

$$\mu(A) := \int_A \frac{dx\,dy}{|x-y|^{n-2\varepsilon}} \tag{1-19}$$

for suitably small $\varepsilon > 0$, whenever $A \subset \mathbb{R}^{2n}$ is a measurable subset. This is a locally finite, doubling Borel measure in $\mathbb{R}^{2n}$. Accordingly, for $x \neq y$, we introduce the function

$$U(x, y) := \frac{|u(x) - u(y)|}{|x - y|^{\alpha+\varepsilon}}. \tag{1-20}$$

The main point here is that the measure $\mu$ and the function $U$ are in duality when $u \in W^{\alpha,2}$ in the sense that for a function $u \in L^2(\mathbb{R}^n)$ we have that $U \in L^2(\mathbb{R}^{2n}; \mu)$ if and only if $u \in W^{\alpha,2}(\mathbb{R}^n)$. This motivates in fact the following:

**Definition 1.2.** Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ and let $\varepsilon \in (0, \frac{1}{2}\alpha)$. The pair $(\mu, U)$ defined in (1-19)–(1-20) is called a dual pair generated by the function $u$.

We then look at the higher $\mu$-integrability for $U$, proving that

$$U \in L^{2+\delta}_{\mathrm{loc}}(\mathbb{R}^{2n}; \mu) \tag{1-21}$$

for some $\delta > 0$. Now, by the very definition of $U$, we have that (1-21) implies the higher differentiability of $u$, that is, (1-8); see Section 6. This is the effect of the cancellations hidden in the definition of fractional norm in (1-5) we mentioned above. In order to prove (1-21), we shall prove decay estimates for the $\mu$-measure of the level sets of $U$. The first step consists of deriving suitable energy estimates (i.e., Caccioppoli-type inequalities) for $U$; see Theorem 3.2. We obtain a kind of reverse Hölder-type inequality, that is,

$$\left(\fint_{\mathcal{B}} U^2 \, d\mu\right)^{1/2} \lesssim \sum_{k=1}^{\infty} 2^{-k(\alpha-\varepsilon)} \left(\fint_{2^k\mathcal{B}} U^q \, d\mu\right)^{1/q} + \text{terms involving } g, f, \tag{1-22}$$

with $q < 2$; see Proposition 4.4. The estimate in (1-22) holds whenever $\mathcal{B} \equiv B \times B$ and $B \subset \mathbb{R}^n$ is a ball. Notice that if we discard from the sum above all the terms but the first one we formally obtain a reverse Hölder-type inequality similar to those that hold for solutions to local problems.

The inequality (1-22) does not seem to be sufficient to proceed, since in order to prove estimates on level sets in $\mathbb{R}^{2n}$ we need information on every ball $\mathcal{B} \subset \mathbb{R}^{2n}$, not only those of diagonal type $B \times B$. To overcome such an apparently decisive lack of information, we have to introduce an extremely delicate localization technique. Consider the level set $\{U > \lambda\}$; we use a Calderón–Zygmund-type exit-time argument in order to cover the level set with (almost disjoint) diagonal balls $B \times B$ and disjoint "off-diagonal" dyadic cubes $\mathcal{K}$:

$$\{U > \lambda\} \subset \bigcup B \times B \cup \bigcup \mathcal{K},$$

on which, for a suitably large number $L$, we have

$$\left(\fint_{B\times B} U^2 \, d\mu\right)^{1/2} \approx \lambda \quad \text{and} \quad \left(\fint_{\mathcal{K}} U^2 \, d\mu\right)^{1/2} \approx L\lambda;$$

see Sections 5A and 5F. We call the cubes $\mathcal{K}$ off-diagonal because they are "far" from the diagonal, in the sense that their distance from the diagonal is larger than their side length. The number $L$ is introduced to make the decomposition along the diagonal predominant with respect to the decomposition outside the diagonal. Indeed, the exit-time balls $B \times B$ will tend to be "larger" than the cubes $\mathcal{K}$, since they have been obtained via an exit time at a lower level $\lambda$, as shown by the first formula in the previous display.

Surprisingly enough, the fact that a cube $\mathcal{K}$ is off-diagonal allows us to prove that a reverse inequality of the type (1-22) also holds on $\mathcal{K}$ (see Lemma 5.8). This inequality, however, incorporates certain correction terms involving diagonal cubes once again. This introduces serious difficulties, since this time such cubes do not come from any exit-time argument, and there is no a priori control on them. Matching the resulting reverse inequalities with those in (1-22) is not an easy task and indeed requires an involved covering/combinatorial argument. See Sections 5I and 5J, and in particular Lemma 5.12.

The final outcome of this lengthy procedure is an inequality on level sets of $U$ — see Proposition 5.1 — that implies the *higher integrability* of $U$, together with the new reverse Hölder-type inequality reported in (1-24) below. This holds for some $\delta > 0$ that does not depend on the solution $u$. See Theorem 6.1. We have therefore proved (1-21). We also remark that treating the complete problem of Theorem 1.1 up to the sharp interpolation range described by (1-17) requires additional ideas. As a matter of fact, the exit-time arguments have to be adapted in order to realize a direct analog of the so-called good-$\lambda$ inequality principle: i.e., no maximal operator is used here. In particular, we employ a simultaneous level set analysis by using the composite quantity $\Psi(\cdot)$ in (5-1), where the number $M$ (appearing in the definition of $\Psi(\cdot)$) is used to adapt the size of the levels at the exit time. This must eventually match with the specific form of the energy estimates available for solutions.

Finally, we would like to remark that, although we are here dealing with the case of scalar, linear growth nonlocal equations, our approach is only based on energy inequalities, and therefore can be extended to more general nonlinear operators of nonlocal type; see for example [Di Castro et al. 2014a; 2014b]. This will be the object of future works.

**1C. *The fractional Gehring lemma for dual pairs.*** The classical Gehring lemma does not simply deal with solutions to equations, but, more generally, with self-improving properties of reverse Hölder-type inequalities. At the core of our approach lies in fact a new, fractional version of the Gehring lemma valid for general fractional Sobolev functions, and not only for solutions to nonlocal equations. Here is a version of it.

**Theorem 1.3** (fractional Gehring lemma). *Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ for $\alpha \in (0,1)$. Let $\varepsilon \in (0, \alpha/2)$ and let $(\mu, U)$ be the dual pair generated by $u$ in the sense of* (1-19)–(1-20) *and Definition 1.2. Assume that the following reverse Hölder-type inequality with the tail holds for every $\sigma \in (0,1)$ and for every ball $B \subset \mathbb{R}^n$:*

$$\left(\fint_{\mathcal{B}} U^2 d\mu\right)^{1/2} \le \frac{c(\sigma)}{\varepsilon^{1/q-1/2}}\left(\fint_{2\mathcal{B}} U^q d\mu\right)^{1/q} + \frac{\sigma}{\varepsilon^{1/q-1/2}}\sum_{k=2}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^q\, d\mu\right)^{1/q}, \tag{1-23}$$

*where $q \in (1,2)$ is a fixed exponent, $\mathcal{B} = B \times B$ and $c(\sigma)$ is a nonincreasing function depending on $\sigma$. Then there exists a positive number $\delta \in (0, 1-\alpha)$, depending only on $n, \alpha, \varepsilon, q$ and the function $c(\cdot)$, such that $U \in L^{2+\delta}_{\mathrm{loc}}(\mathbb{R}^{2n}; \mu)$ and $u \in W^{\alpha+\delta,2+\delta}_{\mathrm{loc}}(\mathbb{R}^n)$. Moreover, the following inequality holds whenever*

$B \subset \mathbb{R}^n$, *again for a constant* $c$ *depending only on* $n, \alpha, \varepsilon, q$ *and the function* $c(\cdot)$:

$$\left( \fint_{\mathcal{B}} U^{2+\delta} \, d\mu \right)^{1/(2+\delta)} \le c \sum_{k=1}^{\infty} 2^{-k(\alpha-\varepsilon)} \left( \fint_{2^k \mathcal{B}} U^2 \, d\mu \right)^{1/2}. \tag{1-24}$$

In the literature there are several extensions of Gehring's lemma in general settings, for instance in metric spaces equipped with a doubling Borel measure, but Theorem 1.3 is completely different. Indeed, its central feature is actually that global higher integrability information is reconstructed from reverse inequalities that do not hold on every ball in $\mathbb{R}^{2n}$, but only on diagonal ones. This is a crucial loss of information that makes Theorem 1.3 hold not for any function $U \in L^2(\mathbb{R}^{2n}; \mu)$, but rather only for dual pairs $(\mu, U)$. Moreover, the presence of the infinite series on the right-hand side of (1-23) gives to this inequality a delicate nonlocal character that adds relevant technical complications. Theorem 1.3 is a particular case of a more general result; we prefer to report this form again to make the basic ideas more transparent. A more comprehensive version including additional functions $F$ and $G$ on the right-hand side of (1-23) can be proved as well; see Theorem 6.1 below.

The results of this paper have been announced in the preliminary research report [Kuusi et al. 2014].

## 2. Preliminaries and notation

In what follows we denote by $c$ a general positive constant, possibly varying from line to line; special occurrences will be denoted by $c_1, c_2, \bar{c}_1, \bar{c}_2$ or the like. All such constants will always be *greater than or equal to one*; moreover, relevant dependencies on parameters will be emphasized using parentheses, i.e., $c_1 \equiv c_1(n, \Lambda, p, \alpha)$ means that $c_1$ depends only on $n, \Lambda, p, \alpha$. We denote by

$$B(x_0, r) \equiv B_r(x_0) := \{x \in \mathbb{R}^n : |x - x_0| < r\}$$

the open ball with center $x_0$ and radius $r > 0$; when not important, or clear from the context, we shall omit denoting the center by writing $B_r \equiv B(x_0, r)$; moreover, with $B$ being a generic ball with radius $r$, we will denote by $\sigma B$ the ball concentric to $B$ having radius $\sigma r$, $\sigma > 0$. Unless otherwise stated, different balls in the same context will have the same center. With $\mathcal{O} \subset \mathbb{R}^k$ being a measurable set with positive $\mu$-measure and with $h$ being a measurable map, we shall denote by

$$(h)_{\mathcal{O}} \equiv \fint_{\mathcal{O}} h \, d\mu := \frac{1}{\mu(\mathcal{O})} \int_{\mathcal{O}} h \, d\mu$$

its integral average. We shall need to consider integrals and functions in $\mathbb{R}^n \times \mathbb{R}^n$. In this respect, instead of dealing with the usual balls in $\mathbb{R}^{2n}$, we prefer to deal with balls generated by a different metric, that is, that relative to the norm (in $\mathbb{R}^{2n}$) defined by

$$\|(x_0, y_0)\| := \max\{|x_0|, |y_0|\}, \tag{2-1}$$

where $|\cdot|$ denotes the standard Euclidean norm in $\mathbb{R}^n$ and $x_0, y_0 \in \mathbb{R}^n$. These balls are denoted by $\mathcal{B}(x_0, y_0, \varrho)$, and are of course of the form

$$\mathcal{B}(x_0, y_0, \varrho) := B(x_0, \varrho) \times B(y_0, \varrho).$$

In the case $x_0 = y_0$ we shall also use the shorter notation $\mathcal{B}(x_0, x_0, \varrho) \equiv \mathcal{B}(x_0, \varrho)$. With obvious meaning, these will be called diagonal balls. Moreover, with $\mathcal{B}(x_0, \varrho)$ being a fixed ball, we shall also denote $\mathcal{B} \equiv \mathcal{B}(x_0, x_0, \varrho)$ when no ambiguity shall arise, and $s\mathcal{B} := \mathcal{B}(x_0, s\varrho)$ for $s > 0$. Needless to say, since they are metric balls, and actually equivalent to the standard ones in $\mathbb{R}^{2n}$, we can apply to them several tools that are available for the usual balls. For instance, we shall later on apply the classical Vitali covering lemma. It follows that

$$B_{\mathbb{R}^{2n}}((x_0, y_0), \varrho) = \{(x_0, y_0) \in \mathbb{R}^{2n} : |(x_0, y_0)| < \varrho\} \subset \mathcal{B}(x_0, y_0, \varrho).$$

Accordingly, we shall denote

$$\mathrm{Diag} := \{(x, x) \in \mathbb{R}^{2n} : x \in \mathbb{R}^n\}.$$

If $A$ is a finite set, the symbol $\#A$ denotes the number of its elements. We shall very often use the elementary inequality

$$2^{2\beta k} \sum_{j=k-1}^{\infty} 2^{-2\beta j} \le \frac{8}{\beta} \quad \text{for } \beta \in (0, 1] \text{ and } k \ge 1. \tag{2-2}$$

Finally, the local fractional Sobolev spaces are defined via the Gagliardo seminorm

$$[u]_{s,\gamma}(\Omega) := \left( \int_\Omega \int_\Omega \frac{|u(x) - u(y)|^\gamma}{|x - y|^{n+\gamma s}} \, dx \, dy \right)^{1/\gamma} \tag{2-3}$$

for $\gamma \ge 1$ and $s \in (0, 1)$. A function $u \in L^\gamma_{\mathrm{loc}}(\mathbb{R}^n)$ belongs to $W^{s,\gamma}_{\mathrm{loc}}(\mathbb{R}^n)$ if $[u]_{s,\gamma}(\Omega)$ is finite whenever $\Omega$ is an open bounded subset of $\mathbb{R}^n$.

The following two lemmas report some classical Poincaré–Sobolev-type inequalities valid in the fractional setting; the proof of the first is exactly the one in [Mingione 2003], for the second we refer to [Kassmann 2009]. See also [Di Nezza et al. 2012; Maz'ya 2011].

**Lemma 2.1** (fractional Poincaré inequality). *Let $v \in L^p(B)$, with $B \subset \mathbb{R}^n$ being a ball of radius $r$, and let $\alpha$ be a real number such that such that $n + p\alpha \ge 0$; then the following inequality holds*:

$$\fint_B |v - (v)_B|^p \, dx \le c r^{p\alpha} \int_B \fint_B \frac{|v(x) - v(y)|^p}{|x - y|^{n+p\alpha}} \, dx \, dy.$$

This inequality in particular applies when $v \in W^{\alpha,p}(B)$, and in this case the quantity on the right-hand side is finite.

**Lemma 2.2** (fractional Poincaré–Sobolev inequality). *Let $v \in W^{\alpha,p}(B)$, for $\alpha \in (0, 1)$, where $B \subset \mathbb{R}^n$ is a ball of radius $r$, or a cube of diameter $r$. If $p\alpha < n$, then the following inequality holds for a constant $c$ depending only on $n$, $\alpha$*:

$$\left( \fint_B |v - (v)_B|^{p^*} \, dx \right)^{1/p^*} \le c r^\alpha \left( \int_B \fint_B \frac{|v(x) - v(y)|^p}{|x - y|^{n+p\alpha}} \, dx \, dy \right)^{1/p},$$

*where $p^* := np/(n - p\alpha)$.*

With $2_*$ being the exponent defined in (1-13), an immediate consequence of the previous lemma is the following inequality, that we report since it will be used several times:

$$\left(\fint_B |v-(v)_B|^2\,dx\right)^{1/2} \le cr^{\alpha}\left(\int_B\fint_B \frac{|v(x)-v(y)|^{2_*}}{|x-y|^{n+2_*\alpha}}\,dx\,dy\right)^{1/2_*}. \tag{2-4}$$

Moreover, if $v$ is compactly supported in $B$, then $v-(v)_B$ above can be replaced by $v$.

## 3. The Caccioppoli inequality

**3A.** ***Preliminary reformulation of the assumptions.*** We start with the assumptions made on $g$, that is, (1-17)–(1-18). In order to give a unified proof for the two cases $2\beta \ge \alpha$ and $2\beta < \alpha$, and to simplify certain computations, we shall make a few preliminary reductions and will restate the assumptions in a more convenient way. First of all let us consider the case $2\beta \ge \alpha$, when (1-17) is in force. Let us notice that, eventually reducing the value of $\delta_0$, and in particular taking $\delta_0 \le \alpha/40$, (1-17) implies the existence of exponents $p, \gamma$ and $\delta_1 > 0$, such that $g \in W^{\gamma(1+\delta_1),p(1+\delta_1)}(\mathbb{R}^n)$ and

$$2\beta > \gamma > 2\beta-\alpha, \quad 2 > p > \frac{2n}{n+2(\gamma-2\beta+\alpha)}, \quad \delta_1 \le \frac{\alpha}{4n}. \tag{3-1}$$

Indeed, let us set $\gamma = 2\beta-\alpha+\delta_0/2$ and recall that $W^{2\beta-\alpha+\delta_0,2}$ embeds in $W^{\gamma(1+\delta_1),p(1+\delta_1)}$ whenever $2\beta-\alpha+\delta_0-n/2 = \gamma(1+\delta_1)-n/[p(1+\delta_1)]$. A lengthy computation then shows that any choice of $p$ as above and $\delta_1 \le 1$ satisfying the inequalities

$$\frac{(1+\delta_1)\delta_0}{n+2\gamma(1+\delta_1)} < \delta_1 < \frac{(2+\delta_1)\delta_0}{n+2\gamma(1+\delta_1)}$$

matches the conditions in (3-1). We now consider the case $2\beta < \alpha$, when (1-18) is in force. In this case we can instead assume the existence of numbers $p > 1$ and $\delta_1 > 0$ such that

$$g \in L^{p(1+\delta_1)}_{\mathrm{loc}}(\mathbb{R}^n), \quad p > \frac{2n}{n+2(\alpha-2\beta)}. \tag{3-2}$$

Let us now unify the previous conditions. In the case $2\beta \ge \alpha$ we clearly have that

$$\int_B\int_B \frac{|g(x)-g(y)|^{p(1+\delta_1)}}{|x-y|^{n+p(1+\delta_1)^2\gamma}}\,dx\,dy + \int_B\int_B \frac{|g(x)-g(y)|^{p}}{|x-y|^{n+p\gamma}}\,dx\,dy < \infty \tag{3-3}$$

for every ball $B \subset \mathbb{R}^n$. This comes by the definition of the space $W^{\gamma(1+\delta_1),p(1+\delta_1)}$. On the other hand, when $2\beta < \alpha$, then assumptions (1-18) do not involve any number $\gamma$. Thanks to the lower bound on $p$ in (1-18), we can find a negative number $\gamma$, such that $|\gamma| \in (0, \frac{1}{10})$ is small enough to still verify (3-1). In this case we note that

$$\begin{aligned}\int_B\int_B \frac{|g(x)-g(y)|^{p(1+\delta_1)}}{|x-y|^{n+p(1+\delta_1)^2\gamma}}\,dx\,dy &\le \int_B\int_B \frac{(|g(x)|+|g(y)|)^{p(1+\delta_1)}}{|x-y|^{n+p(1+\delta_1)^2\gamma}}\,dx\,dy \\ &\le \frac{cr^{-p(1+\delta_1)^2\gamma}}{-\gamma}\int_B |g|^{p(1+\delta_1)}\,dx < \infty,\end{aligned} \tag{3-4}$$

where $r$ denotes the radius of $B$; a similar estimate follows for the second quantity in (3-3). Summarizing, in the rest of the paper we shall always assume that (3-1) and (3-3) hold. In the case $2\beta < \alpha$ the number $\gamma$ is negative.

**Remark 3.1.** We shall denote by $c_b$ a constant that depends on $n, \alpha, \Lambda, p, \beta, \gamma$ and exhibits the blow-up behavior

$$\lim_{p \to 2n/(n+2(\gamma-2\beta+\alpha))} c_b = \infty. \tag{3-5}$$

**3B.** ***The Caccioppoli estimate.*** The Caccioppoli-type inequality stated in the next theorem is an essential tool in the proof of Theorem 1.1.

**Theorem 3.2.** *Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ be a solution to* (1-14) *under the assumptions of Theorem 1.1*; *in particular*, (3-1) *and* (3-3) *are in force. Let $B \equiv B(x_0, r) \subset \mathbb{R}^n$ be a ball, and let $\psi \in C_c^\infty(B(x_0, \frac{3}{4}r))$ be a cutoff function such that $0 \le \psi \le 1$ and $|D\psi| \le c(n)/r$. Then the Caccioppoli-type inequality*

$$\begin{aligned}\int_B \int_B \frac{|u(x)\psi(x) - u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}} dx\, dy \\ \le \frac{c}{r^{2\alpha}} \int_B |u(x)|^2\, dx + c\int_{\mathbb{R}^n \setminus B} \frac{|u(y)|}{|x_0 - y|^{n+2\alpha}}\, dy \int_B |u(x)|\, dx + cr^{n+2\alpha}\biggl(\fint_B |f(x)|^{2_*}\, dx\biggr)^{2/2_*} \\ + c_b r^{n+2(\gamma-2\beta+\alpha)}\left[\sum_{k=0}^{\infty} 2^{(\gamma-2\beta)k}\biggl(\int_{2^k B} \fint_{2^k B} \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\, dx\, dy\biggr)^{1/p}\right]^2\end{aligned} \tag{3-6}$$

*holds for a constant $c \equiv c(n, \Lambda, \alpha)$ which is independent of $p$, and a constant $c_b \equiv c_b(n, \Lambda, \alpha, \beta, \gamma, p)$. The constant $c_b$ exhibits the behavior described in* (3-5); *moreover*, *all the terms appearing on the right-hand side of* (3-6) *are finite.*

*Proof.* In the weak formulation

$$\begin{aligned}\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \varphi(u(x) - u(y))[\eta(x) - \eta(y)]K(x, y)\, dx\, dy \\ = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} [g(x) - g(y)][\eta(x) - \eta(y)]H(x, y)\, dx\, dy + \int_{\mathbb{R}^n} f\eta\, dx,\end{aligned} \tag{3-7}$$

we choose $\eta = u\psi^2$, where $\psi \in C_c^\infty(B)$ is the cutoff function coming from the statement. By a density argument, $\eta$ is an admissible test function. Then we have

$$\begin{aligned}I_1 + I_2 + I_3 := \int_B \int_B \varphi(u(x) - u(y))[u(x)\psi^2(x) - u(y)\psi^2(y)]K(x, y)\, dx\, dy \\ + \int_{\mathbb{R}^n \setminus B} \int_B \varphi(u(x) - u(y))u(x)\psi^2(x)K(x, y)\, dx\, dy \\ - \int_B \int_{\mathbb{R}^n \setminus B} \varphi(u(x) - u(y))u(y)\psi^2(y)K(x, y)\, dx\, dy\end{aligned}$$

$$
\begin{aligned}
&= \int_B \int_B [g(x)-g(y)][u(x)\psi^2(x)-u(y)\psi^2(y)]H(x,y)\,dx\,dy \\
&\quad + \int_{\mathbb{R}^n\setminus B}\int_B [g(x)-g(y)]u(x)\psi^2(x)H(x,y)\,dx\,dy \\
&\quad + \int_B \int_{\mathbb{R}^n\setminus B} [g(y)-g(x)]u(y)\psi^2(y)H(x,y)\,dx\,dy \\
&\quad + \int_B f(x)u(x)\psi^2(x)\,dx =: J_1+J_2+J_3+J_4. \qquad (3\text{-}8)
\end{aligned}
$$

We proceed in estimating the various pieces stemming from this identity.

**Estimation of** $I_1$. Let us first consider the case in which $\psi(x)\ge\psi(y)$. Then we write

$$
\begin{aligned}
&\varphi(u(x)-u(y))[u(x)\psi^2(x)-u(y)\psi^2(y)] \\
&\qquad = \varphi(u(x)-u(y))[u(x)-u(y)]\psi^2(x)+\varphi(u(x)-u(y))u(y)[\psi^2(x)-\psi^2(y)]. \qquad (3\text{-}9)
\end{aligned}
$$

Applying Young's inequality and recalling the first inequality in (1-15), we have

$$
\begin{aligned}
\varphi(u(x)-u(y))u(y)[\psi^2(x)-\psi^2(y)] &= \varphi(u(x)-u(y))u(y)[\psi(x)-\psi(y)][\psi(x)+\psi(y)] \\
&\ge -2|\varphi(u(x)-u(y))||u(y)||\psi(x)-\psi(y)|\psi(x) \\
&\ge -\tfrac{1}{2}|u(x)-u(y)|^2\psi^2(x)-2\Lambda^2u^2(y)[\psi(x)-\psi(y)]^2.
\end{aligned}
$$

Combining the content of the last two displays, and using this time the second inequality in (1-15), yields

$$
\varphi(u(x)-u(y))[u(x)\psi^2(x)-u(y)\psi^2(y)] \ge \tfrac{1}{2}[u(x)-u(y)]^2\psi^2(x)-2\Lambda^2u^2(y)[\psi(x)-\psi(y)]^2.
$$

Now we consider the case in which $\psi(y)\ge\psi(x)$, and we similarly write

$$
\begin{aligned}
&\varphi(u(x)-u(y))[u(x)\psi^2(x)-u(y)\psi^2(y)] \\
&\qquad = \varphi(u(x)-u(y))[u(x)-u(y)]\psi^2(y)+\varphi(u(x)-u(y))u(x)[\psi^2(x)-\psi^2(y)].
\end{aligned}
$$

Proceeding similarly to the case $\psi(x)\ge\psi(y)$, we arrive at

$$
\varphi(u(x)-u(y))[u(x)\psi^2(x)-u(y)\psi^2(y)] \ge \tfrac{1}{2}[u(x)-u(y)]^2\psi^2(y)-2\Lambda^2u^2(x)[\psi(x)-\psi(y)]^2.
$$

In any case, using also (1-4), we conclude that

$$
I_1 \ge \frac{1}{c}\int_B\int_B \frac{|u(x)-u(y)|^2}{|x-y|^{n+2\alpha}}\max\{\psi^2(x),\psi^2(y)\}\,dx\,dy - c\int_B\int_B |u(x)|^2\frac{|\psi(x)-\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy,
$$

where $c$ depends on $\Lambda$. Moreover, by noticing that

$$
[u(x)\psi(x)-u(y)\psi(y)]^2 \le 2[u(x)(\psi(x)-\psi(y))]^2+2[\psi(y)(u(x)-u(y))]^2
$$

and integrating, we conclude that

$$
I_1 \ge \frac{1}{c}\int_B\int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy - c\int_B\int_B |u(x)|^2\frac{|\psi(x)-\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy. \qquad (3\text{-}10)
$$

**Estimation of $I_2$ and $I_3$.** The estimation of the terms $I_2$ and $I_3$ is similar. Indeed, as for $I_2$, we start by observing that a direct computation yields

$$[u(x)-u(y)]u(x)\psi^2(x)K(x,y) \geq -\Lambda \frac{|u(x)||u(y)|\psi^2(x)}{|x-y|^{n+2\alpha}},$$

and therefore, by (1-15) we obtain (we can assume without loss of generality that $u(x) \neq u(y)$) that

$$\begin{aligned}\varphi(u(x)-u(y))u(x)\psi^2(x)K(x,y) &\geq -\Lambda \left|\frac{\varphi(u(x)-u(y))}{u(x)-u(y)}\right| \cdot \frac{|u(x)|\,|u(y)|\psi^2(x)}{|x-y|^{n+2\alpha}} \\ &\geq -\Lambda^2 \frac{|u(x)|\,|u(y)|\psi^2(x)}{|x-y|^{n+2\alpha}}.\end{aligned}$$

Similarly, we obtain

$$-\varphi(u(x)-u(y))u(y)\psi^2(y)K(x,y) \geq -\Lambda^2 \frac{|u(x)|\,|u(y)|\psi^2(y)}{|x-y|^{n+2\alpha}}.$$

We then estimate

$$\begin{aligned}I_2+I_3 &\geq -c\int_{\mathbb{R}^n\setminus B}\int_B \frac{|u(x)||u(y)|\psi^2(x)}{|x-y|^{n+2\alpha}}\,dx\,dy \\ &\geq -c \sup_{z\in \operatorname{supp}\psi}\int_{\mathbb{R}^n\setminus B}\frac{|u(y)|}{|z-y|^{n+2\alpha}}\,dy \int_B |u(x)|\psi^2(x)\,dx \\ &\geq -c\int_{\mathbb{R}^n\setminus B}\frac{|u(y)|}{|x_0-y|^{n+2\alpha}}\,dy\int_B |u(x)|\psi^2(x)\,dx. \qquad (3\text{-}11)\end{aligned}$$

Here we have used the fact that, since $\psi$ is supported in $B(x_0, \frac{3}{4}r)$, we have

$$\frac{|x_0-y|}{|z-y|} \leq 1+\frac{|x_0-z|}{|z-y|} \leq 4 \qquad (3\text{-}12)$$

whenever $z \in \operatorname{supp}\psi$ and $y \in \mathbb{R}^n \setminus B$.

**Estimation of $J_4$.** The fractional Sobolev inequality yields

$$\begin{aligned}J_4 &\leq cr^n \left(\fint_B |u(x)\psi(x)|^{2^*}\,dx\right)^{1/2^*}\left(\fint_B |f(x)|^{2_*}\,dx\right)^{1/2_*} \\ &\leq cr^{n/2+\alpha}\left(\int_B\int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy\right)^{1/2}\left(\fint_B |f(x)|^{2_*}\,dx\right)^{1/2_*},\end{aligned}$$

so that, applying Young's inequality with $\sigma \in (0,1)$, we have

$$J_4 \leq \frac{c}{\sigma}r^{n+2\alpha}\left(\fint_B |f(x)|^{2_*}\,dx\right)^{2/2_*} + \sigma\int_B\int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy. \qquad (3\text{-}13)$$

The constant $c$ depends only on $n, \alpha$.

**Estimation of** $J_1$. We write

$$u(x)\psi^2(x) - u(y)\psi^2(y) = [u(x)\psi(x) - u(y)\psi(y)]\psi(y) + u(x)\psi(x)[\psi(x) - \psi(y)].$$

Therefore, using that $\psi \leq 1$ together with (1-16), we have

$$\begin{aligned} J_1 &\leq \Lambda \int_B \int_B \frac{|g(x)-g(y)|}{|x-y|^{2\beta}} |u(x)\psi(x) - u(y)\psi(y)| \frac{dx\,dy}{|x-y|^n} \\ &\quad + \Lambda \int_B \int_B \frac{|g(x)-g(y)|}{|x-y|^{2\beta}} |u(x)\psi(x)||\psi(x) - \psi(y)| \frac{dx\,dy}{|x-y|^n} \\ &=: J_{1.1} + J_{1.2}. \end{aligned}$$

In turn, we estimate $J_{1.1}$ and $J_{1.2}$ separately. Recalling (3-1), we now set

$$t := 1 - \frac{2\beta - \gamma}{\alpha} \quad \text{and} \quad s := \frac{n}{\alpha}\left[\frac{1}{p} - \frac{1}{2}\right]. \tag{3-14}$$

Observe that $0 < t \leq 1$ if and only if $2\beta - \alpha < \gamma \leq 2\beta$. Then we notice that

$$\begin{aligned} 2\beta \geq \gamma \quad \text{and} \quad 2 > p > \frac{2n}{n + 2(\gamma - 2\beta + \alpha)} \quad &\Longrightarrow \quad 2 > p > \frac{2n}{n+2\alpha} = 2_* \\ &\Longrightarrow \quad 0 < s < 1, \end{aligned} \tag{3-15}$$

and moreover

$$p > \frac{2n}{n + 2(\gamma - 2\beta + \alpha)} \quad \Longrightarrow \quad 0 < s < t. \tag{3-16}$$

We also record the identity $\alpha t = \gamma - (2\beta - \alpha)$. Let us now write

$$J_{1.1} = cr^n \int_B \fint_B \left[ r^{\alpha t} \frac{|g(x)-g(y)|}{|x-y|^{2\beta-\alpha+t\alpha}} \right] \left[ \frac{|u(x)\psi(x) - u(y)\psi(y)|}{|x-y|^{\alpha}} \right]^{1-s} \left[ r^{-\alpha t/s} \frac{|u(x)\psi(x) - u(y)\psi(y)|}{|x-y|^{\alpha(1-t/s)}} \right]^s$$

The definitions in (3-14) imply that

$$\frac{1-s}{2} + \frac{s}{2^*} + \frac{1}{p} = 1;$$

therefore, applying Hölder's inequality with the corresponding choice of the exponents, we have

$$\begin{aligned} J_{1.1} \leq cr^{n+\alpha t} &\left( \int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p(2\beta-\alpha+t\alpha)}} \,dx\,dy \right)^{1/p} \left( \int_B \fint_B \frac{|u(x)\psi(x) - u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}} \,dx\,dy \right)^{(1-s)/2} \\ &\times \left( r^{-2^*\alpha t/s} \int_B \fint_B \frac{|u(x)\psi(x) - u(y)\psi(y)|^{2^*}}{|x-y|^{n+2^*\alpha(1-t/s)}} \,dx\,dy \right)^{s/2^*}. \end{aligned} \tag{3-17}$$

Before going on, let us estimate the last integral:

$$\int_B \fint_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^{2^*}}{|x-y|^{n+2^*\alpha(1-t/s)}}\,dx\,dy \le 2^{2^*-1}\int_B \fint_B \frac{|u(x)\psi(x)|^{2^*}}{|x-y|^{n+2^*\alpha(1-t/s)}}\,dx\,dy$$
$$\le \frac{cr^{-2^*\alpha(1-t/s)}}{t-s}\fint_B |u(x)\psi(x)|^{2^*}\,dx$$
$$\le \frac{cr^{2^*\alpha t/s}}{t-s}\left(\int_B \fint_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy\right)^{2^*/2}. \quad (3\text{-}18)$$

Plugging the inequality into (3-17) yields

$$J_{1.1} \le cr^{n/2+\alpha t}\left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p(2\beta-\alpha+t\alpha)}}\,dx\,dy\right)^{1/p}\left(\int_B \int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy\right)^{1/2}.$$

Using Young's inequality, and keeping in mind that $\alpha t = \gamma - (2\beta - \alpha)$, leads to

$$J_{1.1} \le \frac{c}{\sigma} r^{n+2(\gamma-2\beta+\alpha)}\left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{2/p} + \sigma \int_B \int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy$$

whenever $\sigma \in (0, 1)$. The constant $c$ depends only on $n, \alpha, \Lambda, \beta, \gamma, p$. We then continue with the estimation of $J_{1.2}$. Upon setting $\eta := \frac{1}{2}(1-\alpha)$, using Hölder's inequality with conjugate exponents $(2^*, 2_*)$ we have

$$J_{1.2} \le c\|D\psi\|_{L^\infty} r^n \int_B \fint_B \frac{|g(x)-g(y)|}{|x-y|^{2\beta-1+\eta}}\,\frac{|u(x)\psi(x)|}{|x-y|^{-\eta}}\,\frac{dx\,dy}{|x-y|^n}$$
$$\le c\|D\psi\|_{L^\infty} r^n \left(\int_B \fint_B \frac{|g(x)-g(y)|^{2_*}}{|x-y|^{2_*(2\beta-1+\eta)}}\,\frac{dx\,dy}{|x-y|^n}\right)^{1/2_*}\left(\int_B \fint_B \frac{|u(x)\psi(x)|^{2^*}}{|x-y|^{-2^*\eta}}\,\frac{dx\,dy}{|x-y|^n}\right)^{1/2^*}.$$

In turn, by Lemma 2.2 (see also the remark following it) we have

$$\int_B \fint_B \frac{|u(x)\psi(x)|^{2^*}}{|x-y|^{-2^*\eta}}\,\frac{dx\,dy}{|x-y|^n} \le \frac{cr^{2^*\eta}}{1-\alpha}\fint_B |u(x)\psi(x)|^{2^*}\,dx$$
$$\le cr^{2^*(\eta+\alpha)}\left(\int_B \fint_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy\right)^{2^*/2}$$

and, recalling that $p > 2_*$ by (3-15), we proceed with

$$\int_B \fint_B \frac{|g(x)-g(y)|^{2_*}}{|x-y|^{2_*(2\beta-1+\eta)}}\,\frac{dx\,dy}{|x-y|^n} = \int_B \fint_B \left(\frac{|g(x)-g(y)|}{|x-y|^\gamma}\right)^{2_*}\frac{1}{|x-y|^{2_*(2\beta-1+\eta-\gamma)}}\,\frac{dx\,dy}{|x-y|^n}$$
$$\le \left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{2_*/p}$$
$$\times \left(\int_B \fint_B \frac{1}{|x-y|^{\frac{2_*(2\beta-1+\eta-\gamma)p}{p-2_*}}}\,\frac{dx\,dy}{|x-y|^n}\right)^{1-2_*/p}$$
$$\le \frac{cr^{-2_*(2\beta-1+\eta-\gamma)}}{\gamma-2\beta+\alpha}\left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{2_*/p},$$

where of course we used that $2\beta-1+\eta-\gamma=2\beta-\frac{1}{2}-\frac{1}{2}\alpha-\gamma<2\beta-\alpha-\gamma<0$ due to $\eta:=\frac{1}{2}(1-\alpha)$ and (3-1). Connecting the estimates in the last three displays yields

$$J_{1.2} \le c\|D\psi\|_{L^\infty} r^{n/2+\gamma-2\beta+\alpha+1}\left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p} \times\left(\int_B\int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy\right)^{1/2}.$$

Again using Young's inequality, we conclude that

$$J_{1.2} \le \frac{c}{\sigma} r^2\|D\psi\|^2_{L^\infty} r^{n+2(\gamma-2\beta+\alpha)}\left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{2/p} +\sigma\int_B\int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy,$$

which holds whenever $\sigma\in(0,1)$. Gathering together the estimates found for $J_{1.1}$ and $J_{1.2}$, and using that $r^2\|D\psi\|^2_{L^\infty}\le c(n)$, gives

$$J_1 \le \frac{c}{\sigma} r^{n+2(\gamma-2\beta+\alpha)}\left(\int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{2/p} +2\sigma\int_B\int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy. \tag{3-19}$$

The constant $c$ depends on $n,\alpha,\Lambda,\beta,\gamma,p$.

**Estimation of $J_2$ and** $J_3$. The estimation of the two terms is completely similar, and we therefore confine ourselves to estimating $J_2$. Using (1-16) we have

$$\begin{aligned} J_2 &\le \Lambda\int_{\mathbb{R}^n\setminus B}\int_B \frac{|g(x)-(g)_B|}{|x-y|^{n+2\beta}}|u(x)|\psi^2(x)\,dx\,dy+\Lambda\int_{\mathbb{R}^n\setminus B}\int_B \frac{|g(y)-(g)_B|}{|x-y|^{n+2\beta}}|u(x)|\psi^2(x)\,dx\,dy\\ &=: J_{2.1}+J_{2.2}.\end{aligned}$$

In turn we estimate the two resulting terms. Using that $p\ge 2_*$ by (3-15), we have

$$\begin{aligned} J_{2.1} &\le c\sup_{z\in\operatorname{supp}\psi}\int_{\mathbb{R}^n\setminus B}\frac{dy}{|z-y|^{n+2\beta}}\int_B |g(x)-(g)_B||u(x)|\psi(x)\,dx\\ &\le cr^n\sup_{z\in\operatorname{supp}\psi}\int_{\mathbb{R}^n\setminus B}\frac{dy}{|z-y|^{n+2\beta}}\left(\fint_B |g(x)-(g)_B|^{2_*}\,dx\right)^{1/2_*}\left(\fint_B |u(x)\psi(x)|^{2^*}\,dx\right)^{1/2^*}\\ &\le cr^n\sup_{z\in\operatorname{supp}\psi}\int_{\mathbb{R}^n\setminus B}\frac{dy}{|z-y|^{n+2\beta}}\left(\fint_B |g(x)-(g)_B|^{p}\,dx\right)^{1/p}\left(\fint_B |u(x)\psi(x)|^{2^*}\,dx\right)^{1/2^*}\\ &\le cr^{n/2+\gamma-2\beta+\alpha}\sup_{z\in\operatorname{supp}\psi}\int_{\mathbb{R}^n\setminus B}\frac{r^{2\beta}dy}{|z-y|^{n+2\beta}}\cdot\left(\int_B\fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\\ &\qquad\times\left(\int_B\int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy\right)^{1/2}.\end{aligned}$$

Therefore, using Young's inequality, we have

$$J_{2.1} \le \frac{c}{\sigma} r^{n+2(\gamma-2\beta+\alpha)} \left( \int_B \fint_B \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy \right)^{2/p} + \sigma \int_B \int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy,$$

where we have also used that $\psi \equiv 0$ outside $B(x_0, \frac{3}{4}r)$, and therefore (3-12), to estimate

$$\sup_{z\in \operatorname{supp}\psi} \int_{\mathbb{R}^n\setminus B} \frac{r^{2\beta}\,dy}{|z-y|^{n+2\beta}} \le c(n,\beta).$$

In order to estimate $J_{2.2}$ we need another splitting over annuli. Recalling again that $\psi \le 1$ and that $\psi \equiv 0$ outside $B(x_0, \frac{3}{4}r)$, we have

$$\begin{aligned} J_{2.2} &\le c \sum_{j=0}^{\infty} \int_{2^{j+1}B\setminus 2^j B} \int_B \frac{|g(y)-(g)_B|}{|x-y|^{n+2\beta}} |u(x)|\psi^2(x)\,dx\,dy \\ &\le c r^n \sum_{j=0}^{\infty} (2^j r)^{-2\beta} \fint_{2^{j+1}B} |g(y)-(g)_B|\,dy \fint_B |u(x)\psi(x)|\,dx \\ &\le c r^n \sum_{j=0}^{\infty} (2^j r)^{-2\beta} \left( \fint_{2^{j+1}B} |g(y)-(g)_B|^p\,dy \right)^{1/p} \fint_B |u(x)\psi(x)|\,dx. \end{aligned} \tag{3-20}$$

The estimation of $J_{2.2}$ needs again a splitting; we start with the telescoping summation

$$\begin{aligned} &\left( \fint_{2^{j+1}B} |g(y)-(g)_B|^p\,dy \right)^{1/p} \\ &\quad\le \left( \fint_{2^{j+1}B} |g(y)-(g)_{2^{j+1}B}|^p\,dy \right)^{1/p} + \sum_{k=0}^{j} |(g)_{2^{k+1}B} - (g)_{2^k B}| \\ &\quad\le \left( \fint_{2^{j+1}B} |g(y)-(g)_{2^{j+1}B}|^p\,dy \right)^{1/p} + \sum_{k=0}^{j} \left( \fint_{2^{k+1}B} |g(y)-(g)_{2^k B}|^p\,dy \right)^{1/p} \\ &\quad\le 2 \sum_{k=0}^{j+1} \left( \fint_{2^k B} |g(y)-(g)_{2^k B}|^p\,dy \right)^{1/p}. \end{aligned} \tag{3-21}$$

Then an application of the fractional Poincaré inequality in Lemma 2.1 yields

$$\left( \fint_{2^{j+1}B} |g(y)-(g)_B|^p\,dy \right)^{1/p} \le c \sum_{k=0}^{j+1} (2^k r)^{\gamma} \left( \int_{2^k B} \fint_{2^k B} \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy \right)^{1/p}.$$

Merging the content of the last display with the one of (3-20) gives

$$J_{2.2} \le c r^n \left[ \sum_{j=0}^{\infty} \sum_{k=0}^{j+1} (2^j r)^{-2\beta} (2^k r)^{\gamma} \left( \int_{2^k B} \fint_{2^k B} \frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy \right)^{1/p} \right] \fint_B |u(x)\psi(x)|\,dx.$$

We now manipulate the content of the square brackets above using discrete Fubini's theorem:

$$
\begin{aligned}
&\sum_{j=0}^{\infty}\sum_{k=0}^{j+1}(2^j r)^{-2\beta}(2^k r)^{\gamma}\left(\int_{2^k B}\fint_{2^k B}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\\
&\quad = r^{\gamma-2\beta}\left(\int_{B}\fint_{B}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\sum_{j=0}^{\infty}2^{-2\beta j}\\
&\qquad + r^{\gamma-2\beta}\sum_{k=1}^{\infty}2^{\gamma k}\left(\int_{2^k B}\fint_{2^k B}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\sum_{j=k-1}^{\infty}2^{-2\beta j}\\
&\quad \le \frac{c r^{\gamma-2\beta}}{\beta}\sum_{k=0}^{\infty}2^{(\gamma-2\beta)k}\left(\int_{2^k B}\fint_{2^k B}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}.
\end{aligned}
$$

We remark that in the previous display we have used the elementary inequality in (2-2). All in all we have, by using also Hölder's inequality and Lemma 2.1, that

$$
\begin{aligned}
J_{2.2} &\le c r^{n+\gamma-2\beta}\sum_{k=0}^{\infty}2^{(\gamma-2\beta)k}\left(\int_{2^k B}\fint_{2^k B}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\left(\fint_{B}|u(x)\psi(x)|^{2^*}\,dx\right)^{1/2^*}\\
&\le c r^{n/2+\gamma-2\beta+\alpha}\sum_{k=0}^{\infty}2^{(\gamma-2\beta)k}\left(\int_{2^k B}\fint_{2^k B}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\\
&\quad \times\left(\int_{B}\int_{B}\frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy\right)^{1/2}.
\end{aligned}
$$

Finally, using Young's inequality we conclude that

$$
\begin{aligned}
J_{2.2} \le \frac{c}{\sigma} r^{n+2(\gamma-2\beta+\alpha)}&\left[\sum_{k=0}^{\infty}2^{(\gamma-2\beta)k}\left(\int_{2^k B}\fint_{2^k B}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\right]^2\\
&+\sigma\int_{B}\int_{B}\frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy
\end{aligned}
$$

whenever $\sigma \in (0, 1)$. Connecting the inequalities found for $J_{1.2}$ and $J_{2.2}$, and again recalling that $J_3$ can be estimated in a completely similar way, we have

$$
\begin{aligned}
J_2 + J_3 \le \frac{c}{\sigma} r^{n+2(\gamma-2\beta+\alpha)}&\left[\sum_{k=0}^{\infty}2^{(\gamma-2\beta)k}\left(\int_{2^k B}\fint_{2^k B}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\right]^2\\
&+4\sigma\int_{B}\int_{B}\frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}dx\,dy. \qquad (3\text{-}22)
\end{aligned}
$$

The constant $c$ depends on $n, \Lambda, \alpha, \beta, \gamma, p$.

**Reabsorbing terms.** Inserting the estimates for the terms $I_i$ and $J_i$ into (3-8), we conclude that

$$\begin{aligned}
\frac{1}{c}\int_B\int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy&\\
\le 7\sigma\int_B\int_B \frac{|u(x)\psi(x)-u(y)\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy &+ c\int_B\int_B |u(x)|^2\frac{|\psi(x)-\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy\\
+c\int_{\mathbb{R}^n\setminus B}\frac{|u(y)|}{|x_0-y|^{n+2\alpha}}\,dy\int_B |u(x)|\psi^2(x)\,dx &+ \frac{c}{\sigma}r^{n+2\alpha}\biggl(\fint_B |f(x)|^{2_*}\,dx\biggr)^{2/2_*}\\
+\frac{c_b}{\sigma}r^{n+2(\gamma-2\beta+\alpha)}\Biggl[\sum_{k=0}^{\infty}2^{(\gamma-2\beta)k}\biggl(\int_{2^kB}\fint_{2^kB}&\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\biggr)^{1/p}\Biggr]^2.
\end{aligned}$$

The constant $c$ depends only on $n, \alpha, \Lambda$, and the constant $c_b$ depends only on $n, \Lambda, \alpha, \beta, \gamma, p$. Now, taking $\sigma = 1/(14c)$ and reabsorbing terms finishes the proof, together with the estimate

$$\begin{aligned}
\int_B\int_B |u(x)|^2\frac{|\psi(x)-\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy &\le \|D\psi\|_\infty\int_B|u(x)|^2\int_{B_{2r}(x)}|x-y|^{-n+2-2\alpha}\,dy\,dx\\
&=\frac{c(n)}{1-\alpha}\|D\psi\|_\infty r^{2-2\alpha}\int_B|u(x)|^2\,dx\\
&\le\frac{c(n)}{1-\alpha}\frac{1}{r^{2\alpha}}\int_B|u(x)|^2\,dx.
\end{aligned}$$

The finiteness of the terms appearing on the right in (3-6) follows directly from the fact that $u \in W^{\alpha,2}(\mathbb{R}^n)$ and from Section 4C below. This completes the proof of Theorem 3.2. □

**Remark 3.3.** In the statement of Theorem 3.2, one can replace $u$ with $u-(u)_B$ by testing with $(u-(u)_B)\psi^2$ instead of $u\psi^2$.

**Remark 3.4.** All the constants denoted by $c$ that appear in Theorem 3.2 blow up as $\alpha \to 0$ or as $\alpha \to 1$. The blow-up of the constant $c_b$ is more peculiar, and it is as in (3-5). This appears for instance in estimate (3-18), as in this case $s \to t$; see (3-16). In terms of (1-12) the blow-up of $c_b$ occurs for instance when $\delta_0 \to 0$. Moreover, the constant $c_b$ blows up also when $\beta \to 0$ and $\gamma \to 2\beta - \alpha$.

## 4. The dual pair $(\mu, U)$ and reverse inequalities

**4A.** ***A doubling measure.*** With $\varepsilon$ initially satisfying the condition $0 < \varepsilon < \frac{1}{2}\alpha$, we consider the locally finite measure $\mu$ on $\mathbb{R}^n \times \mathbb{R}^n$ introduced in (1-19). We summarize its basic properties:

**Proposition 4.1.** *With $\mu$ being defined as in* (1-19):

- *Whenever $\mathcal{B} = B \times B$ and $B \subset \mathbb{R}^n$ is a ball with radius $r$, we have*

$$\mu(\mathcal{B}) = \frac{c_\varepsilon(n)r^{n+2\varepsilon}}{\varepsilon}, \tag{4-1}$$

  *where $c_\varepsilon(n)$ denotes a constant depending only on $n$, $\varepsilon$, and it satisfies $1/c(n) \le c_\varepsilon(n) \le c(n)$ for another constant $c(n)$ depending only on $n$.*

- (*doubling diagonal property*) *Whenever* $A \geq 1$, *we have*

$$\sup_{\tilde{x}\in\mathbb{R}^n,\varrho>0} \frac{\mu(\mathcal{B}(\tilde{x}, A\varrho))}{\mu(\mathcal{B}(\tilde{x}, \varrho))} = A^{n+2\varepsilon}. \tag{4-2}$$

- *For every* $A \geq 1$, *there exists a constant* $c_d \equiv c_d(n, A)$ *such that*

$$\frac{\mu(\mathcal{B}(\tilde{x}, \varrho))}{\mu(K_1 \times K_2)} \leq \frac{c_d}{\varepsilon} \tag{4-3}$$

*whenever* $K_1, K_2 \subset B(\tilde{x}, \varrho) \subset \mathbb{R}^n$ *are cubes with sides parallel to the coordinate axes and such that* $|K_1| = |K_2| = \varrho^n/A^n$.

- (*standard doubling property*) *There exists a constant* $c$, *depending only on* $n$, *such that*

$$\sup_{\tilde{x},\tilde{y},\in\mathbb{R}^n,\varrho>0} \frac{\mu(\mathcal{B}(\tilde{x}, \tilde{y}, 2\varrho))}{\mu(\mathcal{B}(\tilde{x}, \tilde{y}, \varrho))} \leq \frac{c}{\varepsilon}. \tag{4-4}$$

*Proof.* The proof of (4-1) follows directly from the definition in (1-19) and a scaling argument, while (4-2) follows from (4-1). The proof of (4-3) is slightly less direct. First, observe that $K_1 \times K_2 \subset \mathcal{B}(\tilde{x}, \varrho)$ and moreover that $|x - y| < 2\varrho$ whenever $x \in K_1$ and $y \in K_2$. Therefore we can estimate

$$\begin{aligned}\mu(\mathcal{B}(\tilde{x}, \varrho)) = \frac{c(n)\varrho^{n+2\varepsilon}}{\varepsilon} &\leq \frac{c(n)A^{2n}}{\varepsilon}\frac{1}{\varrho^{n-2\varepsilon}}\int_{K_1}\int_{K_2} dx\, dy \\ &\leq \frac{c(n)A^{2n}}{\varepsilon}\int_{K_1}\int_{K_2}\frac{dx\, dy}{|x-y|^{n-2\varepsilon}} = \frac{c(n)A^{2n}}{\varepsilon}\mu(K_1 \times K_2),\end{aligned}$$

and the proof of (4-3) is complete. The proof of (4-4) is similar to the one of (4-3); this estimate will not be used in the rest of the paper. □

**4B. *Diagonal reverse Hölder-type inequalities.*** For $(x, y) \in \mathbb{R}^{2n}$, we define the functions

$$U(x, y) := \frac{|u(x) - u(y)|}{|x - y|^{\alpha+\varepsilon}}, \quad G(x, y) := \frac{|g(x) - g(y)|}{|x - y|^{\gamma+2\varepsilon/p}}, \quad F(x, y) := |f(x)|, \tag{4-5}$$

the first two being defined when $x \neq y$. According to Definition 1.2 the function $u$ generates the dual pair $(\mu, U)$. From now on, we shall always assume the following restriction on the number $\varepsilon$:

$$0 < \varepsilon < \min\{\tfrac{1}{2}\alpha, \tfrac{1}{4}|\gamma|(1 + \delta_1), \tfrac{1}{4}(2\beta - \gamma)p\}. \tag{4-6}$$

**Lemma 4.2.** *With the definitions in* (4-5), *it follows that*

$$U \in L^2(\mathbb{R}^{2n}; \mu) \quad \textit{and} \quad F \in L^{2_*+\delta_f}_{\mathrm{loc}}(\mathbb{R}^{2n}; \mu), \quad \textit{with } \delta_f \in [0, \delta_0]. \tag{4-7}$$

*Moreover*, *assuming* (4-6) *it follows that*

$$G \in L^{p+\delta_g}_{\mathrm{loc}}(\mathbb{R}^{2n}; \mu), \quad \textit{where } \delta_g \in [0, p\delta_1]. \tag{4-8}$$

*Proof.* The first inclusion in (4-7) is a direct consequence of the definition in (4-5). As for $F$, for a ball $\mathcal{B} = B \times B$, where $B \subset \mathbb{R}^n$ has radius $r > 0$, we have

$$\int_{\mathcal{B}} F^{2_*+\delta_0}\, d\mu = \int_B \int_B \frac{|f(x)|^{2_*+\delta_0}}{|x-y|^{n-2\varepsilon}}\, dx\, dy \le \frac{cr^{2\varepsilon}}{\varepsilon} \int_B |f|^{2_*+\delta_0}\, dx.$$

This clearly implies that $F \in L^{2_*+\delta_f}_{\mathrm{loc}}(\mathbb{R}^{2n};\mu)$ as long as $\delta_f \le \delta_0$. To prove that $G \in L^{p+\delta_g}_{\mathrm{loc}}(\mathbb{R}^{2n};\mu)$, let us start with the case $2\beta \ge \alpha$, when $\gamma > 0$. By using (4-6) we have

$$\begin{aligned}\int_{\mathcal{B}} G^{p+p\delta_1}\, d\mu &= \int_B \int_B \frac{|g(x)-g(y)|^{p(1+\delta_1)}}{|x-y|^{n+\gamma p(1+\delta_1)+2\varepsilon\delta_1}}\, dx\, dy \\ &\le cr^{\delta_1[\gamma p(1+\delta_1)-2\varepsilon]} \int_B \int_B \frac{|g(x)-g(y)|^{p(1+\delta_1)}}{|x-y|^{n+\gamma p(1+\delta_1)^2}}\, dx\, dy. \end{aligned} \tag{4-9}$$

The last quantity is finite since we are assuming $g \in W^{\gamma(1+\delta_1),p(1+\delta_1)}$, so it follows that $G \in L^{p+\delta_g}_{\mathrm{loc}}(\mathbb{R}^{2n};\mu)$. We finally treat the case $2\beta < \alpha$. In this case, we have $2\varepsilon\delta_1 < 2\varepsilon < |\gamma| p(1+\delta_1) = -\gamma p(1+\delta_1)$, so that $\gamma p(1+\delta_1) + 2\varepsilon\delta_1 < 0$. We can therefore estimate

$$\begin{aligned}\int_{\mathcal{B}} G^{p+p\delta_1}\, d\mu &\le \int_B \int_B \frac{(|g(x)|+|g(y)|)^{p(1+\delta_1)}}{|x-y|^{n+\gamma p(1+\delta_1)+2\varepsilon\delta_1}}\, dx\, dy \\ &\le \frac{cr^{-[\gamma p(1+\delta_1)+2\varepsilon\delta_1]}}{-[\gamma p(1+\delta_1)+2\varepsilon\delta_1]} \int_B |g|^{p(1+\delta_1)}\, dx < \infty, \end{aligned} \tag{4-10}$$

and (4-8) follows again since when $2\beta < \alpha$ we are precisely assuming that $g \in L^{p(1+\delta_1)}_{\mathrm{loc}}(\mathbb{R}^n)$; see (3-2). □

We are now going to state a few inequalities for later use. Let $v \in W^{\tilde{\sigma},q}(B)$ for $\tilde{\sigma} \in (0,1)$ and $q \ge 1$; then the fractional Sobolev inequality

$$\fint_B |v-(v)_B|^2\, dx \le cr^{2\tilde{\sigma}} \left( \int_B \fint_B \frac{|v(x)-v(y)|^q}{|x-y|^{n+\tilde{\sigma}q}}\, dx\, dy \right)^{2/q} \tag{4-11}$$

holds as a consequence of (2-4), provided $q \ge 2n/(n+2\tilde{\sigma})$ and $\tilde{\sigma} > 0$. With $\varepsilon \in (0, \frac{1}{2}\alpha)$ we study the compatibility of the conditions

$$\tilde{\sigma} := \alpha + \varepsilon - \frac{2\varepsilon}{q} \quad \text{and} \quad q \ge \frac{2n}{n+2\tilde{\sigma}} \tag{4-12}$$

in inequality (4-11); this gives $q \ge (2n+4\varepsilon)/(n+2\alpha+2\varepsilon)$. Recalling the definition of the function $U$ in (4-5), and using (4-1), we gain

$$r^{2\tilde{\sigma}} \left( \int_B \fint_B \frac{|u(x)-u(y)|^q}{|x-y|^{n+\tilde{\sigma}q}}\, dx\, dy \right)^{2/q} = \frac{c_\varepsilon(n)^{2/q} r^{2\alpha+2\varepsilon}}{\varepsilon^{2/q}} \left( \fint_{\mathcal{B}} U^q\, d\mu \right)^{2/q},$$

with $c_\varepsilon(n)$ defined in (4-1). We therefore have the following:

**Lemma 4.3.** *Let $\varepsilon \in (0, \frac{1}{2}\alpha)$, and let $q$ be defined by*

$$q := \frac{2n+4\varepsilon}{n+2\alpha+2\varepsilon} < 2. \tag{4-13}$$

*Then the inequality*

$$\fint_B |u-(u)_B|^2\,dx \le \frac{cr^{2(\alpha+\varepsilon)}}{\varepsilon^{2/q}}\left(\fint_{\mathcal{B}} U^q\,d\mu\right)^{2/q} \tag{4-14}$$

*holds for a constant $c$ depending only on $n$ and $\alpha$, whenever $B$ is a ball with radius $r$ and $\mathcal{B}=B\times B$. The same inequality continues to hold when the ball $B$ is replaced by a cube $Q$ with sides of length $r$, and consequently $\mathcal{B}$ is replaced by $Q\times Q$.*

We are now ready for the main result of this section:

**Proposition 4.4** (diagonal reverse Hölder-type inequality). *Let $u\in W^{\alpha,2}(\mathbb{R}^n)$ be a solution to* (1-14) *under the assumptions of Theorem 1.1*; *in particular*, (3-1) *and* (3-3) *are in force. Assume that $\varepsilon$ satisfies* (4-6)*. Then the following reverse Hölder-type inequality with tail holds whenever $\mathcal{B}\subset\mathbb{R}^{2n}$ is a diagonal ball and $\sigma\in(0,1)$:*

$$\begin{aligned}\left(\fint_{\mathcal{B}} U^2 d\mu\right)^{1/2} &\le \frac{c}{\sigma\varepsilon^{1/q-1/2}}\left(\fint_{2\mathcal{B}} U^q d\mu\right)^{1/q} + \frac{\sigma}{\varepsilon^{1/q-1/2}}\sum_{k=1}^{\infty}2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^q\,d\mu\right)^{1/q}\\ &\quad + \frac{c[\mu(\mathcal{B})]^{\eta}}{\varepsilon^{1/2_*-1/2}}\left(\fint_{2\mathcal{B}} F^{2_*}\,d\mu\right)^{1/2_*} + \frac{c_b[\mu(\mathcal{B})]^{\theta}}{\varepsilon^{1/p-1/2}}\sum_{k=1}^{\infty}2^{-k(2\beta-\gamma-2\varepsilon/p)}\left(\fint_{2^k\mathcal{B}} G^p\,d\mu\right)^{1/p},\end{aligned} \tag{4-15}$$

*where $\theta$ and $\eta$ denote the positive exponents*

$$\theta := \frac{\gamma-2\beta+\alpha+\varepsilon(2/p-1)}{n+2\varepsilon} \quad \textit{and} \quad \eta := \frac{\alpha-\varepsilon}{n+2\varepsilon}. \tag{4-16}$$

*The constant $c$ depends only on $n,\alpha,\Lambda$, while the number $q\in(1,2)$ has been defined in* (4-13)*. The constant $c_b$ depends on $n,\alpha,\Lambda,\beta,\gamma,p$ and exhibits the behavior described in* (3-5)*. The infinite sums on the right side of* (4-15) *are finite.*

*Proof.* In the rest of the proof all the constants depend *at least* on $n,\alpha,\Lambda$. We write $\mathcal{B}\equiv B(x_0,r)\times B(x_0,r)$ and apply Theorem 3.2; we choose a cutoff function $\psi\in C_0^\infty(\frac{3}{4}B)$ such that $0\le\psi\le 1$, $|D\psi|\le c(n)/r$ and $\psi\equiv 1$ on $\frac{1}{2}B$. Inequality (3-6) remains valid upon replacing $u$ by $u-(u)_B$; see Remark 3.3. Indeed, notice that for such a function all the integrals on the right-hand side of (3-6) are finite. For this see Section 4C and (4-19) below. All in all, we have

$$\begin{aligned} I_4 &:= \int_B \fint_B \frac{|[u(x)-(u)_B]\psi(x)-[u(y)-(u)_B]\psi(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy\\ &\le cr^{-2\alpha}\fint_B |u(x)-(u)_B|^2\,dx + c\int_{\mathbb{R}^n\setminus B}\frac{|u(y)-(u)_B|}{|x_0-y|^{n+2\alpha}}\,dy\fint_B |u(x)-(u)_B|\,dx\\ &\quad + cr^{2\alpha}\left(\fint_B |f(x)|^{2_*}\,dx\right)^{2/2_*}\\ &\quad + c_b r^{2(\gamma-2\beta+\alpha)}\left[\sum_{k=0}^{\infty}2^{(\gamma-2\beta)k}\left(\int_{2^kB}\fint_{2^kB}\frac{|g(x)-g(y)|^p}{|x-y|^{n+p\gamma}}\,dx\,dy\right)^{1/p}\right]^2\\ &=: J_5+J_6+J_7+J_8. \end{aligned} \tag{4-17}$$

We start by rewriting $I_4$ as

$$I_4 = \frac{1}{|B|}\int_{\mathcal{B}} \frac{\big|[u(x)-(u)_B]\psi(x)-[u(y)-(u)_B]\psi(y)\big|^2}{|x-y|^{2(\alpha+\varepsilon)}}\, d\mu(x,y)$$

so that, with the current choice of $\psi$, we have

$$\frac{r^{2\varepsilon}}{\varepsilon}\fint_{\mathcal{B}/2} U^2 d\mu \le \frac{c(n)}{|B|}\int_{\mathcal{B}/2} U^2\, d\mu \le cI_4.$$

We estimate $J_5$ with the aid of (4-14):

$$J_5 \le \frac{cr^{2\varepsilon}}{\varepsilon^{2/q}}\left(\fint_{\mathcal{B}} U^q d\mu\right)^{2/q}.$$

To estimate $J_6$ we split the term in annuli, and proceed somewhat as in (3-21). As a matter of fact, we will prove that this term is finite; indeed, we have

$$\begin{aligned}\int_{\mathbb{R}^n\setminus B} \frac{|u(y)-(u)_B|}{|x_0-y|^{n+2\alpha}}\, dy &= \sum_{j=0}^{\infty}\int_{2^{j+1}B\setminus 2^jB} \frac{|u(y)-(u)_B|}{|x_0-y|^{n+2\alpha}}\, dy \\ &\le c\sum_{j=0}^{\infty}(2^j r)^{-2\alpha}\fint_{2^{j+1}B}|u(y)-(u)_B|\, dy. \qquad (4\text{-}18)\end{aligned}$$

In turn, we again split every integral in the previous sum, similarly to (3-21), and using Hölder's inequality we estimate

$$\fint_{2^{j+1}B}|u(y)-(u)_B|\, dy \le 2\sum_{k=0}^{j+1}\left(\fint_{2^kB}|u(y)-(u)_{2^kB}|^q\, dy\right)^{1/q}.$$

Each of the previous integrals can be then estimated with the aid of the fractional Poincaré inequality of Lemma 2.1:

$$\begin{aligned}\fint_{2^kB}|u(y)-(u)_{2^kB}|^q\, dy &\le c(2^kr)^{q(\alpha+\varepsilon)-2\varepsilon}\int_{2^kB}\fint_{2^kB}\frac{|u(x)-u(y)|^q}{|x-y|^{n+q\tilde{\sigma}}}\, dx\, dy \\ &= \frac{c(2^kr)^{q(\alpha+\varepsilon)}}{\varepsilon}\fint_{2^k\mathcal{B}} U^q d\mu,\end{aligned}$$

where $\tilde{\sigma}$ is as in (4-12) and $c$ remains independent of $\varepsilon$. As a consequence, we obtain

$$\fint_{2^{j+1}B}|u(y)-(u)_B|\, dy \le \frac{c}{\varepsilon^{1/q}}\sum_{k=0}^{j+1}(2^kr)^{\alpha+\varepsilon}\left(\fint_{2^k\mathcal{B}} U^q d\mu\right)^{1/q}.$$

Connecting the content of the last display to that of (4-18) yields

$$\int_{\mathbb{R}^n\setminus B}\frac{|u(y)-(u)_B|}{|x_0-y|^{n+2\alpha}}\, dy \le \frac{cr^{-\alpha+\varepsilon}}{\varepsilon^{1/q}}\sum_{j=0}^{\infty}\sum_{k=0}^{j+1}2^{-2\alpha j}2^{k(\alpha+\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^q d\mu\right)^{1/q}.$$

Reversing the order of summation gives

$$\begin{aligned}\sum_{j=0}^{\infty}\sum_{k=0}^{j+1} 2^{-2\alpha j} 2^{k(\alpha+\varepsilon)}\left(⨍_{2^k\mathcal{B}} U^q\,d\mu\right)^{1/q} \\ &= \left(⨍_{\mathcal{B}} U^q\,d\mu\right)^{1/q}\sum_{j=0}^{\infty}2^{-2\alpha j}+\sum_{k=1}^{\infty}2^{k(\alpha+\varepsilon)}\left(⨍_{2^k\mathcal{B}} U^q\,d\mu\right)^{1/q}\sum_{j=k-1}^{\infty}2^{-2\alpha j} \\ &\le \frac{c}{\alpha}\sum_{k=0}^{\infty}2^{-k(\alpha-\varepsilon)}\left(⨍_{2^k\mathcal{B}} U^q\,d\mu\right)^{1/q}.\end{aligned}$$

Observe that we have once again used the elementary inequality in (2-2) (with $\beta=\alpha$). All in all, combining the content of the last two displays yields

$$\int_{\mathbb{R}^n\setminus B}\frac{|u(y)-(u)_B|}{|x_0-y|^{n+2\alpha}}\,dy \le \frac{cr^{-\alpha+\varepsilon}}{\varepsilon^{1/q}}\sum_{k=0}^{\infty}2^{-k(\alpha-\varepsilon)}\left(⨍_{2^k\mathcal{B}} U^q\,d\mu\right)^{1/q}, \tag{4-19}$$

so that, via another application of (4-14), we have

$$J_6 \le \frac{cr^{2\varepsilon}}{\varepsilon^{2/q}}\sum_{k=0}^{\infty}2^{-k(\alpha-\varepsilon)}\left(⨍_{2^k\mathcal{B}} U^q\,d\mu\right)^{1/q}\left(⨍_{\mathcal{B}} U^q\,d\mu\right)^{1/q}.$$

With $\sigma\in(0,1)$, using Young's inequality we finally conclude that

$$J_6 \le \frac{cr^{2\varepsilon}}{\sigma^2\varepsilon^{2/q}}\left(⨍_{\mathcal{B}} U^q\,d\mu\right)^{2/q}+\frac{\sigma^2 r^{2\varepsilon}}{\varepsilon^{2/q}}\left[\sum_{k=0}^{\infty}2^{-k(\alpha-\varepsilon)}\left(⨍_{2^k\mathcal{B}} U^q\,d\mu\right)^{1/q}\right]^2.$$

For the estimation of $J_7$ we observe that

$$\begin{aligned}⨍_B |f(x)|^{2_*}\,dx &= ⨍_B ⨍_B |f(x)|^{2_*}\,dx\,dy \\ &\le \frac{c}{r^{n-2\varepsilon+2\varepsilon}}\int_B ⨍_B |f(x)|^{2_*}\,dx\,dy \\ &\le \frac{c}{r^{n+2\varepsilon}}\int_B\int_B \frac{|f(x)|^{2_*}}{|x-y|^{n-2\varepsilon}}\,dx\,dy \\ &\le \frac{c}{\varepsilon}⨍_{\mathcal{B}} F^{2_*}\,d\mu.\end{aligned}$$

Here we have used (4-1) to perform the last estimation and the very definition of the measure $\mu$. By the definition of $J_7$ it then follows that

$$J_7 \le \frac{cr^{2\alpha}}{\varepsilon^{2/2_*}}\left(⨍_{\mathcal{B}} F^{2_*}\,d\mu\right)^{2/2_*}.$$

Next, the definitions of $G(\,\cdot\,)$ and $\mu$ imply

$$J_8 \le \frac{cr^{2(\gamma-2\beta+\alpha+2\varepsilon/p)}}{\varepsilon^{2/p}}\left[\sum_{k=0}^{\infty} 2^{(\gamma-2\beta+2\varepsilon/p)k}\left(\fint_{2^k\mathcal{B}} G^p\,d\mu\right)^{1/p}\right]^2.$$

Finally, connecting the estimates found for $I_4$ and $J_5, \dots, J_8$ to (4-17) yields

$$\begin{aligned}\frac{r^{2\varepsilon}}{\varepsilon}\fint_{\mathcal{B}/2} U^2 d\mu \le \frac{cr^{2\varepsilon}}{\sigma^2\varepsilon^{2/q}}\left(\fint_{\mathcal{B}} U^q d\mu\right)^{2/q} + \frac{\sigma^2 r^{2\varepsilon}}{\varepsilon^{2/q}}\left[\sum_{k=0}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^q d\mu\right)^{1/q}\right]^2 \\ + \frac{cr^{2\alpha}}{\varepsilon^{2/2_*}}\left(\fint_{\mathcal{B}} F^{2_*}\,d\mu\right)^{2/2_*} + \frac{cr^{2(\gamma-2\beta+\alpha+2\varepsilon/p)}}{\varepsilon^{2/p}}\left[\sum_{k=0}^{\infty} 2^{(\gamma-2\beta+2\varepsilon/p)k}\left(\fint_{2^k\mathcal{B}} G^p\,d\mu\right)^{1/p}\right]^2,\end{aligned}$$

from which (4-15) follows immediately (since the ball $\mathcal{B}$ is arbitrary, and we can switch from $\mathcal{B}$ to $2\mathcal{B}$). The right-hand side terms in (4-15) involving infinite sums are finite; we check this in the next remark. □

**Remark 4.5.** A computation based on the definitions in (4-16) gives

$$\frac{2_*\eta}{1-2_*\eta} = \frac{2n(\alpha-\varepsilon)}{n^2+4\varepsilon n+4\alpha\varepsilon} \le \frac{2}{n}$$

and

$$\frac{p\theta}{1-p\theta} = \frac{p(\gamma-2\beta+\alpha)+\varepsilon(2-p)}{n-p(\gamma-2\beta+\alpha)+\varepsilon p} \le \frac{3}{n-p(\gamma-2\beta+\alpha)+\varepsilon p} =: \Lambda_\theta.$$

**4C.** ***The tails are finite.*** We here observe that all the terms on the right-hand sides of (3-6) and (4-15) are finite, obviously confining ourselves to those involving infinite sums. We start with the terms involving $u$. The second term appearing on the right-hand side of (3-6) is seen to be finite by estimating

$$\int_{\mathbb{R}^n\setminus B} \frac{|u(y)|}{|x_0-y|^{n+2\alpha}}\,dy \le \int_{\mathbb{R}^n\setminus B} \frac{|u(y)-(u)_B|}{|x_0-y|^{n+2\alpha}}\,dy + \int_{\mathbb{R}^n\setminus B} \frac{|(u)_B|}{|x_0-y|^{n+2\alpha}}\,dy.$$

The last integral in this display is obviously finite, while the finiteness of the second one can be obtained as in (4-19). In fact, by (2-2) and since $\varepsilon \in (0, \frac{1}{2}\alpha)$, the right-hand side of (4-19) can be further estimated as

$$\begin{aligned}\sum_{k=0}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^q d\mu\right)^{1/q} &\le \sum_{k=0}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}} U^2 d\mu\right)^{1/2} \\ &\le c(\varepsilon,\alpha)\left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|u(x)-u(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy\right)^{1/2}.\end{aligned}$$

This also proves the finiteness of the first infinite sum appearing on the right-hand side of (4-15). We now come to the terms involving $g$, proving that the last series appearing in (4-15) is finite. The finiteness of the last series appearing in (3-6) is therefore implied by looking at the estimate for the term $J_8$ in the proof of Proposition 4.4. We start with the case $2\beta \ge \alpha$, where, using (4-9), we have

$$2^{-k(2\beta-\gamma-2\varepsilon/p)}\left(\fint_{2^k\mathcal{B}} G^p\,d\mu\right)^{1/p} \le c2^{-k(2\beta-\gamma-\delta_1\gamma+n/[p(1+\delta_1)])}[g]_{\gamma(1+\delta_1),p(1+\delta_1)}$$

with $c \equiv c(n, \beta, \gamma, p, \delta_1, r)$, and since by (3-1) we have $\gamma < 2\beta$ and $\delta_1 \gamma p(1+\delta_1) \le n$, the convergence of the series follows. In the case $2\beta < \alpha$ we instead use (4-10) to have the inequality

$$2^{-k(2\beta-\gamma-2\varepsilon/p)} \left( \fint_{2^k \mathcal{B}} G^p \, d\mu \right)^{1/p} \le c 2^{-k(2\beta + n/[p(1+\delta_1)])} \|g\|_{L^{p(1+\delta_1)}(\mathbb{R}^n)},$$

which again implies the convergence of the series in question.

## 5. Level set estimates for dual pairs

In this section we prove a level set estimate which is at the core of the proof of our higher differentiability and integrability results. Let us first define a few functionals. With $\theta$ and $\eta$ as in (4-16), for every $\mathcal{B} \equiv \mathcal{B}(x, \varrho) \subset \mathbb{R}^{2n}$ we define

$$\Psi_{H,M}(\mathcal{B}(x,\varrho)) := \left( \fint_{\mathcal{B}(x,\varrho)} U^2 \, d\mu \right)^{1/2} + \frac{H[\mu(\mathcal{B}(x,\varrho))]^{\eta}}{\varepsilon^{1/2_* - 1/2}} \left( \fint_{\mathcal{B}(x,\varrho)} F^{2_*} \, d\mu \right)^{1/2_*} + \frac{M[\mu(\mathcal{B}(x,\varrho))]^{\theta}}{\varepsilon^{1/p-1/2}} \left( \fint_{\mathcal{B}(x,\varrho)} G^p \, d\mu \right)^{1/p}, \tag{5-1}$$

where $H, M \ge 1$ and $\mathcal{B}(x,\varrho) \subset \mathbb{R}^{2n}$. We also define the functionals

$$\Upsilon_0(\mathcal{B}(x,\varrho)) := \left( \fint_{\mathcal{B}(x,\varrho)} F^{2_*+\delta_f} \, d\mu \right)^{1/(2_*+\delta_f)} + \left( \fint_{\mathcal{B}(x,\varrho)} G^{p+\delta_g} \, d\mu \right)^{1/(p+\delta_g)}, \tag{5-2}$$

$$\Upsilon_1(\mathcal{B}(x,\varrho)) := \sum_{k=0}^{\infty} 2^{-k(\alpha-\varepsilon)} \left( \fint_{\mathcal{B}(x,2^k\varrho)} U^q \, d\mu \right)^{1/q} \tag{5-3}$$

and

$$\Upsilon_{2,M}(\mathcal{B}(x,\varrho)) := \frac{M[\mu(\mathcal{B}(x,\varrho))]^{\theta}}{\varepsilon^{1/p-1/2}} \sum_{k=0}^{\infty} 2^{-k(2\beta-\gamma-2\varepsilon/p)} \left( \fint_{\mathcal{B}(x,2^k\varrho)} G^p \, d\mu \right)^{1/p}. \tag{5-4}$$

We shall denote

$$\Psi(\mathcal{B}(x,\varrho)) := \Psi_{1,1}(\mathcal{B}(x,\varrho)),$$

and shall often use the abbreviations

$$\Psi_{H,M}(\mathcal{B}(x,\varrho)) \equiv \Psi_{H,M}(x,\varrho), \qquad \Upsilon_0(\mathcal{B}(x,\varrho)) \equiv \Upsilon_0(x,\varrho),$$

and so forth. Finally, we can define

$$\mathrm{ADD}(\mathcal{B}(x,\varrho)) \equiv \mathrm{ADD}(x,\varrho) := \Psi(x,\varrho) + \Upsilon_0(x,\varrho) + \Upsilon_1(x,\varrho) + \Upsilon_{2,1}(x,\varrho). \tag{5-5}$$

The aim of this section is to prove the following:

**Proposition 5.1.** *Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ be a solution to* (1-14) *under the assumptions of Theorem 1.1*; *in particular*, (3-1) *and* (3-3) *are in force. Let $\mu$ be the measure defined in* (1-19), *with $\varepsilon$ satisfying* (4-6).

*Consider a ball* $\mathcal{B}(x_0, 2\varrho_0) \subset \mathbb{R}^{2n}$ *such that* $\varrho_0 \le 1$, *and related concentric balls*

$$\mathcal{B}(x_0, \varrho_0) \subset \mathcal{B}(x_0, t) \subset \mathcal{B}(x_0, s) \subset \mathcal{B}(x_0, \tfrac{3}{2}\varrho_0) \tag{5-6}$$

*for* $\varrho_0 < t < s < \frac{3}{2}\varrho_0$. *There exists a constant* $c_s \equiv c_s(n, \alpha, \Lambda)$ *independent of* $\varepsilon$ *and* $p$, *and constants* $c_f \equiv c_f(n, \alpha, \Lambda, \varepsilon) > 1$, $c_g \equiv c_g(n, \alpha, \Lambda, \beta, \gamma, p, \varepsilon) \ge 1$, $\kappa_f \equiv \kappa_f(n, \alpha, \Lambda, \varepsilon) \in (0, 1)$, $\kappa_g \equiv \kappa_g(n, \alpha, \Lambda, \beta, p, \varepsilon) \in (0, 1)$, *such that the inequality*

$$\begin{aligned}\frac{1}{\lambda^2}\int_{\mathcal{B}(x_0,t)\cap\{U>\lambda\}} U^2\,d\mu &\le \frac{c_s}{\varepsilon^{3(2-q)/q}\lambda^q}\int_{\mathcal{B}(x_0,s)\cap\{U>\lambda\}} U^q\,d\mu \\ &\quad + \frac{c_f\lambda_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}}{\lambda^{(1+\eta\delta_f)2_*/(1-2_*\eta)}}\int_{\mathcal{B}(x_0,s)\cap\{F>\kappa_f\lambda\}} F^{2_*}\,d\mu \frac{c_g\lambda_0^{(p+\delta_g)p\theta/(1-p\theta)}}{\lambda^{(1+\theta\delta_g)p/(1-p\theta)}}\int_{\mathcal{B}(x_0,s)\cap\{G>\kappa_g\lambda\}} G^p\,d\mu\end{aligned} \tag{5-7}$$

*holds whenever* $\lambda \ge \lambda_0$, *where*

$$\lambda_0 := \frac{c_a}{\varepsilon}\Big(\frac{\varrho_0}{s-t}\Big)^{2n}\operatorname{ADD}(x_0, 2\varrho_0). \tag{5-8}$$

*The constant* $c_a$ *introduced in the last display depends on* $n, \alpha, \Lambda, \beta, \gamma$, *but is still independent of* $\varepsilon$.

**Remark 5.2.** Unlike $\kappa_f, c_f$, the constants $\kappa_g, c_g$ exhibit the following behavior:

$$\lim_{p\to 2n/[n+2(\gamma-2\beta+\alpha)]}\frac{1}{\kappa_g} = \lim_{\gamma\to 2\beta}\frac{1}{\kappa_g} = \infty = \lim_{p\to 2n/[n+2(\gamma-2\beta+\alpha)]} c_g = \lim_{\gamma\to 2\beta} c_g. \tag{5-9}$$

The proof of Proposition 5.1 is rather delicate and falls into twelve steps. It will take the rest of this section.

**5A.** ***Diagonal balls and Vitali's covering.*** The proof starts with an exit-time argument for the functional $\Psi_{H,M}(\cdot)$, aimed at covering the "diagonal" level set of $U$. The constants $H, M \ge 1$ shall be fixed in due course of the proof, and the whole argument is independent of their particular values until the moment these are fixed. They will be used to give a different weight to the integrals of $F^{2_*}$ and $G^p$: at the exit time, the averages of $F^{2_*}$, $G^p$ will be smaller than the one of $U^2$ provided $H, M$ are chosen to be large enough, respectively. Let us consider concentric diagonal balls as in (5-6). Let $\kappa \in (0, 1]$ be a free parameter to be chosen in the course of the proof, and define

$$\tilde{\lambda}_0 := \kappa^{-1}\sup_{\frac{s-t}{40^n}\le\varrho\le\frac{\varrho_0}{2}}\ \sup_{x\in B(x_0,t)}\big\{\Psi_{H,M}(x,\varrho) + \Upsilon_0(x,\varrho) + \Upsilon_1(x,\varrho) + \Upsilon_{2,M}(x,\varrho)\big\}. \tag{5-10}$$

All the foregoing steps of proofs are independent of the specific choice of $\kappa$ until we fix $\kappa$ in (5-55) below. For the same $\kappa$ (to be defined later) and for $\lambda \ge \tilde{\lambda}_0$, define further the "diagonal level set"

$$D_{\kappa\lambda} := \Big\{(x, x) \in \mathcal{B}(x_0, t) : \sup_{0<\varrho<\frac{s-t}{40^n}} \Psi_{H,M}(x, \varrho) > \kappa\lambda\Big\}. \tag{5-11}$$

Since, by the definition in (5-10), we have

$$\Psi_{H,M}(x, \varrho) \le \kappa\tilde{\lambda}_0 \le \kappa\lambda \qquad \text{if } (x, x) \in \mathcal{B}(x_0, t) \text{ and } \varrho \in [(s-t)/40^n, \varrho_0/2], \tag{5-12}$$

we can find for all $(x,x) \in D_{\kappa\lambda}$ an exit radius $\varrho(x) \in (0, (s-t)/40^n)$ such that

$$\Psi_{H,M}(x, \varrho(x)) \geq \kappa\lambda, \quad \text{while} \quad \sup_{\varrho(x)<\varrho<\frac{s-t}{40^n}} \Psi_{H,M}(x,\varrho) \leq \kappa\lambda. \tag{5-13}$$

Collect the enlarged balls into a covering $\{\mathcal{B}(x, 2\varrho(x)) : (x,x) \in D_{\kappa\lambda}\}$. Balls of the type $\mathcal{B}(x,t,\varrho)$ are, as explained in Section 2, metric balls with respect to the metric (2-1). We therefore apply Vitali's covering theorem to find a countable set $J_D$ and related diagonal points $\{(x_j, x_j)\}_{j\in J_D}$ such that

$$\bigcup_{(x,x)\in D_{\kappa\lambda}} \mathcal{B}(x, 2\varrho(x)) \subset \bigcup_{j\in J_D} \mathcal{B}(x_j, 10\varrho(x_j)) \subset \mathcal{B}(x_0, s) \tag{5-14}$$

and

$$\{\mathcal{B}(x_j, 2\varrho(x_j))\}_{j\in J_D} \quad \text{is a family of mutually disjoint balls.} \tag{5-15}$$

Implicit in (5-14) is the fact that, since $\varrho(x_j) \leq (s-t)/40^n$ and $x_j \in B(x_0,t)$ for every $x_j \in J_D$, then $\mathcal{B}(x_j, 10\varrho(x_j)) \subset \mathcal{B}(x_0, s)$. By (5-12)–(5-13) and the doubling property in (4-2), it follows that

$$\begin{aligned}\sum_{j\in J_D} \int_{\mathcal{B}(x_j,10\varrho(x_j))} U^2\,d\mu &\leq \sum_{j\in J_D} \mu(\mathcal{B}(x_j,10\varrho(x_j)))[\Psi_{H,M}(\mathcal{B}(x_j,10\varrho(x_j)))]^2 \\ &\leq 10^{n+2\varepsilon}\kappa^2\lambda^2 \sum_{j\in J_D} \mu(\mathcal{B}(x_j,\varrho(x_j))).\end{aligned} \tag{5-16}$$

We shall denote in short

$$\mathcal{B}_j := \mathcal{B}(x_j, \varrho(x_j)), \quad \sigma\mathcal{B}_j := \mathcal{B}(x_j, \sigma\varrho(x_j)), \quad \sigma > 0. \tag{5-17}$$

Finally, since we are assuming that $\varrho_0 \leq 1$, by (4-1) we observe that

$$\mu(\mathcal{B}(x_0, 2\varrho_0)) \leq \frac{c2^{n+2\varepsilon}}{\varepsilon} =: L \equiv L(n,\varepsilon). \tag{5-18}$$

**5B.** ***Dyadic cubes, and two constants.*** This section has a very technical nature, and reports a few facts that are true independently of the specify context we are working in. In order to cover the off-diagonal level sets of $U$, we need a more elaborate argument based on classical Calderón–Zygmund coverings. To this aim, we start by recalling basic properties of dyadic cubes in $\mathbb{R}^{2n}$. They differ from the usual ones since they are "centered" at $x_0$ and the size is adapted to the size of the starting ball $\mathcal{B}(x_0, s)$. Define

$$k_0 := \left[-\log_2\left(\frac{s-t}{n10^{10n}}\right)\right]+1, \tag{5-19}$$

where $[\,\cdot\,]$ denotes the integer part of a given number, with the (unnecessarily large) constant $10^{10n}$ having also a symbolic meaning. Let $\Delta_k$, $k \geq k_0$, be the disjoint collection — centered at $x_0$ — of half-open cubes of side length $2^{-k}$ whose closures are touching $\overline{B}(x_0, \frac{1}{2}(s+t))$, i.e.,

$$\Delta_k := \{x_0 + 2^{-k}v + [0, 2^{-k})^n : v \in \mathbb{Z}^n,\ (x_0 + 2^{-k}v + [0,2^{-k}]^n) \cap \overline{B}(x_0, \tfrac{1}{2}(s+t)) \neq \varnothing\}.$$

Notice that, with such a definition, by using (5-19) it follows that $k \geq k_0$ implies

$$B(x_0, t) \subset \bigcup_{K \in \Delta_k} K \subset B(x_0, s). \tag{5-20}$$

The cubes defined above are, up to a translation aimed at centering everything at $x_0$, the standard dyadic cubes in $\mathbb{R}^n$. Let us recall a few basic properties. Let $\Delta$ be the family of all cubes from the families $\Delta_k$, that is, $\Delta := \{K \in \Delta_k : k \geq k_0\}$. Defined this way, every cube $K$ in $\Delta_{k+1}$, $k \geq k_0$, has only one predecessor $\widetilde{K} \in \Delta_k$ such that $K \subset \widetilde{K}$. Moreover, if $K_1 \in \Delta_{k_1}$ and $K_2 \in \Delta_{k_2}$ with $k_0 \leq k_1 < k_2$ and also $K_1 \cap K_2 \neq \varnothing$, then $K_2 \subset K_1$. Starting from the previous cubes, we fix the notation for the corresponding ones in $\mathbb{R}^{2n}$. We set, again for $k \geq k_0$,

$$\Xi_k := \{\mathcal{K} \equiv K_1 \times K_2 : K_1, K_2 \in \Delta_k\}, \qquad \Xi := \bigcup_{k \geq k_0} \Xi_k,$$

while the diagonal cubes build up the family

$$\widetilde{\Xi}_k := \{\mathcal{K} \equiv K \times K : K \in \Delta_k\}. \tag{5-21}$$

With the above definition, it follows from (5-20) that

$$\mathcal{B}(x_0, t) \subset \bigcup_{\mathcal{K} \in \Xi_k} \mathcal{K} \subset \mathcal{B}(x_0, s) \tag{5-22}$$

whenever $k \geq k_0$. Notice that, by defining the product cubes as above, we are actually once again considering dyadic cubes in $\mathbb{R}^{2n}$, with the same properties of the cubes from $\Delta_k$. We also notice that if $\Xi \ni \mathcal{K} = K_1 \times K_2$ then $\widetilde{\mathcal{K}} = \tilde{K}_1 \times \tilde{K}_2$ is its unique predecessor. Finally, let $\mathcal{K} \in \Xi$; then there exist $K_1, K_2 \in \Delta_k$ such that $\mathcal{K} = K_1 \times K_2$; in this case we let

$$k(\mathcal{K}) = k. \tag{5-23}$$

Next, again with $\mathcal{K} = K_1 \times K_2$, we define the *cube projections* as

$$P_1(\mathcal{K}) \equiv P_1\mathcal{K} := K_1 \times K_1 \quad \text{and} \quad P_2(\mathcal{K}) \equiv P_2\mathcal{K} := K_2 \times K_2$$

whenever $K_1, K_2 \in \Delta_k$. In order to shorten the notation, we shall also write $P_h(\mathcal{K}) = P_h\mathcal{K}$ for $h = 1, 2$. It hence follows that

$$P_1(K_1 \times K_2) = P_2(K_2 \times K_1). \tag{5-24}$$

For a given cube $\mathcal{K} \equiv K_1 \times K_2$ we define

$$\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) := \operatorname{dist}(K_1, K_2), \tag{5-25}$$

and its symmetric (or mirror-reflected) cube with respect to the diagonal Diag, is defined by

$$\operatorname{Symm}(\mathcal{K}) = \operatorname{Symm}(K_1 \times K_2) := K_2 \times K_1. \tag{5-26}$$

For future convenience we collect a few basic facts that are a direct consequence of the definitions above, and in particular of (5-23)–(5-26).

**Proposition 5.3.** *Let $\mathcal{K} = K_1 \times K_2 \in \Xi$. The following facts are true*:

- $P_1\mathcal{K}, P_2\mathcal{K} \in \Xi$.
- $\mu(P_1\mathcal{K}) = \mu(P_2\mathcal{K})$ *and* $k(\mathcal{K}) = k(P_1\mathcal{K}) = k(P_2\mathcal{K})$.
- *If* $\Xi \ni \mathcal{H} \subset \mathcal{K}$, *then* $k(\mathcal{K}) \le k(\mathcal{H})$.
- *If* $\widetilde{\mathcal{K}}$ *is the predecessor of* $\mathcal{K}$, *then*

$$\widetilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}}) \le \widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}). \tag{5-27}$$

- *The following relations hold*:

$$\operatorname{dist}(P_1\mathcal{K}, P_2\mathcal{K}) = \sqrt{2}\,\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}), \tag{5-28}$$

$$\operatorname{dist}(\mathcal{K}, \operatorname{Diag}) = \tfrac{1}{2}\operatorname{dist}(P_1\mathcal{K}, P_2\mathcal{K}) = \tfrac{1}{\sqrt{2}}\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}), \tag{5-29}$$

$$\operatorname{dist}(\mathcal{K}, P_1\mathcal{K}) = \operatorname{dist}(\mathcal{K}, P_2\mathcal{K}) = \operatorname{dist}(K_1, K_2) = \widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}), \tag{5-30}$$

$$\widetilde{\operatorname{dist}}(P_1 \operatorname{Symm}(\mathcal{K}), P_2 \operatorname{Symm}(\mathcal{K})) = \widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}).$$

- *Let* $F\colon (x, y) \in \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *be a locally* $\mu$*-integrable function which is symmetric*, *i.e.*, $F(x, y) = F(y, x)$ *holds for every* $x, y \in \mathbb{R}^n$. *Then*

$$\int_{\mathcal{K}} F\,d\mu = \int_{\operatorname{Symm}(\mathcal{K})} F\,d\mu$$

*whenever* $\mathcal{K} \in \Xi$. *In particular*, $\mu(\mathcal{K}) = \mu(\operatorname{Symm}(\mathcal{K}))$ *and*, *moreover*, $k(\mathcal{K}) = k(\operatorname{Symm}(\mathcal{K}))$.

In the next two lemmas we introduce the $\varepsilon$-independent constants $c_{dd}$ and $\tilde{c}_d$, and these will be used very often throughout.

**Lemma 5.4.** *There exists a constant* $c_{dd}$, *depending only on* $n$, *and in particular independent of* $\varepsilon$, *such that for* $h \in \{1, 2\}$ *we have the inequality*

$$c_{dd} \ge \sup_{\mathcal{K}\in\Xi}\left\{\frac{1}{\varepsilon}\left(\frac{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}{2^{-k(\mathcal{K})}}\right)^{n-2\varepsilon}\frac{\mu(\mathcal{K})}{\mu(P_h\mathcal{K})}\right\} + \sup_{\mathcal{K}\in\Xi,\widetilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})\ge 2^{-k(\mathcal{K})}} \varepsilon\left\{\left(\frac{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}{2^{-k(\mathcal{K})}}\right)^{2\varepsilon-n}\frac{\mu(P_h\mathcal{K})}{\mu(\mathcal{K})}\right\} + 1. \tag{5-31}$$

*Proof.* Indeed, observe that using the definition of the measure $\mu$ together with (5-25) (and assuming without loss of generality that $\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) > 0$) we have

$$\mu(P_h\mathcal{K}) = \frac{c(n)}{\varepsilon}2^{-k(\mathcal{K})(n+2\varepsilon)} \quad \text{and} \quad \mu(\mathcal{K}) \le \frac{2^{-2k(\mathcal{K})n}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})^{n-2\varepsilon}}.$$

This allows us to bound the first quantity in (5-31) in a universal way:

$$\frac{1}{\varepsilon}\left(\frac{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}{2^{-k(\mathcal{K})}}\right)^{n-2\varepsilon}\frac{\mu(\mathcal{K})}{\mu(P_h\mathcal{K})} \le c(n).$$

On the other hand, again by (5-25), if $x \in K_1$ and $y \in K_2$ then

$$\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \le |x-y| \le 2\sqrt{n}\,[2^{-k(\mathcal{K})} + \widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})],$$

so that the very definition of the measure $\mu$ yields

$$\mu(\mathcal{K}) \ge \frac{2^{-2k(\mathcal{K})n}}{(2\sqrt{n})^{n-2\varepsilon}[2^{-k(\mathcal{K})} + \widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})]^{n-2\varepsilon}}.$$

Then we have

$$\varepsilon\left(\frac{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}{2^{-k(\mathcal{K})}}\right)^{2\varepsilon-n} \frac{\mu(P_h\mathcal{K})}{\mu(\mathcal{K})} \le c(n)\left(\frac{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}{2^{-k(\mathcal{K})}}\right)^{2\varepsilon-n} \frac{[2^{-k(\mathcal{K})} + \widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})]^{n-2\varepsilon}}{2^{-2k(\mathcal{K})n+k(\mathcal{K})(n+2\varepsilon)}} \le c(n),$$

where we have used that $\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \ge 2^{-k(\mathcal{K})}$. We have therefore proved that (5-31) holds for a constant $c_{dd}$ depending only on $n$. □

The second constant is presented in the next lemma.

**Lemma 5.5.** *There exists a constant $\tilde{c}_d$, depending only on $n$, in particular independent of $\varepsilon$, such that the following inequality holds*:

$$\sup\left\{\frac{\mu(\widetilde{\mathcal{K}})}{\mu(\mathcal{K})} : \widetilde{\mathcal{K}} \text{ is the predecessor of } \mathcal{K},\ \widetilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}}) \ge 2^{-k(\mathcal{K})}\right\} \le \tilde{c}_d. \tag{5-32}$$

*Proof.* Let us consider a dyadic cube $\mathcal{K} = K_1 \times K_2 \subset \mathbb{R}^{2n}$, with $\widetilde{\mathcal{K}}$ being its predecessor, and such that $\widetilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}}) \ge 2^{-k(\mathcal{K})}$. The triangle inequality gives

$$|x-y| \le 2\sqrt{n}\,2^{-k(\mathcal{K})+1} + \operatorname{dist}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}}) \le 8\sqrt{n}\operatorname{dist}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}})$$

whenever $(x, y) \in K_1 \times K_2$. By the very definition of $\mu$ and (5-25), and finally using the inequality in the previous line when performing the final estimation, we get

$$\begin{aligned}\mu(\widetilde{\mathcal{K}}) &\le \widetilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}})^{-(n-2\varepsilon)}|\widetilde{K_1} \times \widetilde{K_2}|\\ &= 4^n\widetilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}})^{-(n-2\varepsilon)}|K_1 \times K_2|\\ &\le c(n)\mu(\mathcal{K}),\end{aligned}$$

and the proof of the lemma is complete. □

**5C.** ***Off-diagonal cubes and Calderón–Zygmund coverings.*** We start by reporting an adaptation of the classical Calderón–Zygmund decomposition lemma. The argument is completely similar to the classical one and for a proof we refer for instance to [Stein 1993], taking into account that the measure $\mu$ is doubling and absolutely continuous with respect to the Lebesgue measure.

**Theorem 5.6.** *Let $Q_0$ be a cube in $\mathbb{R}^{2n}$ and let $\widetilde{U}$ be a nonnegative function in $L^1(Q_0)$. Let $\tilde{\lambda}$ be a real number such that*

$$⨍_{Q_0} \widetilde{U}\, d\mu \le \tilde{\lambda}.$$

*There exists a countable, but possibly finite, family of pairwise disjoint dyadic cubes* $\{Q_i\}$, *with sides parallel to those of* $Q_0$, *such that*

$$\tilde{\lambda} < \fint_{Q_i} \widetilde{U}\, d\mu \ \ \textit{and} \ \ \fint_{\widetilde{Q}_i} \widetilde{U}\, d\mu \leq \tilde{\lambda} \quad \textit{for every } Q_i,$$

*where* $\widetilde{Q}_i$ *denotes the predecessor of* $Q_i$, *and*

$$\widetilde{U} \leq \tilde{\lambda} \quad \textit{a.e. in } Q_0 \setminus \bigcup_i Q_i.$$

We now start to cover the off-diagonal part of the level set of $U$. To this end, let us consider the cubes from the family $\Xi_{k_0}$ and, accordingly, the quantity

$$\lambda_1 := \max\left\{\tilde{\lambda}_0, \sup_{\mathcal{K}\in\Xi_{k_0}} \left(\fint_{\mathcal{K}} U^2\, d\mu\right)^{1/2}\right\}. \tag{5-33}$$

We recall that the numbers $\tilde{\lambda}_0$ and $k_0$ have been determined in (5-10) and (5-19), respectively. Let us observe that (5-22) implies that the family $\{\mathcal{K}\}_{\mathcal{K}\in\Xi_{k_0}}$ forms a disjoint covering of $\mathcal{B}(x_0, t)$. With $\lambda \geq \lambda_1$ we now apply Theorem 5.6 with the choice $Q_0 \equiv \mathcal{K}_0$, for every single cube $\mathcal{K}_0 \in \Xi_{k_0}$; we therefore obtain a family of disjoint dyadic cubes $Q_i(\mathcal{K}_0)$ such that

$$\lambda^2 < \fint_{Q_i(\mathcal{K}_0)} U^2\, d\mu \quad \text{and} \quad \fint_{\widetilde{Q}_i(\mathcal{K}_0)} U^2\, d\mu \leq \lambda^2 \quad \text{for every } Q_i,$$

where, as usual, $\widetilde{Q}_i(\mathcal{K}_0)$ denotes the predecessor of $Q_i(\mathcal{K}_0)$, and

$$U \leq \lambda \quad \text{a.e. in } \mathcal{K}_0 \setminus \bigcup_i Q_i(\mathcal{K}_0).$$

Putting all such families of cubes together, we get a countable family

$$\mathcal{U}_\lambda := \bigcup_{\mathcal{K}_0\in\Xi_{k_0}} \{Q_i(\mathcal{K}_0)\} \equiv \{\mathcal{K}\}$$

of disjoint dyadic cubes $\mathcal{K}$ which are such that

$$\lambda^2 < \fint_{\mathcal{K}} U^2\, d\mu \quad \text{and} \quad \fint_{\widetilde{\mathcal{K}}} U^2\, d\mu \leq \lambda^2 \quad \text{for every } \mathcal{K} \in \mathcal{U}_\lambda, \tag{5-34}$$

where $\widetilde{\mathcal{K}}$ denotes the predecessor of $\mathcal{K}$, and such that

$$U \leq \lambda \quad \text{a.e. in } \mathcal{B}(x_0, t) \setminus \bigcup_{\mathcal{K}\in\mathcal{U}_\lambda} \mathcal{K}. \tag{5-35}$$

**Remark 5.7.** The symmetry of the function $U$ and Proposition 5.3 imply that

$$\int_{\mathcal{K}} U^2\, d\mu = \int_{\mathrm{Symm}(\mathcal{K})} U^2\, d\mu$$

whenever $\mathcal{K} \in \Xi$. It then follows that $\mathcal{K} \in \mathcal{U}_\lambda$ if and only if $\mathrm{Symm}(\mathcal{K}) \in \mathcal{U}_\lambda$.

**5D.** ***First removal of nearly diagonal cubes.*** In this step we are going to show that, in order to cover the level sets of $U^2$, it is sufficient to restrict our attention to those dyadic cubes that are "far" from the diagonal in a suitably quantified sense. Specifically, the word far refers to the fact that for such cubes it happens that their distance to the diagonal is larger than their size. These are really the relevant cubes to analyze, since we shall see that the remaining ones can be covered by the balls considered in (5-14)–(5-15). We therefore start by considering the family of *nearly diagonal* cubes

$$\mathcal{U}^d_\lambda := \{\mathcal{K} \in \mathcal{U}_\lambda : \widetilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}}) < 2^{-k(\mathcal{K})},\ \widetilde{\mathcal{K}} \text{ is the predecessor of } \mathcal{K}\}.$$

With $\mathcal{K} \in \mathcal{U}^d_\lambda$, consider now a point $(\tilde{x}, \tilde{x}) \in \operatorname{Diag}$ such that $\operatorname{dist}((\tilde{x}, \tilde{x}), \widetilde{\mathcal{K}}) = \operatorname{dist}(\operatorname{Diag}, \widetilde{\mathcal{K}})$ and a diagonal ball $\mathcal{B}(\tilde{x}, \varrho) \subset \mathbb{R}^{2n}$ with radius $\varrho$ greater than or equal to

$$\frac{5\sqrt{n}\,\widetilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}})}{2} + 5\sqrt{n}2^{-k(\mathcal{K})+1}.$$

Keeping (5-29) in mind and applying it to $\widetilde{\mathcal{K}}$, it follows that $\widetilde{\mathcal{K}} \subset \mathcal{B}(\tilde{x}, \varrho)$. Ultimately, we can find a diagonal ball $\mathcal{B} \equiv \mathcal{B}(\tilde{x}, 24\sqrt{n}2^{-k(\mathcal{K})})$ such that $\mathcal{K} \subset \mathcal{B}$. Notice that, in this case, by using (4-3) from Proposition 4.1 and recalling that $(\tilde{x}, \tilde{x}) \in \operatorname{Diag}$, we conclude there exists a constant $c_d$, which depends only on $n$, such that

$$1 \le \frac{\mu(\mathcal{B})}{\mu(\mathcal{K})} \le \frac{c_d}{\varepsilon} \equiv \frac{c_d(n)}{\varepsilon}. \tag{5-36}$$

Therefore, if $\mathcal{K} \in \mathcal{U}^d_\lambda$, then the lower bound in (5-34) yields

$$\lambda^2 < \fint_{\mathcal{K}} U^2\, d\mu \le \frac{\mu(\mathcal{B})}{\mu(\mathcal{K})} \fint_{\mathcal{B}} U^2\, d\mu \le \frac{c_d}{\varepsilon} \fint_{\mathcal{B}} U^2\, d\mu.$$

Assuming that the number $\kappa \in (0, 1]$ introduced in (5-10) satisfies

$$\kappa \in (0, \kappa_0], \quad \kappa_0 := \frac{\varepsilon^{1/2}}{\sqrt{2c_d}}, \tag{5-37}$$

all in all we have proved that

$$\text{for all } \mathcal{K} \in \mathcal{U}^d_\lambda,\ \text{there exists } \mathcal{B}^{\mathcal{K}} \equiv B^{\mathcal{K}} \times B^{\mathcal{K}} \quad \text{such that} \quad \fint_{\mathcal{B}^{\mathcal{K}}} U^2\, d\mu > \kappa^2\lambda^2 \text{ and } \mathcal{K} \subset \mathcal{B}^{\mathcal{K}}.$$

This means that, if $\tilde{x}$ is the center of $\mathcal{B}^{\mathcal{K}}$, by the exit-time condition (5-13) it follows that $(\tilde{x}, \tilde{x}) \in D_{\kappa\lambda}$ and then $\mathcal{B}^{\mathcal{K}} \subset B(\tilde{x}, \varrho(\tilde{x}))$. By (5-14) it hence follows that

$$\bigcup_{\mathcal{K} \in \mathcal{U}^d_\lambda} \mathcal{K} \subset \bigcup_{j \in J_D} 10\mathcal{B}_j. \tag{5-38}$$

Notice that here, in order to find the ball $\mathcal{B}^{\mathcal{K}}$ and apply the exit-time condition in (5-13), we have used that the radius of the diagonal ball $\mathcal{B} \equiv \mathcal{B}(\tilde{x}, 24\sqrt{n}2^{-k(\mathcal{K})})$ is smaller than $(s-t)/40^n$. In turn, this is a consequence of the fact that $k(\mathcal{K}) \ge k_0$ and of the fact that $k_0$ is large enough, as prescribed in (5-19).

**5E.** ***Off-diagonal reverse Hölder inequalities.*** As we saw in the previous section, $\mathcal{U}_\lambda^d$ has already been covered by the diagonal cover. Thus, we shall now only consider so-called off-diagonal cubes:

$$\mathcal{U}_\lambda^{nd} := \{\mathcal{K} \in \mathcal{U}_\lambda : \widetilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}}) \geq 2^{-k(\mathcal{K})},\ \widetilde{\mathcal{K}} \text{ is the predecessor of } \mathcal{K}\}. \tag{5-39}$$

We notice that (5-27) implies

$$\mathcal{K} \in \mathcal{U}_\lambda^{nd} \quad \Longrightarrow \quad \widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \geq 2^{-k(\mathcal{K})}.$$

The goal is thus to sort and estimate suitable off-diagonal sums of the measures of cubes belonging to $\mathcal{U}_\lambda^{nd}$. The following lemma is our basic tool. It roughly says that, for nondiagonal cubes, reverse Hölder inequalities hold automatically, and independently of the fact that the function solves an equation. The price to pay is the appearance of certain diagonal correction terms, and this is eventually treated by some combinatorial lemmas.

**Lemma 5.8** (off-diagonal reverse inequality). *Let $k \geq k_0$, and suppose that $\mathcal{K} \in \Xi_k$. There exists a constant $c_{nd} \equiv c_{nd}(n, \alpha)$, independent of $\varepsilon$, such that whenever $\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \geq 2^{-k}$, the inequality*

$$\left(\fint_{\mathcal{K}} U^2\,d\mu\right)^{1/2} \leq c_{nd}\left(\fint_{\mathcal{K}} U^q\,d\mu\right)^{1/q} + \frac{c_{nd}}{\varepsilon^{1/q}}\left(\frac{2^{-k}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{\alpha+\varepsilon}\left(\fint_{P_1\mathcal{K}} U^q\,d\mu\right)^{1/q} + \frac{c_{nd}}{\varepsilon^{1/q}}\left(\frac{2^{-k}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{\alpha+\varepsilon}\left(\fint_{P_2\mathcal{K}} U^q\,d\mu\right)^{1/q}$$

*holds, with the number $q$ being defined in* (4-13). *In particular, this inequality holds whenever $\mathcal{K} \in \mathcal{U}_\lambda^{nd}$.*

*Proof.* Let $\mathcal{K} \equiv K_1 \times K_2 \in \Xi_k$, and find points $x_1 \in \overline{K}_1$ and $y_1 \in \overline{K}_2$ such that $\operatorname{dist}(K_1, K_2) = |x_1 - y_1|$. By the triangle inequality we obtain, whenever $x, y \in \mathcal{K}$,

$$\begin{aligned} |x-y| &\leq \operatorname{dist}(K_1, K_2) + |x_1 - x| + |y_1 - y| \\ &\leq \operatorname{dist}(K_1, K_2) + 2\sqrt{n}\,2^{-k} \\ &\leq 3\sqrt{n}\,\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) = 3\sqrt{n}\operatorname{dist}(K_1, K_2). \end{aligned}$$

Therefore we have

$$1 \leq \frac{|x-y|}{\operatorname{dist}(K_1, K_2)} \leq 3\sqrt{n} \quad \text{for all } (x, y) \in \mathcal{K}, \tag{5-40}$$

where the first inequality is a trivial consequence of the definition of $\operatorname{dist}(K_1, K_2)$. Next, thanks to (5-40), the very definition of $\mu$ yields

$$\mu(\mathcal{K}) \approx \frac{4^{-nk}}{\operatorname{dist}(K_1, K_2)^{n-2\varepsilon}}, \tag{5-41}$$

with the constant involved being independent of $\varepsilon$, but just depending on $n$. By using (5-40) and (5-41) we then have

$$\begin{aligned}\left(\fint_{\mathcal{K}} U^2\,d\mu\right)^{1/2} &= \left(\frac{1}{\mu(\mathcal{K})}\int_{K_1}\int_{K_2}\frac{|u(x)-u(y)|^2}{|x-y|^{n+2\alpha}}\,dx\,dy\right)^{1/2} \\ &\le c\left(\frac{\operatorname{dist}(K_1,K_2)^{n-2\varepsilon-(n+2\alpha)}}{4^{-nk(\mathcal{K})}}\int_{K_1}\int_{K_2}|u(x)-u(y)|^2\,dx\,dy\right)^{1/2} \\ &\le c\operatorname{dist}(K_1,K_2)^{-(\alpha+\varepsilon)}\left(\fint_{K_1}\fint_{K_2}|u(x)-u(y)|^2\,dx\,dy\right)^{1/2}, \qquad (5\text{-}42)\end{aligned}$$

where $c$ depends only on $n$. We further estimate the integral on the right using Minkowski's inequality:

$$\begin{aligned}\left(\fint_{K_1}\fint_{K_2}|u(x)-u(y)|^2\,dx\,dy\right)^{1/2} &\le \left(\fint_{K_1}|u(x)-(u)_{K_1}|^2\,dx\right)^{1/2} \\ &\quad + \left(\fint_{K_2}|u(x)-(u)_{K_2}|^2\,dx\right)^{1/2} + |(u)_{K_1}-(u)_{K_2}|.\end{aligned}$$

By using the fractional Poincaré inequality of Lemma 4.3 applied to cubes, and recalling that $P_h\mathcal{K} = K_h \times K_h$ for $h \in \{1, 2\}$, we deduce that

$$\left(\fint_{K_h}|u(x)-(u)_{K_h}|^2\,dx\right)^{1/2} \le \frac{c2^{-k(\alpha+\varepsilon)}}{\varepsilon^{1/q}}\left(\fint_{P_h\mathcal{K}} U^q\,d\mu\right)^{1/q}, \quad h \in \{1,2\},$$

with the implied constant $c$ depending only on $n$ and $\alpha$. Finally, by Hölder's inequality, and using (5-40) and (5-41) repeatedly, we get

$$\begin{aligned}|(u)_{K_1}-(u)_{K_2}| &\le \fint_{K_1}\fint_{K_2}|u(x)-u(y)|\,dx\,dy \\ &\le \left(\fint_{K_1}\fint_{K_2}|u(x)-u(y)|^q\,dx\,dy\right)^{1/q} \\ &\le c\left(\frac{1}{\operatorname{dist}(K_1,K_2)^{n-2\varepsilon}\mu(\mathcal{K})}\int_{K_1}\int_{K_2}|u(x)-u(y)|^q\,dx\,dy\right)^{1/q} \\ &\le c\left(\fint_{\mathcal{K}}|u(x)-u(y)|^q\,d\mu\right)^{1/q} \\ &\le c\operatorname{dist}(K_1,K_2)^{\alpha+\varepsilon}\left(\fint_{\mathcal{K}} U^q\,d\mu\right)^{1/q},\end{aligned}$$

with $c \equiv c(n)$. Combining the content of the last four displays and recalling the definition in (5-25) finishes the proof. □

We remark that the previous lemma works for any function $u \in W^{\alpha,2}$ and does not require that $u$ solves any equation; moreover, the lemma works for every positive integer $k$. Applying it in the present situation gives the next result:

**Corollary 5.9.** *Let $k \geq k_0$ be an integer, and suppose that $\mathcal{K} \in \Xi_k$ is such that $\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \geq 2^{-k}$. Assume that*

$$\left(\fint_{\mathcal{K}} U^2 \, d\mu\right)^{1/2} \geq \lambda$$

*and that the number $\kappa$ introduced in* (5-10) *satisfies*

$$\kappa \in (0, \kappa_1], \quad \kappa_1 := \frac{\varepsilon^{1/q}}{2^{1/q} 3 c_{nd}}, \tag{5-43}$$

*where $c_{nd} \equiv c_{nd}(n, \alpha)$ has been defined in Lemma 5.8. Then we have*

$$\begin{aligned}\mu(\mathcal{K}) \leq \frac{3^q c_{nd}^q}{\lambda^q} \int_{\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu + \frac{3^q c_{nd}^q}{\varepsilon\lambda^q} \frac{\mu(\mathcal{K})}{\mu(P_1\mathcal{K})} \left(\frac{2^{-k}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)} \int_{P_1\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu \\ + \frac{3^q c_{nd}^q}{\varepsilon\lambda^q} \frac{\mu(\mathcal{K})}{\mu(P_2\mathcal{K})} \left(\frac{2^{-k}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)} \int_{P_2\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu.\end{aligned} \tag{5-44}$$

*In particular, the inequality* (5-44) *holds whenever $\mathcal{K} \in \mathcal{U}_\lambda^{nd}$.*

*Proof.* Appealing to Lemma 5.8, and using the elementary inequality $(a+b+c)^q \leq 3^{q-1}(a^q+b^q+c^q)$ valid for all nonnegative numbers $a, b, c \in \mathbb{R}$, we get

$$\frac{\lambda^q}{3^{q-1}c_{nd}^q} \leq \fint_{\mathcal{K}} U^q \, d\mu + \frac{1}{\varepsilon}\left(\frac{2^{-k}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)} \left(\fint_{P_1\mathcal{K}} U^q \, d\mu + \fint_{P_2\mathcal{K}} U^q \, d\mu\right).$$

To estimate the integrals appearing on the right-hand side, we note that by (5-43) we have

$$\fint_E U^q \, d\mu \leq \kappa_1^q \lambda^q + \frac{1}{\mu(E)} \int_{E\cap\{U>\kappa\lambda\}} U^q \, d\mu$$

with $E \in \{\mathcal{K}, P_1\mathcal{K}, P_2\mathcal{K}\}$ so that, recalling that $\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \geq 2^{-k}$, we gain

$$\begin{aligned}\frac{\lambda^q}{3^{q-1}c_{nd}^q} \leq \frac{3\kappa_1^q\lambda^q}{\varepsilon} + \frac{1}{\mu(\mathcal{K})} \int_{\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu + \frac{1}{\varepsilon\mu(P_1\mathcal{K})} \left(\frac{2^{-k}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)} \int_{P_1\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu \\ + \frac{1}{\varepsilon\mu(P_2\mathcal{K})} \left(\frac{2^{-k}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)} \int_{P_2\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu.\end{aligned}$$

Now (5-44) follows by inserting (5-43) in the last estimate and reabsorbing terms. □

**5F.** ***Families of off-diagonal cubes.*** With $\mathcal{U}_\lambda^{nd}$ as defined in (5-39), consider now the families

$$\mathcal{M}_\lambda^h := \left\{\mathcal{K} \in \mathcal{U}_\lambda^{nd} : \fint_{P_h\mathcal{K}} U^q \, d\mu \leq (10n)^{n+2}\kappa^q\lambda^q\right\} \tag{5-45}$$

and

$$\mathcal{N}_\lambda^h := \left\{\mathcal{K} \in \mathcal{U}_\lambda^{nd} : \fint_{P_h\mathcal{K}} U^q \, d\mu > (10n)^{n+2}\kappa^q\lambda^q\right\} \tag{5-46}$$

for $h \in \{1,2\}$, with the number $\kappa$ defined as in (5-10) and $q$ defined as in (4-13). Furthermore, define

$$\mathcal{M}_\lambda := \mathcal{M}_\lambda^1 \cap \mathcal{M}_\lambda^2 \quad \text{and} \quad \mathcal{N}_\lambda := \mathcal{N}_\lambda^1 \cup \mathcal{N}_\lambda^2, \tag{5-47}$$

so that we have the decomposition into disjoint families

$$\mathcal{U}_\lambda^{nd} = \mathcal{M}_\lambda \cup \mathcal{N}_\lambda. \tag{5-48}$$

**Lemma 5.10** (soft off-diagonal summation). *The inequality*

$$\sum_{\mathcal{K}\in\mathcal{M}_\lambda} \mu(\mathcal{K}) \le \frac{6^q c_{nd}^q}{\lambda^q} \int_{\mathcal{B}(x_0,s)\cap\{U>\kappa\lambda\}} U^q \, d\mu \tag{5-49}$$

*holds whenever the number* $\kappa$ *in* (5-10) *satisfies*

$$\kappa \in (0,\kappa_2], \quad \kappa_2 := \frac{\varepsilon^{1/q}}{8^{1/q} 3 c_{nd} (10n)^{(n+2)/q}}. \tag{5-50}$$

*The constant* $c_{nd} \equiv c_{nd}(n,\alpha)$ *was defined in Lemma 5.8 and appears in Corollary 5.9*; *it is independent of* $\varepsilon$.

*Proof.* It is sufficient to prove that if $\mathcal{K} \in \mathcal{M}_\lambda$, then

$$\mu(\mathcal{K}) \le \frac{6^q c_{nd}^q}{\lambda^q} \int_{\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu. \tag{5-51}$$

After this, (5-49) follows, since the initial family $\mathcal{U}_\lambda$ is disjoint and (5-22) holds. For the proof of (5-51), notice that if $\mathcal{K} \in \mathcal{M}_\lambda^h$, then we have, for $h \in \{1,2\}$, that

$$\begin{aligned}\frac{3^q c_{nd}^q}{\varepsilon\lambda^q} \frac{\mu(\mathcal{K})}{\mu(P_h\mathcal{K})} \left( \frac{2^{-k(\mathcal{K})}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})} \right)^{q(\alpha+\varepsilon)} \int_{P_h\mathcal{K}\cap\{U>\kappa\lambda\}} U^q \, d\mu \\ \le \mu(\mathcal{K}) \frac{3^q c_{nd}^q}{\varepsilon\lambda^q} \fint_{P_h\mathcal{K}} U^q \, d\mu \le \mu(\mathcal{K}) \frac{3^q c_{nd}^q}{\varepsilon\lambda^q} (10n)^{n+2} \kappa^q \lambda^q \le \frac{\mu(\mathcal{K})}{8}.\end{aligned} \tag{5-52}$$

Using this last estimate for $h \in \{1,2\}$ in combination with (5-44), and reabsorbing terms, gives (5-51); the proof is therefore complete. □

It remains to study the family $\mathcal{N}_\lambda$ defined in (5-47). To this aim, we introduce the family of diagonal cubes defined by

$$P_h\mathcal{N}_\lambda := \{P_h\mathcal{K} : \mathcal{K} \in \mathcal{N}_\lambda^h\}, \quad h \in \{1,2\}.$$

Keeping (5-24) and Remark 5.7 in mind, we have that

$$\mathcal{K} \in \mathcal{N}_\lambda^1 \iff \operatorname{Symm}(\mathcal{K}) \in \mathcal{N}_\lambda^2 \tag{5-53}$$

whenever $\mathcal{K} \in \Xi$. Now, let us make a remark: consider $T \in P_1\mathcal{N}_\lambda$, so that $T = P_1(\mathcal{K})$ for some $\mathcal{K} \in \mathcal{N}_\lambda^1$. Therefore $T = P_2(\operatorname{Symm}(\mathcal{K}))$ by (5-24), and by (5-53) we have $\operatorname{Symm}(\mathcal{K}) \in \mathcal{N}_\lambda^2$. We conclude that

$T \in P_2\mathcal{N}_\lambda$ and eventually that $P_1\mathcal{N}_\lambda \subset P_2\mathcal{N}_\lambda$. In a similar way it follows that $P_2\mathcal{N}_\lambda \subset P_1\mathcal{N}_\lambda$. We therefore conclude that $P_1\mathcal{N}_\lambda = P_2\mathcal{N}_\lambda = P_1\mathcal{N}_\lambda \cup P_2\mathcal{N}_\lambda$. Let $P\mathcal{N}_\lambda$ be a disjoint subfamily of $P_1\mathcal{N}_\lambda \cup P_2\mathcal{N}_\lambda$ such that

$$\bigcup_{\mathcal{H}\in P\mathcal{N}_\lambda} \mathcal{H} = \bigcup_{\mathcal{K}\in P_1\mathcal{N}_\lambda \cup P_2\mathcal{N}_\lambda} \mathcal{K}. \tag{5-54}$$

Note that, since all the cubes of the family $P\mathcal{N}_\lambda$ are themselves dyadic cubes, such an extracted disjoint covering always exists. We remark that a straightforward consequence of the definitions is that all cubes from $P\mathcal{N}_\lambda$ obviously belong to $P_1\mathcal{N}_\lambda \cup P_2\mathcal{N}_\lambda$ and are therefore diagonal cubes.

**5G.** ***Determining*** $\boldsymbol{\kappa}$**.** We here determine the parameter $\kappa$ in (5-10). By choosing

$$\kappa := \min\{\kappa_0, \kappa_1, \kappa_2\} \equiv \min\left\{\frac{\varepsilon^{1/2}}{\sqrt{2c_d}}, \frac{\varepsilon^{1/q}}{2^{1/q}3c_{nd}}, \frac{\varepsilon^{1/q}}{8^{1/q}3c_{nd}(10n)^{(n+2)/q}}\right\}, \tag{5-55}$$

conditions (5-37), (5-43) and (5-50) are all satisfied, so the content and the results of Sections 5D–5F are at our disposal. Recalling that $c_d$ in (5-36) (coming from Proposition 4.1) depends only on $n$, and that $c_{nd}$ from Lemma 5.8 depends only on $n, \alpha$, we conclude there exists a new constant $c_\kappa$ such that

$$\kappa \geq \varepsilon^{1/q}/c_\kappa, \quad c_\kappa \equiv c_\kappa(n, \alpha). \tag{5-56}$$

**5H.** ***Further removal of nearly diagonal cubes.*** We recall that our final goal is to estimate the measure of the level sets of $U$. Since the nearly diagonal part has already been covered, we proceed in excluding from the subsequent analysis those cubes covered by the balls in (5-14)–(5-15). Therefore, we introduce

$$\mathcal{N}_{\lambda,d} := \left\{\mathcal{K} \in \mathcal{N}_\lambda : \mathcal{K} \subset \bigcup_{j\in J_D} 10\mathcal{B}_j\right\} \tag{5-57}$$

and, accordingly,

$$\mathcal{N}_{\lambda,nd} := \mathcal{N}_\lambda \setminus \mathcal{N}_{\lambda,d} \quad \text{and} \quad \mathcal{N}^h_{\lambda,nd} := \mathcal{N}_{\lambda,nd} \cap \mathcal{N}^h_\lambda, \quad \text{for } h \in \{1, 2\}. \tag{5-58}$$

We observe that the main difficulty in handling the cubes from the family $P\mathcal{N}_\lambda$ stems from the fact that they do not belong to the family $\mathcal{U}_\lambda$, i.e., they do not come from an exit-time argument and therefore no control is available on the values taken by $U^2$ on such cubes. This will be bypassed via a very delicate combinatorial argument. The next lemma is instrumental to that.

**Lemma 5.11.** *Let $\mathcal{K} \in \mathcal{N}_{\lambda,nd}$ be such that $P_h\mathcal{K} \subset \mathcal{H}$ for some $\mathcal{H} \in P\mathcal{N}_\lambda$ and some $h \in \{1, 2\}$. Then $\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \geq 2^{-k(\mathcal{H})}$.*

*Proof.* First, let us consider a cube $\mathcal{H} \in P\mathcal{N}_\lambda$; take the diagonal ball $\mathcal{B}(\mathcal{H}) \equiv \mathcal{B}(x_\mathcal{H}, 2^{-(k(\mathcal{H})+1)})$, $(x_\mathcal{H}, x_\mathcal{H})$ being the center of $\mathcal{H}$. It follows that

$$\mathcal{B}(\mathcal{H}) \subset \mathcal{H} \subset \sqrt{n}\mathcal{B}(\mathcal{H}). \tag{5-59}$$

Therefore we have by Hölder's inequality and the definition of $P\mathcal{N}_\lambda$ that

$$
\begin{aligned}
(10n)^{(n+2)/q}\kappa\lambda &< \left(\fint_{\mathcal{H}} U^q \, d\mu\right)^{1/q} \\
&\le \left(\frac{\mu(10n\mathcal{B}(\mathcal{H}))}{\mu(\mathcal{B}(\mathcal{H}))} \fint_{10n\mathcal{B}(\mathcal{H})} U^q \, d\mu\right)^{1/q} \\
&\le (10n)^{(n+2)/q}\left(\fint_{10n\mathcal{B}(\mathcal{H})} U^2 \, d\mu\right)^{1/2}. \qquad (5\text{-}60)
\end{aligned}
$$

By the definition of $D_{\kappa\lambda}$ in (5-11) it follows that $(x_{\mathcal{H}}, x_{\mathcal{H}}) \in D_{\kappa\lambda}$, and then the exit-time condition (5-13) gives $\mathcal{B}(\mathcal{H}) \subset \mathcal{B}(x_{\mathcal{H}}, \varrho(x_{\mathcal{H}}))$. We are using that the radius of the ball $10n\mathcal{B}(\mathcal{H})$ is smaller than $(s-t)/40^n$. In turn, this is a consequence of the fact that $k(\mathcal{H})+1 \ge k_0$ and of (5-19). Then (5-14) implies

$$
10n\,\mathcal{B}(\mathcal{H}) \subset \bigcup_{j\in J_D} 10\mathcal{B}_j. \qquad (5\text{-}61)
$$

Now, in order to prove the lemma, assume by contradiction that $\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) < 2^{-k(\mathcal{H})}$ and let $\mathcal{B}(\mathcal{H})$ be the ball determined in (5-59), and for which (5-61) holds. We are going to show that

$$
\mathcal{K} \subset 10n\,\mathcal{B}(\mathcal{H}), \qquad (5\text{-}62)
$$

and this then contradicts the assumption $\mathcal{K} \in \mathcal{N}_{\lambda,nd}$ by (5-61). In order to show (5-62), we observe that Proposition 5.3 and the fact that $P_h\mathcal{K} \subset \mathcal{H}$ give

$$
\operatorname{dist}(\mathcal{K}, \mathcal{H}) \le \operatorname{dist}(\mathcal{K}, P_h\mathcal{K}) = \widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \le 2^{-k(\mathcal{H})}.
$$

Again by Proposition 5.3 we have $k(P_h\mathcal{K}) = k(\mathcal{K})$ and $k(\mathcal{K}) \ge k(\mathcal{H})$. Therefore, since $\mathcal{H} \subset \sqrt{n}\mathcal{B}(\mathcal{H})$ and the radius of $\mathcal{B}(\mathcal{H})$ is $2^{-(k(\mathcal{H})+1)}$, then (5-62) must hold. The proof of the lemma is complete. □

**5I.** ***Summation in $\mathcal{N}_{\lambda,nd}$.*** The aim of this section is to prove the following:

**Lemma 5.12** (hard off-diagonal summation). *There exists a constant $c$, depending only on $n$, $\alpha$, such that the estimate*

$$
\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}} \mu(\mathcal{K}) \le \frac{c}{\lambda^q} \int_{\mathcal{B}(x_0,s)\cap\{U>\kappa\lambda\}} U^q \, d\mu \qquad (5\text{-}63)
$$

*holds, where $\kappa$ has been determined in* (5-55).

*Proof. Step 1: Classifying cubes.* Here we classify the cubes from $\mathcal{N}_{\lambda,nd}$ according to their projections, thereby partitioning $\mathcal{N}_{\lambda,nd}$ into suitable disjoint subfamilies. For every $\mathcal{H} \in P\mathcal{N}_\lambda$, set

$$
\mathcal{N}^h_{\lambda,nd}(\mathcal{H}) := \{\mathcal{K} \in \mathcal{N}_{\lambda,nd} : P_h\mathcal{K} \subset \mathcal{H}\}, \quad h \in \{1, 2\}.
$$

Since $P\mathcal{N}_\lambda$ is a disjoint covering of $P_1\mathcal{N}_\lambda \cup P_2\mathcal{N}_\lambda = P_1\mathcal{N}_\lambda = P_2\mathcal{N}_\lambda$, we have the decomposition in mutually disjoint families

$$
\mathcal{N}^h_{\lambda,nd} = \bigcup_{\mathcal{H}\in P\mathcal{N}_\lambda} \mathcal{N}^h_{\lambda,nd}(\mathcal{H}). \qquad (5\text{-}64)
$$

This means that for $\mathcal{H}_1, \mathcal{H}_2 \in P\mathcal{N}_\lambda$ it follows that $\mathcal{N}^h_{\lambda,nd}(\mathcal{H}_1) \cap \mathcal{N}^h_{\lambda,nd}(\mathcal{H}_2) \neq \varnothing$ implies $\mathcal{H}_1 = \mathcal{H}_2$. Indeed, assume that a cube $\mathcal{K}$ lies in $\mathcal{N}^h_{\lambda,nd}(\mathcal{H}_1) \cap \mathcal{N}^h_{\lambda,nd}(\mathcal{H}_2)$ and that $\mathcal{H}_1 \neq \mathcal{H}_2$; then we would have that $P_h\mathcal{K} \subset \mathcal{H}_1 \cap \mathcal{H}_2$ against the fact that $\mathcal{H}_1$ and $\mathcal{H}_2$ have a nonempty intersection, being elements of the disjoint covering $P\mathcal{N}_\lambda$. Next, let us recall that for every $\mathcal{K} \in \mathcal{N}^h_{\lambda,nd}(\mathcal{H})$ we have $k(\mathcal{K}) = k(P_h\mathcal{K}) \geq k(\mathcal{H})$, and this leads us to define the classes

$$[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_i := \{\mathcal{K} \in \mathcal{N}^h_{\lambda,nd}(\mathcal{H}) : k(\mathcal{K}) = i + k(\mathcal{H})\}$$

for $h \in \{1, 2\}$ and for every integer $i \geq 0$. Therefore, the decomposition in mutually disjoint families

$$\mathcal{N}^h_{\lambda,nd}(\mathcal{H}) = \bigcup_{i\geq 0} [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_i$$

holds, in the sense that $[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_i \cap [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_j \neq \varnothing$ implies that $i = j$. Next, take $\mathcal{H} \in P\mathcal{N}_\lambda$; by Lemma 5.11 we have that if $\mathcal{K} \in \mathcal{N}^h_{\lambda,nd}(\mathcal{H})$, that is, if $P_h\mathcal{K} \subset \mathcal{H}$, then it follows that $\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \geq 2^{-k(\mathcal{H})}$, and this finally leads us to classify elements of $[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_i$ in the following way:

$$[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j} := \{\mathcal{K} \in [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_i : 2^{j-k(\mathcal{H})} \leq \widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) < 2^{j+1-k(\mathcal{H})}\}$$

for $h \in \{1, 2\}$ and for integers $i, j \geq 0$. Again, we have the decomposition

$$\mathcal{N}^h_{\lambda,nd}(\mathcal{H}) = \bigcup_{i,j\geq 0} [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}, \tag{5-65}$$

and these are disjoint classes in the sense that, if $[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i_1,j_1} \cap [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i_2,j_2} \neq \varnothing$, then $(i_1, j_1) = (i_2, j_2)$. All in all, in view of (5-64) and (5-65), we have the decomposition into mutually disjoint classes

$$\mathcal{N}^h_{\lambda,nd} = \bigcup_{\mathcal{H}\in P\mathcal{N}_\lambda} \bigcup_{i,j\geq 0} [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}. \tag{5-66}$$

*Step 2: Sums and further partitions.* Let us fix $\mathcal{H} \in P\mathcal{N}_\lambda$; our aim here is to prove that the following inequality holds for $h \in \{1, 2\}$:

$$\frac{1}{\varepsilon} \sum_{\mathcal{K}\in\mathcal{N}^h_{\lambda,nd}(\mathcal{H})} \frac{\mu(\mathcal{K})}{\mu(P_h\mathcal{K})} \left(\frac{2^{-k(\mathcal{K})}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)} \int_{P_h\mathcal{K}\cap\{U>\kappa\lambda\}} U^q\, d\mu \leq \frac{c(n)}{\alpha^2} \int_{\mathcal{H}\cap\{U>\kappa\lambda\}} U^q\, d\mu. \tag{5-67}$$

We start by recalling that, by the very definitions in (5-46) and (5-47), and again (5-27), we have that $\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K}) \geq 2^{-k(\mathcal{K})}$ as soon as $\mathcal{K} \in \mathcal{N}_{\lambda,nd}$; (5-31) yields

$$\frac{1}{\varepsilon} \frac{\mu(\mathcal{K})}{\mu(P_h\mathcal{K})} \leq c_{dd} \left(\frac{2^{-k(\mathcal{K})}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})}\right)^{n-2\varepsilon}$$

for $h \in \{1, 2\}$, and, moreover, if $\mathcal{K} \in [\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}$, we also have that

$$\frac{2^{-k(\mathcal{K})}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})} = \frac{1}{2^i} \frac{2^{-k(\mathcal{H})}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K}, P_2\mathcal{K})} \leq \frac{1}{2^{i+j}}.$$

Using the inequalities in the last two displays we can estimate

$$\begin{aligned}
&\frac{1}{\varepsilon}\sum_{\mathcal{K}\in\mathcal{N}^h_{\lambda,nd}(\mathcal{H})}\frac{\mu(\mathcal{K})}{\mu(P_h\mathcal{K})}\left(\frac{2^{-k(\mathcal{K})}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\right)^{q(\alpha+\varepsilon)}\int_{P_h\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\
&\quad\le c_{dd}\sum_{\mathcal{K}\in\mathcal{N}^h_{\lambda,nd}(\mathcal{H})}\left(\frac{2^{-k(\mathcal{K})}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\right)^{n+q(\alpha+\varepsilon)-2\varepsilon}\int_{P_h\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\
&\quad= c_{dd}\sum_{i,j=0}^{\infty}\sum_{\mathcal{K}\in[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}}\left(\frac{2^{-k(\mathcal{K})}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\right)^{n+q(\alpha+\varepsilon)-2\varepsilon}\int_{P_h\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\
&\quad\le c(n)\sum_{i,j=0}^{\infty}\left(\frac{1}{2^{i+j}}\right)^{n+q(\alpha+\varepsilon)-2\varepsilon}\sum_{\mathcal{K}\in[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}}\int_{P_h\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu.
\end{aligned}\tag{5-68}$$

In order to evaluate the last sum we have to further decompose $[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}$. For each integer $i\ge 0$, $\mathcal{H}$ contains precisely $4^{ni}=2^{2ni}$ disjoint cubes from $\Xi_{i+k(\mathcal{H})}$ and exactly $2^{ni}$ disjoint cubes from $\widetilde{\Xi}_{i+k(\mathcal{H})}$; see the definition in (5-21) and in the preceding display. As a consequence, it contains at most $2^{ni}$ disjoint (diagonal) cubes from the class $\widetilde{\Xi}_{i+k(\mathcal{H})}\cap(P_1\mathcal{N}_\lambda\cup P_2\mathcal{N}_\lambda)$. We anyway consider all the diagonal cubes $\widetilde{\Xi}_{i+k(\mathcal{H})}$ from $\mathcal{H}$, and relabel them as

$$\{\widetilde{\mathcal{H}}\in\widetilde{\Xi}_{i+k(\mathcal{H})}:\widetilde{\mathcal{H}}\subset\mathcal{H}\}=\{\mathcal{H}_i^m:1\le m\le 2^{ni}\},\tag{5-69}$$

so that

$$\sum_{m=1}^{2^{ni}}\int_{\mathcal{H}_i^m\cap\{U>\kappa\lambda\}}U^q\,d\mu\le\int_{\mathcal{H}\cap\{U>\kappa\lambda\}}U^q\,d\mu.\tag{5-70}$$

Now, let us concentrate one moment on the elements of $[\mathcal{N}^1_{\lambda,nd}(\mathcal{H})]_{i,j}$; a similar argument then applies to $[\mathcal{N}^2_{\lambda,nd}(\mathcal{H})]_{i,j}$. For any $\mathcal{K}\in[\mathcal{N}^1_{\lambda,nd}(\mathcal{H})]_{i,j}$, there is the unique cube from the diagonal class (5-21), which we denote by $\mathcal{H}_i^m(\mathcal{K})$, such that $P_1\mathcal{K}=\mathcal{H}_i^m(\mathcal{K})$. Now, note that for $h\in\{1,2\}$ one can split $[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}$ into subsets

$$[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j,m}:=\{\mathcal{K}\in[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}:P_h\mathcal{K}=\mathcal{H}_i^m\},\quad m\in\{1,\dots,2^{ni}\}.$$

Since $\mathcal{N}^1_{\lambda,nd}$ is a family of dyadic cubes, if $\mathcal{K}_1,\mathcal{K}_2\in[\mathcal{N}^1_{\lambda,nd}(\mathcal{H})]_{i,j,m}$ and $\mathcal{K}_1\ne\mathcal{K}_2$, then $P_2\mathcal{K}_1\cap P_2\mathcal{K}_2=\varnothing$, i.e., the second components are disjoint (otherwise the two cubes would coincide). A similar argument holds when looking at $\mathcal{N}^2_{\lambda,nd}$. It then follows that

$$\#[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j,m}\le c(n)2^{n(i+j)},\quad h\in\{1,2\},\tag{5-71}$$

for every choice of $i,\ j\geq 0$ and $m\in\{1,\ldots,2^{ni}\}$. We use now use (5-70)–(5-71) to estimate

$$\begin{aligned}\sum_{\mathcal{K}\in[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}}\int_{P_h\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu&=\sum_{m=1}^{2^{ni}}\sum_{\mathcal{K}\in[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j,m}}\int_{\mathcal{H}^m_i\cap\{U>\kappa\lambda\}}U^q\,d\mu\\&\leq c(n)2^{n(i+j)}\sum_{m=1}^{2^{ni}}\int_{\mathcal{H}^m_i\cap\{U>\kappa\lambda\}}U^q\,d\mu\\&\leq c(n)2^{n(i+j)}\int_{\mathcal{H}\cap\{U>\kappa\lambda\}}U^q\,d\mu.\end{aligned}$$

Using also (2-2) it follows that

$$\begin{aligned}&\sum_{i,j=0}^{\infty}\Big(\frac{1}{2^{i+j}}\Big)^{n+q(\alpha+\varepsilon)-2\varepsilon}\sum_{\mathcal{K}\in[\mathcal{N}^h_{\lambda,nd}(\mathcal{H})]_{i,j}}\int_{P_h\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\&\quad\leq c(n)\sum_{i,j=0}^{\infty}\Big(\frac{1}{2^{i+j}}\Big)^{q(\alpha+\varepsilon)-2\varepsilon}\int_{\mathcal{H}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\&\quad\leq\frac{c(n)}{[q(\alpha+\varepsilon)-2\varepsilon]^2}\int_{\mathcal{H}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\&\quad\leq\frac{c(n)}{\alpha^2}\int_{\mathcal{H}\cap\{U>\kappa\lambda\}}U^q\,d\mu.\end{aligned}$$

Notice that we have used that, since $q>1$ and $\varepsilon<\frac{1}{2}\alpha$, we have $q(\alpha+\varepsilon)-2\varepsilon>\frac{1}{2}\alpha$. Combining the inequality in the last display with (5-68) yields (5-67).

*Step 3: Summation.* Let now $\mathcal{K}\in\mathcal{N}^1_{\lambda,nd}$. There are then two cases: either $\mathcal{K}\in\mathcal{M}^2_\lambda$ or $\mathcal{K}\in\mathcal{N}^2_\lambda$ (the relevant definitions are in (5-45), (5-46) and (5-58)). Now, if $\mathcal{K}\in\mathcal{M}^2_\lambda$, then using (5-44) and (5-52), and reabsorbing terms, we obtain that

$$\mu(\mathcal{K})\leq\frac{6^q c^q_{nd}}{\lambda^q}\int_{\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu+\frac{6^q c^q_{nd}}{\varepsilon\lambda^q}\frac{\mu(\mathcal{K})}{\mu(P_1\mathcal{K})}\bigg(\frac{2^{-k}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\bigg)^{q(\alpha+\varepsilon)}\int_{P_1\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu.$$

If, on the other hand, $\mathcal{K}\in\mathcal{N}^2_\lambda$, then using (5-44) we get

$$\begin{aligned}\mu(\mathcal{K})\leq\frac{3^q c^q_{nd}}{\lambda^q}\int_{\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu+\frac{3^q c^q_{nd}}{\varepsilon\lambda^q}\frac{\mu(\mathcal{K})}{\mu(P_1\mathcal{K})}\bigg(\frac{2^{-k}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\bigg)^{q(\alpha+\varepsilon)}\int_{P_1\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\+\frac{3^q c^q_{nd}}{\varepsilon\lambda^q}\frac{\mu(\mathcal{K})}{\mu(P_2\mathcal{K})}\bigg(\frac{2^{-k}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\bigg)^{q(\alpha+\varepsilon)}\int_{P_2\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu.\end{aligned}$$

A similar reasoning holds if $\mathcal{K}\in\mathcal{N}^2_{\lambda,nd}$. Summing up over the cubes $\mathcal{K}\in\mathcal{N}_{\lambda,nd}=\mathcal{N}^1_{\lambda,nd}\cup\mathcal{N}^2_{\lambda,nd}$ then yields

$$\begin{aligned}\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\mu(\mathcal{K})\le\frac{6^q c_{nd}^q}{\lambda^q}&\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\int_{\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\&+\frac{6^q c_{nd}^q}{\varepsilon\lambda^q}\sum_{\mathcal{K}\in\mathcal{N}^1_{\lambda,nd}}\frac{\mu(\mathcal{K})}{\mu(P_1\mathcal{K})}\Big(\frac{2^{-k(\mathcal{K})}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\Big)^{q(\alpha+\varepsilon)}\int_{P_1\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\&+\frac{6^q c_{nd}^q}{\varepsilon\lambda^q}\sum_{\mathcal{K}\in\mathcal{N}^2_{\lambda,nd}}\frac{\mu(\mathcal{K})}{\mu(P_2\mathcal{K})}\Big(\frac{2^{-k(\mathcal{K})}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\Big)^{q(\alpha+\varepsilon)}\int_{P_2\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu.\end{aligned}\tag{5-72}$$

Observe that a key point in the previous inequality, due to the argument at the beginning of Step 3, is that terms involving integrals over $P_h\mathcal{K}$ appear on the right-hand side if and only if $\mathcal{K}\in\mathcal{N}^h_{\lambda,nd}$, for $h\in\{1,2\}$. By the symmetry of $U$ and $\mu$, by (5-53) and subsequent remarks, and using Proposition 5.3, we have that if $\mathcal{K}\in\mathcal{N}^2_{\lambda,nd}$, then $\operatorname{Symm}(\mathcal{K})\in\mathcal{N}^1_{\lambda,nd}$, and vice versa; moreover, again by Proposition 5.3, we have

$$\int_{P_2\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu=\int_{P_1\operatorname{Symm}(\mathcal{K})\cap\{U>\kappa\lambda\}}U^q\,d\mu.$$

Hence the last two terms in (5-72) coincide. Therefore, also recalling (5-64), (5-72) can be rewritten as

$$\begin{aligned}\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\mu(\mathcal{K})\le\frac{c}{\lambda^q}&\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\int_{\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\\&+\frac{c}{\varepsilon\lambda^q}\sum_{\mathcal{H}\in P\mathcal{N}_\lambda}\sum_{\mathcal{K}\in\mathcal{N}^1_{\lambda,nd}(\mathcal{H})}\frac{\mu(\mathcal{K})}{\mu(P_1\mathcal{K})}\Big(\frac{2^{-k(\mathcal{K})}}{\widetilde{\operatorname{dist}}(P_1\mathcal{K},P_2\mathcal{K})}\Big)^{q(\alpha+\varepsilon)}\int_{P_1\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu\end{aligned}$$

for a constant $c$ depending on $n,\alpha$. To estimate the last term we make use of (5-67), and this yields

$$\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\mu(\mathcal{K})\le\frac{c}{\lambda^q}\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\int_{\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu+\frac{c}{\lambda^q}\sum_{\mathcal{H}\in P\mathcal{N}_\lambda}\int_{\mathcal{H}\cap\{U>\kappa\lambda\}}U^q\,d\mu.$$

At this stage (5-63) follows, observing that

$$\sum_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}}\int_{\mathcal{K}\cap\{U>\kappa\lambda\}}U^q\,d\mu+\sum_{\mathcal{H}\in P\mathcal{N}_\lambda}\int_{\mathcal{H}\cap\{U>\kappa\lambda\}}U^q\,d\mu\le 2\int_{\mathcal{B}(x_0,s)\cap\{U>\kappa\lambda\}}U^q\,d\mu.$$

This is in turn true since the families $P\mathcal{N}_\lambda$ and $\mathcal{N}_{\lambda,nd}$ are made of mutually disjoint cubes and all their members are contained in $\mathcal{B}(x_0,s)$ (since these families are contained in $\Xi$ and (5-22) holds). The proof of Lemma 5.12 is complete. □

**5J.** ***Conclusion of the off-diagonal analysis.*** The next lemma summarizes the decomposition results in the off-diagonal case:

**Lemma 5.13** (off-diagonal level set inequality). *The inequality*

$$\int_{\mathcal{B}(x_0,t)\cap\{U>\lambda\}} U^2\,d\mu \le 10^{n+2}\kappa^2\lambda^2 \sum_{j\in J_D} \mu(\mathcal{B}_j) + c\lambda^{2-q} \int_{\mathcal{B}(x_0,s)\cap\{U>\kappa\lambda\}} U^q\,d\mu \tag{5-73}$$

*holds for a constant $c$ depending only on $n, \alpha$, while the number $\kappa$ has been defined in* (5-55) *and exhibits the dependence displayed in* (5-56).

*Proof.* We have the decompositions in disjoint classes $\mathcal{U}_\lambda = \mathcal{U}_\lambda^d \cup \mathcal{U}_\lambda^{nd}$ and $\mathcal{U}_\lambda^{nd} = \mathcal{M}_\lambda \cup \mathcal{N}_{\lambda,d} \cup \mathcal{N}_{\lambda,nd}$, and we recall that all the cubes from $\mathcal{U}_\lambda^{nd}$ are mutually disjoint. Moreover, by (5-38) and (5-57) it follows that

$$\bigcup_{\mathcal{K}\in\mathcal{U}_\lambda^d} \mathcal{K} \cup \bigcup_{\mathcal{K}\in\mathcal{N}_{\lambda,d}} \mathcal{K} \subset \bigcup_{j\in J_D} 10\mathcal{B}_j.$$

Therefore

$$\bigcup_{\mathcal{K}\in\mathcal{U}_\lambda} \mathcal{K} \subset \bigcup_{j\in J_D} 10\mathcal{B}_j \cup \bigcup_{\mathcal{K}\in\mathcal{M}_\lambda} \mathcal{K} \cup \bigcup_{\mathcal{K}\in\mathcal{N}_{\lambda,nd}} \mathcal{K}.$$

Keeping this in mind, and recalling (5-35), we start by estimating

$$\int_{\mathcal{B}(x_0,t)\cap\{U>\lambda\}} U^2\,d\mu \le \sum_j \int_{10\mathcal{B}_j\cap\{U>\lambda\}} U^2\,d\mu + \sum_{\mathcal{K}\in\mathcal{M}_\lambda\cup\mathcal{N}_{\lambda,nd}} \int_{\mathcal{K}\cap\{U>\lambda\}} U^2\,d\mu.$$

By (5-34) it follows that, if $\mathcal{K}\in\mathcal{M}_\lambda\cup\mathcal{N}_{\lambda,nd}\subset\mathcal{U}_\lambda^{nd}$, then

$$\fint_{\mathcal{K}} U^2\,d\mu \le \frac{\mu(\widetilde{\mathcal{K}})}{\mu(\mathcal{K})} \fint_{\widetilde{\mathcal{K}}} U^2\,d\mu \le \tilde{c}_d\lambda^2.$$

Note that we have used (5-32), since $\mathcal{K}\in\mathcal{U}_\lambda^{nd}$ implies by the definition in (5-39) that $\widetilde{\operatorname{dist}}(P_1\widetilde{\mathcal{K}}, P_2\widetilde{\mathcal{K}}) \ge 2^{-k(\mathcal{K})}$. Therefore we conclude that

$$\mathcal{K}\in\mathcal{M}_\lambda\cup\mathcal{N}_{\lambda,nd} \quad\Longrightarrow\quad \int_{\mathcal{K}\cap\{U>\lambda\}} U^2\,d\mu \le \tilde{c}_d\lambda^2\mu(\mathcal{K}).$$

Using this last inequality together with (5-16) yields

$$\int_{\mathcal{B}(x_0,t)\cap\{U>\lambda\}} U^2\,d\mu \le 10^{n+2\varepsilon}\kappa^2\lambda^2 \sum_{j\in J_D} \mu(\mathcal{B}_j) + \tilde{c}_d\lambda^2 \sum_{\mathcal{K}\in\mathcal{M}_\lambda\cup\mathcal{N}_{\lambda,nd}} \mu(\mathcal{K}),$$

and (5-73) follows by just using Lemmas 5.10 and 5.12. □

**Remark 5.14.** An interesting point of Lemma 5.13 is that it does not make use of the fact that $u$ is a solution. All the estimates just rely on the fact that $u$ belongs to the Sobolev space $W^{\alpha,2}$. This is ultimately linked to the fact that the analysis in Sections 5B–5J is made in a zone where the kernel of the operator, that is, $|x-y|^{-(n+2\alpha)}$, is not very singular. The ultimate outcome is that the whole issue reduces to estimating $\sum_j \mu(\mathcal{B}_j)$. Therefore, it remains to perform the analysis close to the diagonal, and this will be done in the next section.

**5K. *Diagonal estimates.*** Whenever $\mathcal{B}_j$ is a ball from the covering determined in (5-14)–(5-15), from (5-13) it follows that $\Psi_{H,M}(\mathcal{B}_j) \geq \kappa\lambda$. By the very definition of $\Psi_{H,M}(\cdot)$ in (5-1) it then follows that at least one of the following three inequalities must hold:

$$\left(\fint_{\mathcal{B}_j} U^2 d\mu\right)^{1/2} \geq \frac{\kappa\lambda}{3}, \tag{5-74}$$

$$\frac{H[\mu(\mathcal{B}_j)]^\eta}{\varepsilon^{1/2_*-1/2}}\left(\fint_{\mathcal{B}_j} F^{2_*}\, d\mu\right)^{1/2_*} \geq \frac{\kappa\lambda}{3}, \tag{5-75}$$

$$\frac{M[\mu(\mathcal{B}_j)]^\theta}{\varepsilon^{1/p-1/2}}\left(\fint_{\mathcal{B}_j} G^p\, d\mu\right)^{1/p} \geq \frac{\kappa\lambda}{3}, \tag{5-76}$$

where $\kappa$ has been defined in (5-55). We now examine the occurrence of each of the three cases separately.

*Occurrence of* (5-74) *(and estimate of the tail at the exit time).* When (5-74) holds, using (4-15) we have

$$\begin{aligned}\kappa\lambda \leq \frac{c}{\sigma\varepsilon^{1/q-1/2}}&\left(\fint_{2\mathcal{B}_j} U^q d\mu\right)^{1/q}\\ &+\frac{\sigma}{\varepsilon^{1/q-1/2}}\sum_{k=1}^{\infty}2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}_j} U^q\, d\mu\right)^{1/q}+\frac{c_1[\mu(\mathcal{B}_j)]^\eta}{\varepsilon^{1/2_*-1/2}}\left(\fint_{2\mathcal{B}_j} F^{2_*}\, d\mu\right)^{1/2_*}\\ &+\frac{c_2[\mu(\mathcal{B}_j)]^\theta}{\varepsilon^{1/p-1/2}}\sum_{k=1}^{\infty}2^{-k(2\beta-\gamma-2\varepsilon/p)}\left(\fint_{2^k\mathcal{B}_j} G^p\, d\mu\right)^{1/p}\end{aligned} \tag{5-77}$$

for all $\sigma \in (0, 1]$. The constants $c_1, c$ depend only on $n, \alpha, \Lambda$, while $c_2 := 3c_b$ and therefore it depends on $n, \alpha, \Lambda, \beta, \gamma, p$ and exhibits the behavior described in (3-5). With $\mathcal{B}_j \equiv \mathcal{B}(x_j, \varrho(x_j))$ we determine the integer $m \geq 0$ such that

$$2^{-m}\varrho_0/2 \leq \varrho(x_j) < 2^{-m+1}\varrho_0/2. \tag{5-78}$$

Notice that since $\varrho(x_j) < (s-t)/40^n$, we have $m \geq 3$ and moreover $(s-t)/40^n \leq \varrho_0/40^n \leq 2^{m-1}\varrho(x_j)$, so that (5-10) implies

$$\Upsilon_0(2^{m-1}\mathcal{B}_j) + \Upsilon_1(2^{m-1}\mathcal{B}_j) + \Upsilon_{2,M}(2^{m-1}\mathcal{B}_j) \leq \kappa\tilde{\lambda}_0. \tag{5-79}$$

On the other hand, the terms indexed before $m$ can be estimated using Hölder's inequality and the exit-time condition in (5-13) as

$$\left(\fint_{2^k\mathcal{B}_j} U^q\, d\mu\right)^{1/q} \leq \Psi_{H,M}(2^k\mathcal{B}_j) \leq \kappa\lambda \quad \text{if } 1 \leq k \leq m-1.$$

By using the inequalities in the last two displays we then have

$$\begin{aligned}\sum_{k=1}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}_j} U^q\,d\mu\right)^{1/q} &= \sum_{k=1}^{m-2} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^k\mathcal{B}_j} U^q\,d\mu\right)^{1/q}\\ &\quad + 2^{-(m-1)(\alpha-\varepsilon)}\sum_{k=0}^{\infty} 2^{-k(\alpha-\varepsilon)}\left(\fint_{2^{k+m-1}\mathcal{B}_j} U^q\,d\mu\right)^{1/q}\\ &\le \kappa\lambda\sum_{k=1}^{m-2} 2^{-k(\alpha-\varepsilon)} + 2^{-(m-1)(\alpha-\varepsilon)}\Upsilon_1(2^{m-1}\mathcal{B}_j)\\ &\le \kappa\lambda\sum_{k=1}^{m-2} 2^{-k(\alpha-\varepsilon)} + 2^{-(m-1)(\alpha-\varepsilon)}\kappa\tilde{\lambda}_0\\ &\le \kappa\lambda\sum_{k=1}^{\infty} 2^{-k(\alpha-\varepsilon)} \le \frac{4\kappa\lambda}{\alpha-\varepsilon} \le \frac{8\kappa\lambda}{\alpha},\end{aligned}$$

where we have used (2-2) and that $\varepsilon < \frac{1}{2}\alpha$. In a completely similar way, again using (5-79), we have

$$\frac{c_2[\mu(\mathcal{B}_j)]^{\theta}}{\varepsilon^{1/p-1/2}}\sum_{k=1}^{\infty} 2^{-k(2\beta-\gamma-2\varepsilon/p)}\left(\fint_{2^k\mathcal{B}_j} G^p\,d\mu\right)^{1/p} \le \frac{4c_2\kappa\lambda}{(2\beta-\gamma-2\varepsilon/p)M} \le \frac{8c_2\kappa\lambda}{(2\beta-\gamma)M},$$

where we also used the upper bound on $\varepsilon$ in (4-6). By (5-78) and the fact that $m \ge 3$ we gain that $2\varrho(x_j) \le \frac{1}{2}\varrho_0$ so that (5-13) and Hölder's inequality yield

$$\frac{c_1[\mu(\mathcal{B}_j)]^{\eta}}{\varepsilon^{1/2_*-1/2}}\left(\fint_{2\mathcal{B}_j} F^{2_*}\,d\mu\right)^{1/2_*} \le \frac{c_1\Psi_{H,M}(2\mathcal{B}_j)}{H} \le \frac{c_1\kappa\lambda}{H}.$$

By merging the inequalities in the last three displays with (5-77) we obtain

$$\kappa\lambda \le \frac{c}{\sigma\varepsilon^{1/q-1/2}}\left(\fint_{2\mathcal{B}_j} U^q\,d\mu\right)^{1/q} + \frac{\sigma}{\varepsilon^{1/q-1/2}}\frac{8\kappa\lambda}{\alpha} + \frac{c_1\kappa\lambda}{H} + \frac{8c_2\kappa\lambda}{(2\beta-\gamma)M}. \tag{5-80}$$

We recall that up to now the parameters $H, M \ge 1$ in the definition in (5-1) have not yet been chosen, and neither has $\sigma \in (0,1)$. Hence, taking

$$\sigma := \frac{\varepsilon^{1/q-1/2}\alpha}{56}, \quad H := 6c_1, \quad M := \frac{56c_2}{2\beta-\gamma}, \tag{5-81}$$

and reabsorbing terms in (5-80), we conclude that

$$\kappa\lambda \le \frac{c}{\varepsilon^{2/q-1}}\left(\fint_{2\mathcal{B}_j} U^q\,d\mu\right)^{1/q} \implies \mu(\mathcal{B}_j) \le \frac{c}{\varepsilon^{2-q}(\kappa\lambda)^q}\int_{2\mathcal{B}_j} U^q\,d\mu,$$

where $c$ depends on $n, \alpha, \Lambda$. Now, select a number $\kappa_3 > 0$; also using (4-2), we estimate

$$\begin{aligned}\frac{c}{\varepsilon^{2-q}(\kappa\lambda)^q}\int_{2\mathcal{B}_j} U^q\,d\mu &\le \frac{c}{\varepsilon^{2-q}(\kappa\lambda)^q}\int_{2\mathcal{B}_j\cap\{U\le\kappa_3\kappa\lambda\}} U^q\,d\mu + \frac{c}{\varepsilon^{2-q}(\kappa\lambda)^q}\int_{2\mathcal{B}_j\cap\{U>\kappa_3\kappa\lambda\}} U^q\,d\mu \\ &\le \frac{\tilde{c}\mu(\mathcal{B}_j)\kappa_3^q}{\varepsilon^{2-q}} + \frac{\tilde{c}}{\varepsilon^{2-q}(\kappa\lambda)^q}\int_{2\mathcal{B}_j\cap\{U>\kappa_3\kappa\lambda\}} U^q\,d\mu, \end{aligned} \tag{5-82}$$

again for $\tilde{c}$ depending only on $n, \alpha, \Lambda$. By choosing

$$\kappa_3 \le \left(\frac{\varepsilon^{2-q}}{2\tilde{c}}\right)^{1/q}, \tag{5-83}$$

we arrive at

$$\mu(\mathcal{B}_j) \le \frac{c_3}{(\kappa\lambda)^q}\int_{2\mathcal{B}_j\cap\{U>\kappa_3\kappa\lambda\}} U^q\,d\mu, \quad \text{where } c_3 := \frac{2\tilde{c}}{\varepsilon^{2-q}}, \tag{5-84}$$

and $\tilde{c}$ is independent of $\varepsilon$ and only depends on $n, \alpha, \Lambda$.

*Occurrence of* (5-75)–(5-76). In case of (5-75), we have

$$\left(\frac{\kappa\lambda}{3}\right)^{2_*} \le \frac{H^{2_*}[\mu(\mathcal{B}_j)]^{2_*\eta-1}}{\varepsilon^{1-2_*/2}}\int_{\mathcal{B}_j} F^{2_*}\,d\mu,$$

which readily implies

$$\mu(\mathcal{B}_j) \le \left(\frac{3H}{\varepsilon^{1/2_*-1/2}\kappa\lambda}\right)^{2_*/(1-2_*\eta)}\left(\int_{\mathcal{B}_j} F^{2_*}\,d\mu\right)^{1/(1-2_*\eta)}.$$

Observe that by the definitions given in (4-16) we have that $2_*\eta < \frac{1}{2}$. With $\kappa_4 \in (0,1)$ being a positive number to be chosen in a few lines, we further split the support of the right-hand side integral as already done in (5-82):

$$\begin{aligned}\left(\int_{\mathcal{B}_j} F^{2_*}\,d\mu\right)^{1/(1-2_*\eta)} &\le \left[\int_{\mathcal{B}_j\cap\{F>\kappa_4\kappa\lambda\}} F^{2_*}\,d\mu + (\kappa_4\kappa\lambda)^{2_*}\mu(\mathcal{B}_j)\right]^{1/(1-2_*\eta)} \\ &\le 2^{2_*\eta/(1-2_*\eta)}\left(\int_{\mathcal{B}_j\cap\{F>\kappa_4\kappa\lambda\}} F^{2_*}\,d\mu\right)^{1/(1-2_*\eta)} \\ &\quad + [2(L+1)]^{2_*/(1-2_*\eta)}(\kappa_4\kappa\lambda)^{2_*/(1-2_*\eta)}\mu(\mathcal{B}_j).\end{aligned}$$

Observe that, in view of $\mathcal{B}_j \subset \mathcal{B}(x_0, 2\varrho_0)$ and (5-18), we have estimated

$$[\mu(\mathcal{B}_j)]^{1/(1-2_*\eta)} \le [\mu(\mathcal{B}(x_0, 2\varrho_0))]^{2_*\eta/(1-2_*\eta)}\mu(\mathcal{B}_j) \le L^{2_*\eta/(1-2_*\eta)}\mu(\mathcal{B}_j). \tag{5-85}$$

We now take $\kappa_4 \in (0,1)$ in order to satisfy

$$\left[\frac{6H(L+1)\kappa_4}{\varepsilon^{1/2_*-1/2}}\right]^{2_*/(1-2_*\eta)} \le \frac{1}{2} \quad \Longrightarrow \quad \kappa_4 \le \left(\frac{1}{2}\right)^{(1-2_*\eta)/2_*}\frac{\varepsilon^{1/2_*-1/2}}{6H(L+1)}. \tag{5-86}$$

Using this choice and combining the content of the last four displays (and recalling that $2_*\eta/(1-2_*\eta)\le 1$) then yields that

$$\mu(\mathcal{B}_j)\le 4\left(\frac{3H}{\varepsilon^{1/2_*-1/2}\kappa\lambda}\right)^{2_*/(1-2_*\eta)}\left(\int_{\mathcal{B}_j\cap\{F>\kappa_4\kappa\lambda\}}F^{2_*}\,d\mu\right)^{1/(1-2_*\eta)}.$$

Now, by means of (5-78)–(5-79), we have

$$\begin{aligned}\int_{\mathcal{B}_j\cap\{F>\kappa_4\kappa\lambda\}}F^{2_*}\,d\mu&\le(\kappa_4\kappa\lambda)^{2_*}\int_{\mathcal{B}_j\cap\{F>\kappa_4\kappa\lambda\}}\left(\frac{F}{\kappa_4\kappa\lambda}\right)^{2_*+\delta_f}d\mu\\&\le\frac{\mu(2^{m-1}\mathcal{B}_j)}{(\kappa_4\kappa\lambda)^{\delta_f}}\fint_{2^{m-1}\mathcal{B}_j}F^{2_*+\delta_f}\,d\mu\\&\le\frac{\mu(\mathcal{B}(x_0,2\varrho_0))}{(\kappa_4\kappa\lambda)^{\delta_f}}[\Upsilon_0(2^{m-1}\mathcal{B}_j)]^{2_*+\delta_f}\le\frac{L\tilde{\lambda}_0^{2_*+\delta_f}}{(\kappa_4\kappa\lambda)^{\delta_f}},\end{aligned}\tag{5-87}$$

and hence

$$\mu(\mathcal{B}_j)\le\frac{c_4\tilde{\lambda}_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}}{(\kappa_4\kappa\lambda)^{(1+\eta\delta_f)2_*/(1-2_*\eta)}}\int_{\mathcal{B}_j\cap\{F>\kappa_4\kappa\lambda\}}F^{2_*}\,d\mu,\tag{5-88}$$

where

$$c_4:=4\left[\frac{3H(L+1)}{\varepsilon^{1/2_*-1/2}}\right]^{2_*/(1-2_*\eta)},\tag{5-89}$$

and $H$ has been defined in (5-81). A similar argument can be used in case (5-76) holds. Specifically, we have

$$\mu(\mathcal{B}_j)\le\left(\frac{3M}{\varepsilon^{1/p-1/2}\kappa\lambda}\right)^{p/(1-p\theta)}\left(\int_{\mathcal{B}_j}G^p\,d\mu\right)^{1/(1-p\theta)},$$

and then

$$\left(\int_{\mathcal{B}_j}G^p\,d\mu\right)^{\frac{1}{1-p\theta}}\le 2^{p\theta/(1-p\theta)}\left(\int_{\mathcal{B}_j\cap\{G>\kappa_5\kappa\lambda\}}G^p\,d\mu\right)^{\frac{1}{1-p\theta}}+[2(L+1)]^{p/(1-p\theta)}(\kappa_5\kappa\lambda)^{p/(1-p\theta)}\mu(\mathcal{B}_j).$$

This time we select a number $\kappa_5\in(0,1)$ such that

$$\kappa_5\le\left(\frac{1}{2}\right)^{(1-p\theta)/p}\frac{\varepsilon^{1/p-1/2}}{6M(L+1)}\tag{5-90}$$

and recall Remark 4.5 in order to get

$$\mu(\mathcal{B}_j)\le 2^{\Lambda_\theta+1}\left(\frac{3M}{\varepsilon^{1/p-1/2}\kappa\lambda}\right)^{p/(1-p\theta)}\left(\int_{\mathcal{B}_j\cap\{G>\kappa_5\kappa\lambda\}}G^p\,d\mu\right)^{1/(1-p\theta)}.$$

We then estimate as in (5-87), thereby obtaining

$$\int_{\mathcal{B}_j\cap\{G>\kappa_5\kappa\lambda\}}G^p\,d\mu\le\frac{\mu(\mathcal{B}(x_0,2\varrho_0))}{(\kappa_5\kappa\lambda)^{\delta_g}}[\Upsilon_0(2^{m-1}\mathcal{B}_j)]^{p+\delta_g}\le\frac{L\tilde{\lambda}_0^{p+\delta_g}}{(\kappa_5\kappa\lambda)^{\delta_g}},$$

and we conclude that

$$\mu(\mathcal{B}_j) \le \frac{c_5 \tilde{\lambda}_0^{(p+\delta_g)p\theta/(1-p\theta)}}{(\kappa_5\kappa\lambda)^{(1+\theta\delta_g)p/(1-p\theta)}} \int_{\mathcal{B}_j \cap \{G>\kappa_5\kappa\lambda\}} G^p \, d\mu, \tag{5-91}$$

where

$$c_5 := 2^{\Lambda_\theta+1}\left[\frac{3M(L+1)}{\varepsilon^{1/p-1/2}}\right]^{p/(1-p\theta)}. \tag{5-92}$$

All in all, taking (5-84), (5-88) and (5-91) into account, we obtain

$$\mu(\mathcal{B}_j) \le \frac{c_3}{(\kappa\lambda)^q} \int_{2\mathcal{B}_j \cap \{U>\kappa_3\kappa\lambda\}} U^q \, d\mu + \frac{c_4 \tilde{\lambda}_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}}{(\kappa_4\kappa\lambda)^{(1+\eta\delta_f)2_*/(1-2_*\eta)}} \int_{\mathcal{B}_j \cap \{F>\kappa_4\kappa\lambda\}} F^{2_*} \, d\mu \\ + \frac{c_5 \tilde{\lambda}_0^{(p+\delta_g)p\theta/(1-p\theta)}}{(\kappa_5\kappa\lambda)^{(1+\theta\delta_g)p/(1-p\theta)}} \int_{\mathcal{B}_j \cap \{G>\kappa_5\kappa\lambda\}} G^p \, d\mu.$$

Since $\{2\mathcal{B}_j\}_j$ is a disjoint family and all members belong to $\mathcal{B}(x_0, s)$, we have that

$$\sum_{j\in J_D} \mu(\mathcal{B}_j) \le \frac{c_3}{(\kappa\lambda)^q} \int_{\mathcal{B}(x_0,s) \cap \{U>\kappa_3\kappa\lambda\}} U^q \, d\mu + \frac{c_4 \tilde{\lambda}_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}}{(\kappa_4\kappa\lambda)^{(1+\eta\delta_f)2_*/(1-2_*\eta)}} \int_{\mathcal{B}(x_0,s) \cap \{F>\kappa_4\kappa\lambda\}} F^{2_*} \, d\mu \\ + \frac{c_5 \tilde{\lambda}_0^{(p+\delta_g)p\theta/(1-p\theta)}}{(\kappa_5\kappa\lambda)^{(1+\theta\delta_g)p/(1-p\theta)}} \int_{\mathcal{B}(x_0,s) \cap \{G>\kappa_5\kappa\lambda\}} G^p \, d\mu. \tag{5-93}$$

The constants $c_3, c_4, c_5$ have been defined in (5-84), (5-89) and (5-92), respectively, while the numbers $\kappa, \kappa_3, \kappa_4, \kappa_5 \in (0, 1)$ must be taken in order to satisfy (5-55), (5-83), (5-86) and (5-90), respectively.

**5L.** ***Conclusion of the proof.*** We start by combining (5-73) and (5-93). Using the elementary estimate

$$\int_{\mathcal{B}(x_0,t) \cap \{U>\kappa_3\kappa\lambda\}} U^2 \, d\mu \le \lambda^{2-q} \int_{\mathcal{B}(x_0,t) \cap \{U>\kappa_3\kappa\lambda\}} U^q \, d\mu + \int_{\mathcal{B}(x_0,t) \cap \{U>\lambda\}} U^2 \, d\mu,$$

(5-73) and (5-93) yield, after a few elementary manipulations, the estimate

$$\int_{\mathcal{B}(x_0,t) \cap \{U>\kappa_3\kappa\lambda\}} U^2 \, d\mu \le \frac{c}{(\kappa_3\kappa)^{2-q}} (\kappa_3\kappa\lambda)^{2-q} \int_{\mathcal{B}(x_0,s) \cap \{U>\kappa_3\kappa\lambda\}} U^q \, d\mu \\ + \frac{c_4 \tilde{\lambda}_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}}{\kappa_4^2 (\kappa_4\kappa\lambda)^{(1+\eta\delta_f)2_*/(1-2_*\eta)-2}} \int_{\mathcal{B}(x_0,s) \cap \{F>\kappa_4\kappa\lambda\}} F^{2_*} \, d\mu \\ + \frac{c_5 \tilde{\lambda}_0^{(p+\delta_g)p\theta/(1-p\theta)}}{\kappa_5^2 (\kappa_5\kappa\lambda)^{(1+\theta\delta_g)p/(1-p\theta)-2}} \int_{\mathcal{B}(x_0,s) \cap \{G>\kappa_5\kappa\lambda\}} G^p \, d\mu. \tag{5-94}$$

The constant $c$ appearing above depends on $n, \alpha, \Lambda$, but is still independent of $\varepsilon$, and we have also used the fact that $\kappa, \kappa_3 \in (0,1)$. We can therefore reformulate estimate (5-94) as

$$\int_{\mathcal{B}(x_0,t)\cap\{U>\lambda\}} U^2\,d\mu \le \frac{c\lambda^{2-q}}{(\kappa_3\kappa\varepsilon)^{2-q}}\int_{\mathcal{B}(x_0,s)\cap\{U>\lambda\}} U^q\,d\mu + \frac{c_6\tilde{\lambda}_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}}{\lambda^{(1+\eta\delta_f)2_*/(1-2_*\eta)-2}}\int_{\mathcal{B}(x_0,s)\cap\{F>\kappa_4\lambda/\kappa_3\}} F^{2_*}\,d\mu + \frac{c_7\tilde{\lambda}_0^{(p+\delta_g)p\theta/(1-p\theta)}}{\lambda^{(1+\theta\delta_g)p/(1-p\theta)-2}}\int_{\mathcal{B}(x_0,s)\cap\{G>\kappa_5\lambda/\kappa_3\}} G^p\,d\mu. \tag{5-95}$$

The constant $c \equiv c(n, \alpha, \Lambda)$ is independent of $\varepsilon$, while

$$c_6 \equiv c_6(n, \alpha, \Lambda, L, \varepsilon) \quad \text{and} \quad c_7 \equiv c_7(n, \alpha, \Lambda, \beta, \gamma, p, L, \varepsilon);$$

the constant $c_7$ exhibits a blow-up behavior with respect to $p$ as described in (3-5). Since estimate (5-94) holds for $\lambda \ge \lambda_1$ — and $\lambda_1$ has been defined in (5-33) — we have that (5-95) holds whenever $\lambda \ge \kappa\kappa_3\lambda_1$. We remark that the previous inequality holds for a choice of $\kappa, \kappa_3, \kappa_4, \kappa_5 \in (0,1)$ that satisfy (5-55), (5-83), (5-86) and (5-90), respectively. In order to conclude with (5-7) we now need to estimate a few constants. We are primarily interested in an explicit dependence on $\varepsilon$ in the second integral appearing in (5-95). We therefore look at (5-55) and (5-83), and we infer that we can in fact choose $\kappa, \kappa_3$ in order to have

$$\kappa_3\kappa \approx \frac{\varepsilon^{3/q-1}}{c_*}, \tag{5-96}$$

for a constant $c_*$ which is now independent of $\varepsilon$, but just depends on $n, \alpha, \Lambda$. We next find an upper bound for the numbers $\tilde{\lambda}_0$ and $\lambda_1$ introduced in (5-10) and (5-33), respectively; this will allow us to verify estimate (5-7) in the range dictated by (5-8). Let us notice that if $x \in B(x_0,t)$ and $(s-t)/40^n \le \varrho \le \frac{1}{2}\varrho_0$, then $\mathcal{B}(x,\varrho) \subset \mathcal{B}(x_0, 2\varrho_0)$. Therefore, recalling (4-2), whenever $\widetilde{U}$ is a $\mu$-integrable function we can estimate

$$\fint_{\mathcal{B}(x,\varrho)} \widetilde{U}\,d\mu \le \frac{\mu(\mathcal{B}(x_0,2\varrho_0))}{\mu(\mathcal{B}(x,\varrho))}\fint_{\mathcal{B}(x_0,2\varrho_0)} \widetilde{U}\,d\mu \le c\Big(\frac{\varrho_0}{s-t}\Big)^{n+2\varepsilon}\fint_{\mathcal{B}(x_0,2\varrho_0)} \widetilde{U}\,d\mu \tag{5-97}$$

for a constant $c$ depending on $n$ but independent of $\varepsilon$. Applying the inequality in the last display to $U^2, G^p, F^{2_*}, G^{p+\delta_g}$ and $F^{2_*+\delta_f}$ — and eventually on different balls $2^k\mathcal{B}(x,\varrho) \subset 2^k\mathcal{B}(x_0,2\varrho_0)$ — yields

$$\begin{aligned}
&\kappa^{-1}\{\Psi_{H,M}(x,\varrho) + \Upsilon_0(x,\varrho) + \Upsilon_1(x,\varrho) + \Upsilon_{2,M}(x,\varrho)\} \\
&\quad \le \frac{c}{\varepsilon^{1/q}}\Big(\frac{\varrho_0}{s-t}\Big)^{n+2\varepsilon}\{\Psi_{H,M}(x_0,2\varrho_0) + \Upsilon_0(x_0,2\varrho_0) + \Upsilon_1(x_0,2\varrho_0) + \Upsilon_{2,M}(x_0,2\varrho_0)\} \\
&\quad \le \frac{c}{\varepsilon^{1/q}}\Big(\frac{\varrho_0}{s-t}\Big)^{n+2\varepsilon}\,\mathrm{ADD}(x_0,2\varrho_0).
\end{aligned} \tag{5-98}$$

In order, we have also used (5-56), (5-81) to get rid of the presence of $M$ and $H$ and that $\varrho_0/(s-t)$ is bounded away from zero. We recall that the functional $\mathrm{ADD}(\cdot)$ has been introduced in (5-5). We now obtain an upper bound for $\lambda_1$ defined in (5-33). The quantity appearing on the right-hand side of (5-98) provides an upper bound on $\tilde{\lambda}_0$. In a similar way, if $\mathcal{K} = K_1 \times K_2 \in \Xi_{k_0}$, with $k_0$ as in (5-19), then

$\mathcal{K} \subset \mathcal{B}(x_0, s) \subset \mathcal{B}(x_0, 2\varrho_0)$ and therefore we have

$$\mu(\mathcal{K}) \geq \frac{c}{\varrho_0^{n-2\varepsilon}} \int_{K_1} \int_{K_2} dx\, dy = \frac{c(s-t)^{2n}}{\varrho_0^{n-2\varepsilon}}.$$

Hence, as for (5-97), we have

$$\fint_{\mathcal{K}} \widetilde{U}\, d\mu \leq \frac{\mu(\mathcal{B}(x_0, 2\varrho_0))}{\mu(\mathcal{K})} \fint_{\mathcal{B}(x_0,2\varrho_0)} \widetilde{U}\, d\mu \leq \frac{c}{\varepsilon}\Big(\frac{\varrho_0}{s-t}\Big)^{2n} \fint_{\mathcal{B}(x_0,2\varrho_0)} \widetilde{U}\, d\mu. \tag{5-99}$$

By using (5-98)-(5-99), and recalling that $\varepsilon < 1$, we get

$$\lambda_1 \leq \frac{c}{\varepsilon}\Big(\frac{\varrho_0}{s-t}\Big)^{2n} \operatorname{ADD}(x_0, 2\varrho_0)$$

where $c$ depends only on $n, \alpha, \Lambda, \beta, p, \gamma, \varepsilon$. Summarizing the content of these manipulations, we can finally arrive at (5-7), with the restriction on $\lambda$ described in (5-8). Specifically, we use (5-96) to estimate the constant in front of the second integral appearing in (5-95), and the bounds found for $\tilde{\lambda}_0$ and $\lambda_1$ to conclude with the admissible range of values $\lambda \geq \lambda_0$ described via (5-8). Needless to say, we are taking $\kappa_f := \kappa_4/\kappa_3$ and $\kappa_g := \kappa_5/\kappa_3$.

## 6. Self-improving inequalities

This section is dedicated to the proof of a fractional reverse Hölder-type inequality on diagonal balls with increasing supports, that is, the estimate (6-1) below. This will eventually imply Theorem 1.1 at the end of the section.

**Theorem 6.1** (reverse Hölder-type inequality). *Let $u \in W^{\alpha,2}(\mathbb{R}^n)$ be a solution to* (1-14) *under the assumptions of Theorem 1.1*; *in particular*, (3-1) *and* (3-3) *are in force. Define the functions $U$, $F$ and $G$ as in* (4-5). *Then there exist positive constants $\varepsilon \in (0, 1-\alpha)$, $\delta \in (0, 1)$ and $c_8 \geq 1$, depending on $n, \alpha, \Lambda, \beta, p, \gamma, \delta_1$, such that whenever $\mathcal{B} \equiv \mathcal{B}(x_0, \varrho_0) \subset \mathbb{R}^{2n}$ we have the inequality*

$$\begin{aligned}\Big(\fint_{\mathcal{B}} U^{2+\delta}\, d\mu\Big)^{1/(2+\delta)} &\leq c_8 \sum_{k=1}^{\infty} 2^{-k(\alpha-\varepsilon)} \Big(\fint_{2^k\mathcal{B}} U^2\, d\mu\Big)^{1/2} \\ &\quad + c_8 \varrho_0^{\alpha-\varepsilon} \Big(\fint_{2\mathcal{B}} F^{2_*+\delta_0}\, d\mu\Big)^{1/(2_*+\delta_0)} + c_8 \varrho_0^{\gamma-2\beta+\alpha+\varepsilon(2/p-1)} \Big(\fint_{2\mathcal{B}} G^{p(1+\delta_1)}\, d\mu\Big)^{1/[p(1+\delta_1)]} \\ &\quad + c_8 \varrho_0^{\gamma-2\beta+\alpha+\varepsilon(2/p-1)} \sum_{k=1}^{\infty} 2^{-k(2\beta-\gamma-2\varepsilon/p)} \Big(\fint_{2^k\mathcal{B}} G^p\, d\mu\Big)^{1/p}. \end{aligned} \tag{6-1}$$

*All the terms on the right-hand side of this inequality are finite.*

*Proof. Step 1: Determining the exponents.* Let us observe that, whenever $\varepsilon \in (0, \frac{1}{2}\alpha)$, we have

$$\frac{8\varepsilon}{n+2\varepsilon} < \frac{2\varepsilon(n+2\alpha)}{n(\alpha-\varepsilon)}.$$

Therefore, we can always find two positive numbers $\varepsilon \in (0, \frac{1}{2}\alpha)$ and $\delta_f > 0$, satisfying (4-6) and $\delta_f \le \delta_0$, respectively, such that

$$\frac{8\varepsilon}{n+2\varepsilon} < \delta_f \le \frac{2\varepsilon(n+2\alpha)}{n(\alpha-\varepsilon)} \quad \text{and} \quad \varepsilon < 1-\alpha. \tag{6-2}$$

We recall that $F \in L^{2_*+\delta_f}_{\mathrm{loc}}(\mathbb{R}^n; \mu)$ by (4-7). Next, we determine the positive number $\delta > 0$ by imposing different restrictions on it; we start by assuming that

$$\delta \le \frac{4\varepsilon(n+2\alpha)}{n^2+4\varepsilon(n+\alpha)} \quad \text{and} \quad \delta \le \frac{(\gamma-2\beta+\alpha)\delta_g}{4n}. \tag{6-3}$$

Let us briefly discuss a few consequences of the two conditions above, starting with the first one. Specifically, we start by showing that

$$\delta \le \delta_f\left[\frac{(n+2\alpha)(n+2\varepsilon)}{n^2+4\varepsilon(n+\alpha)}\right] - \frac{4\varepsilon(n+2\alpha)}{n^2+4\varepsilon(n+\alpha)}. \tag{6-4}$$

Indeed, using the first inequality in (6-3), we have

$$\begin{aligned}\delta \le \frac{4\varepsilon(n+2\alpha)}{n^2+4\varepsilon(n+\alpha)} &= \frac{8\varepsilon}{n+2\varepsilon}\frac{(n+2\alpha)(n+2\varepsilon)}{n^2+4\varepsilon(n+\alpha)} - \frac{4\varepsilon(n+2\alpha)}{n^2+4\varepsilon(n+\alpha)}\\ &\le \delta_f\left[\frac{(n+2\alpha)(n+2\varepsilon)}{n^2+4\varepsilon(n+\alpha)}\right] - \frac{4\varepsilon(n+2\alpha)}{n^2+4\varepsilon(n+\alpha)}.\end{aligned}$$

Next, the definition in (4-16) and the fact that $\varepsilon < \frac{1}{2}\alpha$ gives that $1 > \theta > (\gamma-2\beta+\alpha)/(n+\alpha)$. Then the fact that the function $t \to t/(1-t)$ is increasing in the interval $(0, 1)$ allows to estimate

$$\frac{\gamma-2\beta+\alpha}{2n} \le \frac{\gamma-2\beta+\alpha}{n-\gamma+2\beta} \le \frac{\theta}{1-\theta} < \frac{p\theta}{1-p\theta},$$

so that, from the second inequality in (6-3), it follows that

$$\delta < \frac{(\gamma-2\beta+\alpha)\delta_g}{4n} \le \frac{\delta_g}{2}\frac{p\theta}{1-p\theta}. \tag{6-5}$$

Finally, for $t \in (0, 1)$, we define the function

$$S(t) := \frac{2c_s(n+4)}{4\alpha t^6} \ge \frac{2c_s}{(2-q)t^{3(2-q)/q}}, \tag{6-6}$$

where $c_s$ is the constant introduced in Proposition 5.1 and $q$ has been introduced in (4-13); in the last estimation we have used that $\varepsilon \in (0, \frac{1}{2}\alpha)$. We then impose the last restriction on $\delta$, that is,

$$\delta S(\varepsilon) \le \tfrac{1}{4}. \tag{6-7}$$

All in all, the choices made in (6-3) and (6-7) allow us to determine $\delta$ as a positive number depending only on $n, \alpha, \Lambda, \beta, p, \gamma, \delta_1$, as required in the statement of Theorem 6.1.

*Step 2: Reverse Hölder-type inequalities*. In this step, by applying Proposition 5.1 with the numbers $\varepsilon, \delta, \delta_f$ as chosen in Step 1, we are going to prove that $U \in L^{2+\delta}_{\mathrm{loc}}(\mathbb{R}^{2n}; \mu)$. The finiteness of the terms on the right-hand side of (6-1) has already been discussed in Section 4C. First of all, we show that we can reduce to the case $\varrho_0 = 1$ and $\mathcal{B} = B(0, 1) \times B(0, 1)$; this eventually allows us to apply Proposition 5.1. Indeed, notice that the rescaled functions

$$\tilde{u}(x) := u(x_0 + \varrho_0 x), \quad \tilde{g}(x) := \varrho_0^{2\alpha-2\beta} g(x_0 + \varrho_0 x), \quad \tilde{f}(x) := \varrho_0^{2\alpha} f(x_0 + \varrho_0 x),$$

still solve (1-14). Therefore, applying (6-1) in this case and in $B(0, 1) \times B(0, 1)$, and scaling back to the original functions and to the original diagonal ball $\mathcal{B}$, leads to (6-1) in the general case. We now pass to the proof of (6-1) when $\mathcal{B} = B(0, 1) \times B(0, 1)$. We define the truncated function $U_m := \min\{U, m\}$ for $m$ being a positive integer, and the measure $d\nu = U^2\, d\mu$. Moreover, we use the abbreviation $\mathcal{B}_s := \mathcal{B}(0, s)$. With the aim of applying Proposition 5.1, we then consider balls

$$\mathcal{B} \equiv \mathcal{B}_1 \subset \mathcal{B}_t \subset \mathcal{B}_s \subset \mathcal{B}_2$$

as in (5-6), while $\lambda_0$ is accordingly defined as in (5-8). We shall derive uniform higher integrability for the functions $U_m$ and will recover the final result by letting $m \to \infty$. With $\delta \in (0, 1)$ being the number determined in Step 1, by Cavalieri's principle we have that

$$\begin{aligned}\int_{\mathcal{B}_t} U_m^\delta U^2\, d\mu &= \int_{\mathcal{B}_t} U_m^\delta\, d\nu \\ &= \delta \int_0^\infty \lambda^{\delta-1} \nu(\mathcal{B}_t \cap \{U_m > \lambda\})\, d\lambda \\ &= \delta \int_0^m \lambda^{\delta-1} \int_{\mathcal{B}_t \cap \{U > \lambda\}} U^2\, d\mu\, d\lambda \\ &\le \lambda_0^\delta \int_{\mathcal{B}_t} U^2\, d\mu + \delta \int_{\lambda_0}^m \lambda^{\delta-1} \int_{\mathcal{B}_t \cap \{U > \lambda\}} U^2\, d\mu\, d\lambda. \end{aligned} \tag{6-8}$$

The second-last integral appearing in this display can be easily estimated by recalling the definition of $\lambda_0$ in (5-8) and that $\varrho_0/(s-t) \ge 1$, and using (4-2):

$$\lambda_0^\delta \int_{\mathcal{B}_t} U^2\, d\mu \le \mu(\mathcal{B}_2) \lambda_0^\delta \fint_{2\mathcal{B}} U^2\, d\mu \le c\mu(\mathcal{B}_1) \lambda_0^{2+\delta}. \tag{6-9}$$

We proceed with the remaining term in (6-8); using (5-7) we gain

$$\begin{aligned}\delta \int_{\lambda_0}^m \lambda^{\delta-1} \int_{\mathcal{B}_t \cap \{U > \lambda\}} U^2\, d\mu\, d\lambda &\le \frac{c_s \delta}{\varepsilon^{3(2-q)/q}} \int_{\lambda_0}^m \lambda^{\delta+1-q} \int_{\mathcal{B}_s \cap \{U > \lambda\}} U^q\, d\mu\, d\lambda \\ &\quad + c_f \delta \int_{\lambda_0}^m \frac{\lambda_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}}{\lambda^{(1+\eta\delta_f)2_*/(1-2_*\eta)-1-\delta}} \int_{\mathcal{B}_s \cap \{F > \kappa_f \lambda\}} F^{2_*}\, d\mu\, d\lambda \\ &\quad + c_g \delta \int_{\lambda_0}^m \frac{\lambda_0^{(p+\delta_g)p\theta/(1-p\theta)}}{\lambda^{(1+\theta\delta_g)p/(1-p\theta)-1-\delta}} \int_{\mathcal{B}_s \cap \{G > \kappa_g \lambda\}} G^p\, d\mu\, d\lambda \\ &=: \mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3. \end{aligned} \tag{6-10}$$

Using (6-6)–(6-7) and Fubini's theorem, we get

$$
\begin{aligned}
\mathcal{J}_1 &\le \frac{c_s\delta}{\varepsilon^{3(2-q)/q}}\int_0^\infty \lambda^{\delta+1-q}\int_{\mathcal{B}_s\cap\{U_m>\lambda\}} U^q\,d\mu\,d\lambda \\
&= \frac{c_s\delta}{(\delta+2-q)\varepsilon^{3(2-q)/q}}\int_{\mathcal{B}_s} U_m^{2+\delta-q}U^q\,d\mu \\
&\le \delta S(\varepsilon)\int_{\mathcal{B}_s} U_m^\delta U^2\,d\mu \le \frac{1}{4}\int_{\mathcal{B}_s} U_m^\delta U^2\,d\mu. \qquad (6\text{-}11)
\end{aligned}
$$

We next estimate $\mathcal{J}_2$. Changing variables, using Fubini's theorem, and recalling the dependence $\kappa_f \equiv \kappa_f(n,\alpha,\Lambda,\varepsilon)$, we have

$$
\begin{aligned}
&\int_{\lambda_0}^m \lambda^{\delta+1-(1+\eta\delta_f)2_*/(1-2_*\eta)}\int_{\mathcal{B}_s\cap\{F>\kappa_f\lambda\}} F^{2_*}\,d\mu\,d\lambda \\
&\quad\le c\int_0^\infty \lambda^{\delta+1-(1+\eta\delta_f)2_*/(1-2_*\eta)}\int_{\mathcal{B}_s\cap\{F>\lambda\}} F^{2_*}\,d\mu\,d\lambda \\
&\quad= \frac{c\mu(\mathcal{B}_2)}{\delta+2-(1+\eta\delta_f)2_*/(1-2_*\eta)}\fint_{\mathcal{B}_2} F^{\delta+2-(1+\eta\delta_f)2_*/(1-2_*\eta)+2_*}\,d\mu \\
&\quad\le \frac{c\mu(\mathcal{B}_2)}{\delta}\fint_{\mathcal{B}_2} F^{\delta+2-(1+\eta\delta_f)2_*/(1-2_*\eta)+2_*}\,d\mu, \qquad (6\text{-}12)
\end{aligned}
$$

again for a constant depending on $n$, $\alpha$, $\Lambda$ and $\varepsilon$. In writing the last inequality we have used that (6-2) is in force and the fact that

$$
\delta_f \le \frac{2\varepsilon(n+2\alpha)}{n(\alpha-\varepsilon)} \quad\Longleftrightarrow\quad 2-\frac{(1+\eta\delta_f)2_*}{1-2_*\eta}\ge 0.
$$

The last integral appearing in (6-12) is finite if $\delta+2-(1+\eta\delta_f)2_*/(1-2_*\eta)+2_*\le 2_*+\delta_f$, and a lengthy computation shows that this is equivalent to (6-4). Therefore, using Hölder's inequality, we can estimate

$$
\begin{aligned}
\mathcal{J}_2 &\le c\mu(\mathcal{B}_2)\lambda_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)}\left(\fint_{\mathcal{B}_2} F^{2_*+\delta_f}\,d\mu\right)^{\frac{\delta+2-(1+\eta\delta_f)2_*/(1-2_*\eta)+2_*}{2_*+\delta_f}} \\
&\le c\mu(\mathcal{B}_1)\lambda_0^{(2_*+\delta_f)2_*\eta/(1-2_*\eta)+\delta+2-(1+\eta\delta_f)2_*/(1-2_*\eta)+2_*} \\
&= c\mu(\mathcal{B}_1)\lambda_0^{2+\delta}, \qquad (6\text{-}13)
\end{aligned}
$$

where $c$ depends only on $n$, $\alpha$, $\Lambda$ and $\varepsilon$. We finally come to the estimation of $\mathcal{J}_3$. For this we notice that the definitions of $p$ and $\theta$ give, independently of $\varepsilon$, that

$$
p \ge \frac{2n}{n+2(\gamma-2\beta+\alpha)} \quad\Longleftrightarrow\quad \frac{p}{1-p\theta}\ge 2, \qquad (6\text{-}14)
$$

and then, recalling that $\kappa_g \equiv \kappa_g(n, \alpha, \Lambda, \varepsilon, \gamma, \beta, p)$, we have

$$\begin{aligned}\int_{\lambda_0}^{m} \lambda^{\delta+1-(1+\theta\delta_g)p/(1-p\theta)} \int_{\mathcal{B}_s\cap\{G>\kappa_g\lambda\}} G^p \, d\mu \, d\lambda &\le \int_{\lambda_0}^{\infty} \lambda^{\delta+1-(1+\theta\delta_g)p/(1-p\theta)} \, d\lambda \int_{\mathcal{B}_s} G^p \, d\mu \\ &\le \frac{c\lambda_0^{\delta+2-(1+\theta\delta_g)p/(1-p\theta)}\mu(\mathcal{B}_2)}{(1+\theta\delta_g)p/(1-p\theta)-\delta-2} \fint_{\mathcal{B}_2} G^p \, d\mu \\ &\le \frac{c\lambda_0^{\delta+2-(1+\theta\delta_g)p/(1-p\theta)}\mu(\mathcal{B}_2)}{\theta\delta_g p/(1-p\theta)-\delta} \left(\fint_{\mathcal{B}_2} G^{p+\delta_g} \, d\mu\right)^{p/(p+\delta_g)} \\ &\le \frac{c}{\delta}\lambda_0^{\delta+2-(1+\theta\delta_g)p/(1-p\theta)+p}\mu(\mathcal{B}_2).\end{aligned}$$

Observe that in order to perform the last two estimations we have also used (6-14) and (6-5), respectively. Therefore we can estimate as in (6-13), that is,

$$\mathcal{J}_3 \le c\mu(\mathcal{B}_2)\lambda_0^{(p+\delta_g)p\theta/(1-p\theta)+\delta+2-(1+\theta\delta_g)p/(1-p\theta)+p} = c\mu(\mathcal{B}_2)\lambda_0^{2+\delta}, \tag{6-15}$$

with $c \equiv c(n, \alpha, \Lambda, \varepsilon, \gamma, \beta, p)$. Connecting (6-11), (6-13) and (6-15) to (6-10), and combining the resulting inequality with (6-8) and (6-9), we get

$$\int_{\mathcal{B}_t} U_m^\delta U^2 \, d\mu \le \frac{1}{4}\int_{\mathcal{B}_s} U_m^\delta U^2 \, d\mu + c\mu(\mathcal{B}_1)\lambda_0^{2+\delta}.$$

By recalling the definition of $\lambda_0$ in (5-8), and using several times the doubling property of $\mu$, after a few elementary manipulations we come to

$$\left(\fint_{\mathcal{B}_t} U_m^\delta U^2 \, d\mu\right)^{1/(2+\delta)} \le \frac{1}{2}\left(\fint_{\mathcal{B}_s} U_m^\delta U^2 \, d\mu\right)^{1/(2+\delta)} + \frac{c}{\varepsilon}\left(\frac{\varrho_0}{s-t}\right)^{2n} \mathrm{ADD}(2\mathcal{B}).$$

We can therefore rewrite the above inequality as

$$\phi(t) \le \tfrac{1}{2}\phi(s) + \frac{c}{\varepsilon}\left(\frac{\varrho_0}{s-t}\right)^{2n} \mathrm{ADD}(2\mathcal{B})$$

for a constant $c \equiv c(n, \alpha, \Lambda, \varepsilon, \gamma, \beta, p)$ which is still independent of $m \in \mathbb{N}$, and where, obviously, we have set

$$\phi(\varrho) := \left(\fint_{\mathcal{B}_\varrho} U_m^\delta U^2 \, d\mu\right)^{1/(2+\delta)}$$

for $\varrho \in [\varrho_0, \frac{3}{2}\varrho_0]$. We are therefore in position to apply the standard iteration Lemma 6.2 below, which gives, after returning to the full notation,

$$\left(\fint_{\mathcal{B}} U_m^\delta U^2 \, d\mu\right)^{1/(2+\delta)} \le c \, \mathrm{ADD}(2\mathcal{B}).$$

The previous inequality holds for a constant $c \equiv c(n, \alpha, \Lambda, \varepsilon, \gamma, \beta, p)$ which is independent of $m \in \mathbb{N}$. Therefore, letting $m \to \infty$ yields

$$\left( \fint_{\mathcal{B}} U^{2+\delta} \, d\mu \right)^{1/(2+\delta)} \le c \, \mathrm{ADD}(2\mathcal{B}).$$

At this point (6-1) follows by recalling the definition of $\mathrm{ADD}(2\mathcal{B})$ in (5-5) and using a few elementary manipulations involving Hölder's inequality. In particular, we use the fact that $2_* + \delta_f \le 2_* + \delta_0$ and $p + \delta_g \le p(1 + \delta_1)$; see Lemma 4.2. □

**Lemma 6.2.** *Let $\phi\colon [\varrho_0, \frac{3}{2}\varrho_0] \to [0, \infty)$ be a function such that*

$$\phi(t) \le \tfrac{1}{2}\phi(s) + \frac{A}{(s-t)^\gamma}$$

*whenever $\varrho_0 < t < s < \frac{3}{2}\varrho_0$, where $A$ and $\gamma$ are positive constants. Then the inequality*

$$\phi(\varrho_0) \le \frac{cA}{\varrho_0^\gamma}$$

*holds for a constant $c \equiv c(\gamma)$.*

For a proof of this lemma, see for instance [Giusti 2003, Chapter 6].

*Proof of Theorem 1.1.* The proof is now a simple consequence of Theorem 6.1, that gives that $U \in L^{2+\delta}(\mathcal{B}; \mu)$ whenever $\mathcal{B} = B \times B$ and $B \subset \mathbb{R}^n$ is a ball (that for simplicity we take to be centered at the origin). We now translate this information in terms of fractional norms of the original function $u$. In fact, this means that, whenever $B \subset \mathbb{R}^n$ is a ball centered at the origin, we have

$$\int_{B\times B} U^{2+\delta} \, d\mu = \int_B \int_B \frac{|u(x) - u(y)|^{2+\delta}}{|x-y|^{n+(2+\delta)\alpha+\varepsilon\delta}} \, dx \, dy < \infty.$$

Rewriting the last integral, we find

$$\int_B \int_B \frac{|u(x) - u(y)|^{2+\delta}}{|x-y|^{n+(2+\delta)[\alpha+\varepsilon\delta/(2+\delta)]}} \, dx \, dy < \infty$$

whenever $B \subset \mathbb{R}^n$ is a ball, and this means that $u \in W^{\alpha+\varepsilon\delta/(2+\delta),2+\delta}_{\mathrm{loc}}(\mathbb{R}^n)$; observe that since $\varepsilon < 1 - \alpha$ then $\alpha + \varepsilon\delta/(2+\delta) < 1$. We have therefore improved the regularity of $u$ both in the fractional and in the differentiability scale, and Theorem 1.1 follows by suitably renaming (via embedding theorems) the number $\delta$ considered in its statement. □

*Proof of Theorem 1.3.* The proof is just a consequence of the arguments developed to prove Theorem 6.1. In fact the only thing needed there is Proposition 4.4, whose content is now considered as an assumption in (1-23), provided that we take $F = G = 0$; the rest of the argument then remains unchanged. □

## Acknowledgments

We wish to thank Vladimir Maz'ya for a useful discussion and Paolo Baroni for a careful reading of a preliminary version of the manuscript. We also thank the referees for their extremely valuable work and their careful reading of the first draft of the paper; their comments led to an improved version.

## References

[Bass and Kassmann 2005] R. F. Bass and M. Kassmann, "Hölder continuity of harmonic functions with respect to operators of variable order", *Comm. Partial Differential Equations* **30**:7-9 (2005), 1249–1259. MR 2006i:31005 Zbl 1087.45004

[Bass and Ren 2013] R. F. Bass and H. Ren, "Meyers inequality and strong stability for stable-like operators", *J. Funct. Anal.* **265**:1 (2013), 28–48. MR 3049880 Zbl 1295.47047

[Bjorland et al. 2012] C. Bjorland, L. Caffarelli, and A. Figalli, "Non-local gradient dependent operators", *Adv. Math.* **230**:4-6 (2012), 1859–1894. MR 2927356 Zbl 1252.35099

[Bojarski and Iwaniec 1983] B. Bojarski and T. Iwaniec, "Analytical foundations of the theory of quasiconformal mappings in $\mathbf{R}^n$", *Ann. Acad. Sci. Fenn. Ser. A I Math.* **8**:2 (1983), 257–324. MR 85h:30023 Zbl 0548.30016

[Cabré and Cinti 2014] X. Cabré and E. Cinti, "Sharp energy estimates for nonlinear fractional diffusion equations", *Calc. Var. Partial Differential Equations* **49**:1-2 (2014), 233–269. MR 3148114 Zbl 1282.35399

[Cabré and Roquejoffre 2013] X. Cabré and J.-M. Roquejoffre, "The influence of fractional diffusion in Fisher-KPP equations", *Comm. Math. Phys.* **320**:3 (2013), 679–722. MR 3057187 Zbl 06179645

[Caffarelli and Silvestre 2011] L. Caffarelli and L. Silvestre, "The Evans-Krylov theorem for nonlocal fully nonlinear equations", *Ann. of Math.* (2) **174**:2 (2011), 1163–1187. MR 2831115 Zbl 1232.49043

[Caffarelli et al. 2011] L. Caffarelli, C. H. Chan, and A. Vasseur, "Regularity theory for parabolic nonlinear integral operators", *J. Amer. Math. Soc.* **24**:3 (2011), 849–869. MR 2012c:45024 Zbl 1223.35098

[Di Castro et al. 2014a] A. Di Castro, T. Kuusi, and G. Palatucci, "Local behavior of fractional $p$-minimizers", preprint, 2014, Available at http://cvgmt.sns.it/paper/2379/.

[Di Castro et al. 2014b] A. Di Castro, T. Kuusi, and G. Palatucci, "Nonlocal Harnack inequalities", *J. Funct. Anal.* **267**:6 (2014), 1807–1836. MR 3237774 Zbl 06330975

[Di Nezza et al. 2012] E. Di Nezza, G. Palatucci, and E. Valdinoci, "Hitchhiker's guide to the fractional Sobolev spaces", *Bull. Sci. Math.* **136**:5 (2012), 521–573. MR 2944369 Zbl 1252.46023

[Dorronsoro 1985] J. R. Dorronsoro, "A characterization of potential spaces", *Proc. Amer. Math. Soc.* **95**:1 (1985), 21–31. MR 86k:46046 Zbl 0577.46035

[Felsinger and Kassmann 2013] M. Felsinger and M. Kassmann, "Local regularity for parabolic nonlocal operators", *Comm. Partial Differential Equations* **38**:9 (2013), 1539–1573. MR 3169755 Zbl 1277.35090

[Fusco and Sbordone 1990] N. Fusco and C. Sbordone, "Higher integrability of the gradient of minimizers of functionals with nonstandard growth conditions", *Comm. Pure Appl. Math.* **43**:5 (1990), 673–683. MR 91b:49015 Zbl 0727.49021

[Gehring 1973] F. W. Gehring, "The $L^p$-integrability of the partial derivatives of a quasiconformal mapping", *Acta Math.* **130** (1973), 265–277. MR 53 #5861 Zbl 0258.30021

[Giusti 2003] E. Giusti, *Direct methods in the calculus of variations*, World Scientific Publishing Co., River Edge, NJ, 2003. MR 2004g:49003 Zbl 1028.49001

[Iwaniec 1998] T. Iwaniec, "The Gehring lemma", pp. 181–204 in *Quasiconformal mappings and analysis* (Ann Arbor, MI, 1995), edited by P. Duren et al., Springer, New York, 1998. MR 99e:30012 Zbl 0888.30017

[Kassmann 2009] M. Kassmann, "A priori estimates for integro-differential operators with measurable kernels", *Calc. Var. Partial Differential Equations* **34**:1 (2009), 1–21. MR 2010b:35474 Zbl 1158.35019

[Kinnunen and Lewis 2000] J. Kinnunen and J. L. Lewis, "Higher integrability for parabolic systems of $p$-Laplacian type", *Duke Math. J.* **102**:2 (2000), 253–271. MR 2001b:35152 Zbl 0994.35036

[Kuusi and Mingione 2012] T. Kuusi and G. Mingione, "Universal potential estimates", *J. Funct. Anal.* **262**:10 (2012), 4205–4269. MR 2900466 Zbl 1252.35097

[Kuusi et al. 2014] T. Kuusi, G. Mingione, and Y. Sire, "A fractional Gehring lemma, with applications to nonlocal equations", *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.* **25**:4 (2014), 345–358. MR 3283394

[Maz'ya 2011] V. Maz'ya, *Sobolev spaces with applications to elliptic partial differential equations*, 2nd ed., Grundlehren der Mathematischen Wissenschaften **342**, Springer, Heidelberg, 2011. MR 2012a:46056 Zbl 1217.46002

[Meyers 1963] N. G. Meyers, "An $L^p$-estimate for the gradient of solutions of second order elliptic divergence equations", *Ann. Scuola Norm. Sup. Pisa* (3) **17** (1963), 189–206. MR 28 #2328 Zbl 0127.31904

[Meyers and Elcrat 1975] N. G. Meyers and A. Elcrat, "Some results on regularity for solutions of non-linear elliptic systems and quasi-regular functions", *Duke Math. J.* **42** (1975), 121–136. MR 54 #5618 Zbl 0347.35039

[Mingione 2003] G. Mingione, "The singular set of solutions to non-differentiable elliptic systems", *Arch. Ration. Mech. Anal.* **166**:4 (2003), 287–301. MR 2004b:35070 Zbl 1142.35391

[Stein 1961] E. M. Stein, "The characterization of functions arising as potentials", *Bull. Amer. Math. Soc.* **67** (1961), 102–104. MR 23 #A1051

[Stein 1993] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series **43**, Princeton University Press, 1993. MR 95c:42002 Zbl 0821.42001

Received 19 Feb 2014. Accepted 22 Oct 2014.

TUOMO KUUSI: tuomo.kuusi@aalto.fi
*Institute of Mathematics, Aalto University, P.O. Box 11100, FI-00076 Aalto, Finland*

GIUSEPPE MINGIONE: giuseppe.mingione@math.unipr.it
*Dipartimento di Matematica e Informatica, Universitá di Parma, Parco Area delle Scienze 53/a, Campus, I-43100 Parma, Italy*

YANNICK SIRE: yannick.sire@univ-amu.fr
*Institut de Mathématique de Marseille, CMI-UMR CNRS 7353, Université Aix-Marseille, 9 rue F. Joliot Curie, 13453 Marseille Cedex 13, France*

ANALYSIS AND PDE
Vol. 8, No. 1, 2015
dx.doi.org/10.2140/apde.2015.8.115

# SYMBOL CALCULUS FOR OPERATORS OF LAYER POTENTIAL TYPE ON LIPSCHITZ SURFACES WITH VMO NORMALS, AND RELATED PSEUDODIFFERENTIAL OPERATOR CALCULUS

STEVE HOFMANN, MARIUS MITREA AND MICHAEL E. TAYLOR

We show that operators of layer potential type on surfaces that are locally graphs of Lipschitz functions with gradients in vmo are equal, modulo compacts, to pseudodifferential operators (with rough symbols), for which a symbol calculus is available. We build further on the calculus of operators whose symbols have coefficients in $L^\infty \cap \text{vmo}$, and apply these results to elliptic boundary problems on domains with such boundaries, which in turn we identify with the class of Lipschitz domains with normals in vmo. This work simultaneously extends and refines classical work of Fabes, Jodeit and Rivière, and also work of Lewis, Salvaggi and Sisto, in the context of $\mathscr{C}^1$ surfaces.



Work supported in part by grants from the US National Science Foundation, the Simons Foundation, and the University of Missouri.
*MSC2010:* primary 31B10, 35S05, 35S15, 42B20, 35J57; secondary 42B37, 45B05, 58J05, 58J32.
*Keywords:* singular integral operator, compact operator, pseudodifferential operator, rough symbol, symbol calculus, single layer potential operator, strongly elliptic system, boundary value problem, nontangential maximal function, nontangential boundary trace, Dirichlet problem, regularity problem, Poisson problem, oblique derivative problem, regular elliptic boundary problem, elliptic first-order system, Hodge–Dirac operator, Cauchy integral, Hardy spaces, Sobolev spaces, Besov spaces, Lipschitz domain, SKT domain.

## 1. Introduction

We produce a symbol calculus for a class of operators of layer potential type, of the form

$$Kf(x) = \mathrm{PV}\int_{\partial\Omega} k(x, x-y) f(y)\, d\sigma(y), \quad x \in \partial\Omega, \tag{1.1}$$

in the following setting. First,

$$k \in \mathscr{C}^\infty(\mathbb{R}^{n+1} \times (\mathbb{R}^{n+1} \setminus 0)) \tag{1.2}$$

with $k(x,z)$ homogeneous of degree $-n$ in $z$ and $k(x,-z) = -k(x,z)$. Next, $\Omega \subset \mathbb{R}^{n+1}$ is a bounded Lipschitz domain with a little extra regularity. Namely, $\Omega$ is locally the upper graph of a function $\varphi_0 : \mathbb{R}^n \to \mathbb{R}$ satisfying

$$\nabla\varphi_0 \in L^\infty(\mathbb{R}^n) \cap \mathrm{vmo}(\mathbb{R}^n). \tag{1.3}$$

We say $\Omega$ is a $\mathrm{Lip} \cap \mathrm{vmo}_1$ domain.

Since we will be dealing with a number of variants of BMO, we recall some definitions. First,

$$\mathrm{BMO}(\mathbb{R}^n) := \{f \in L^1_{\mathrm{loc}}(\mathbb{R}^n) : f^\# \in L^\infty(\mathbb{R}^n)\}, \tag{1.4}$$

where

$$f^\#(x) := \sup_{B \in \mathscr{B}(x)} \frac{1}{V(B)} \int_B |f(y) - f_B|\, dy, \tag{1.5}$$

with $\mathscr{B}(x) := \{B_r(x) : 0 < r < \infty\}$, $B_r(x)$ being the ball centered at $x$ of radius $r$, and $f_B$ the mean value of $f$ on $B$. There are variants giving the same space. For example, one could use cubes containing $x$ instead of balls centered at $x$, and one could replace $f_B$ in (1.5) by $c_B$, chosen to minimize the integral. We set

$$\|f\|_{\mathrm{BMO}} := \|f^\#\|_{L^\infty}. \tag{1.6}$$

This is not a norm, since $\|c\|_{\mathrm{BMO}} = 0$ if $c$ is a constant; it is a seminorm. The space $\mathrm{bmo}(\mathbb{R}^n)$ is defined by

$$\mathrm{bmo}(\mathbb{R}^n) := \{f \in L^1_{\mathrm{loc}}(\mathbb{R}^n) : {}^\# f \in L^\infty(\mathbb{R}^n)\}, \tag{1.7}$$

where

$${}^\# f(x) := \sup_{B \in \mathscr{B}_1(x)} \frac{1}{V(B)} \int_B |f(y) - f_B|\, dy + \frac{1}{V(B_1(x))} \int_{B_1(x)} |f(y)|\, dy \tag{1.8}$$

with $\mathcal{B}_1(x) := \{B_r(x) : 0 < r \le 1\}$. We set

$$\|f\|_{\mathrm{bmo}} := \|^{\#}f\|_{L^\infty}. \tag{1.9}$$

This is a norm, and $\mathrm{bmo}(\mathbb{R}^n)$ has good localization properties.

Now, $\mathrm{VMO}(\mathbb{R}^n)$ is the closure in $\mathrm{BMO}(\mathbb{R}^n)$ of $\mathrm{UC}(\mathbb{R}^n) \cap \mathrm{BMO}(\mathbb{R}^n)$, where $\mathrm{UC}(\mathbb{R}^n)$ is the space of uniformly continuous functions on $\mathbb{R}^n$, and $\mathrm{vmo}(\mathbb{R}^n)$ is the closure in $\mathrm{bmo}(\mathbb{R}^n)$ of $\mathrm{UC}(\mathbb{R}^n) \cap \mathrm{bmo}(\mathbb{R}^n)$. One can use local coordinates and partitions of unity to define $\mathrm{bmo}(M)$ and $\mathrm{vmo}(M)$ on a class of Riemannian manifolds $M$ (see [Taylor 2009]). See also Appendix C of this paper for a discussion of $\mathrm{BMO}(M)$ and $\mathrm{VMO}(M)$ on spaces $M$ of homogeneous type. If $M$ is compact, $\mathrm{BMO}(M)$ coincides with $\mathrm{bmo}(M)$ and $\mathrm{VMO}(M)$ coincides with $\mathrm{vmo}(M)$.

With this in mind, $\Omega$ could be an open set in a compact $(n+1)$-dimensional Riemannian manifold $M$, whose boundary, in local coordinates on $M$, is locally a graph as in (1.3), and $k(x, x-y)$ in (1.1) could be the integral kernel of a pseudodifferential operator on $M$ of order $-1$ with odd symbol. In fact, lower-order terms in $k(x, x-y)$ yield weakly singular integral operators on functions on $L^p(\partial\Omega)$, which are compact on $L^p(\partial\Omega)$, for $1 < p < \infty$, on elementary grounds. Thus it suffices for the principal symbol to have this property.

The analysis of operators of the form (1.1) as bounded operators on $L^p(\partial\Omega)$ for $p \in (1, \infty)$, together with nontangential maximal function estimates for

$$\mathcal{K} f(x) = \int_{\partial\Omega} k(x, x-y) f(y)\, d\sigma(y), \quad x \in \mathbb{R}^{n+1} \setminus \partial\Omega, \tag{1.10}$$

and nontangential convergence, was done for general Lipschitz domains in [Coifman et al. 1982], carrying through the breakthrough initiated in [Calderón 1977], at least for $k = k(x-y)$.

Also key was [Fabes et al. 1978], which treated (1.1) (again with $k = k(x-y)$) when $\Omega$ has a $\mathscr{C}^1$ boundary and gave some applications to PDE. These applications involved looking at double layer potentials

$$K_d f(x) = \mathrm{PV}\int_{\partial\Omega} \nu(x)\cdot(x-y)E(x-y) f(y)\, d\sigma(y), \quad x \in \partial\Omega, \tag{1.11}$$

where $\nu(x)$ is the unit normal to $\partial\Omega$ and $E(z) = c_n|z|^{-(n+1)}$. Such an operator is of the form $K_d f(x) = \nu(x)\cdot Kf(x)$, where $K$ is as in (1.1) with $k(z) = zE(z)$ vector-valued. In [Fabes et al. 1978] it was shown that $K_d$ is compact when $\Omega$ is a bounded domain of class $\mathscr{C}^1$. (See Section 3D of this paper for a proof that $K_d$ is compact more generally when $\Omega$ is a bounded $\mathrm{Lip} \cap \mathrm{vmo}_1$ domain.) This compactness was applied to the Dirichlet problem for the Laplace operator on bounded $\mathscr{C}^1$ domains. In fact, if

$$\mathcal{K}_d f(x) = \int_{\partial\Omega} \nu(x)\cdot(x-y)E(x-y) f(y)\, d\sigma(y), \quad x \in \Omega, \tag{1.12}$$

one has

$$\mathcal{K}_d f|_{\partial\Omega} = \left(\tfrac{1}{2}I + K_d\right) f, \tag{1.13}$$

so solving the Dirichlet problem $\Delta u = 0$ on $\Omega$, $u|_{\partial\Omega} = g$, in the form $u = \mathcal{K}_d f$, leads to solving

$$\left(\tfrac{1}{2}I + K_d\right) f = g, \tag{1.14}$$

and the compactness of $K_d$ implies $\frac{1}{2}I + K_d$ is Fredholm of index 0.

For a general bounded Lipschitz domain $\Omega \subset \mathbb{R}^{n+1}$, (1.12)–(1.14) still hold but $K_d$ is typically not compact. However, it was shown in [Verchota 1984] that $\frac{1}{2}I + K_d$ is still Fredholm of index 0, using Rellich identities as a tool. This led to much work on other elliptic boundary problems, including boundary problems for the Stokes system, linear elasticity systems, and the Hodge Laplacian. In [Mitrea and Taylor 1999] a program was initiated that extended the study of (1.1) from $k = k(x - y)$ to $k = k(x, x - y)$, a development that enabled the authors to work on Lipschitz domains in Riemannian manifolds. This led to a series of papers, including [Mitrea and Taylor 2000; Mitrea et al. 2001], in which variants of Rellich identities also played major roles.

Meanwhile, [Hofmann 1994] established compactness of $K_d$ in (1.11) when $\Omega \subset \mathbb{R}^{n+1}$ is a bounded $\mathrm{VMO}_1$ domain, i.e., its boundary is locally a graph of a function $\varphi_0$ satisfying

$$\nabla\varphi_0 \in \mathrm{VMO}(\mathbb{R}^n), \tag{1.15}$$

which is weaker than (1.3). This led [Hofmann et al. 2010] to establish compactness of a somewhat broader class of operators called regular SKT domains, not just $\mathrm{VMO}_1$ domains; this class was introduced by [Semmes 1991; Kenig and Toro 1997], who called them chord–arc domains with vanishing constant. This was applied in [Hofmann et al. 2010] to the Dirichlet boundary problem for the Laplace operator, on regular SKT domains in Riemannian manifolds, and also to a variety of boundary problems for other second-order elliptic systems.

In these works on various elliptic boundary problems, both on Lipschitz domains and on regular SKT domains, each elliptic system seemed to need a separate treatment. This is in striking contrast to the now-standard theory of regular elliptic boundary problems on smoothly bounded domains for operators with smooth coefficients. Such cases yield operators of the form (1.1) that are pseudodifferential operators on $\partial\Omega$, for which a symbol calculus is effective to power the analysis. One can, for example, see the treatment of regular elliptic boundary problems in [Taylor 1996, Chapter 7, §12].

Our goal here is to develop a symbol calculus for operators of the form (1.1) in $\mathrm{Lip} \cap \mathrm{vmo}_1$ domains, and to apply this symbol calculus to the analysis of some elliptic boundary problems.

We work in local graph coordinates, in which (1.1) takes the form

$$Kf(x) = \mathrm{PV}\int_{\mathbb{R}^n} k(\varphi(x), \varphi(x) - \varphi(y))\, f(y)\,\Sigma(y)\, dy, \quad x \in \mathbb{R}^n, \tag{1.16}$$

where $\varphi(x) = (x, \varphi_0(x))$ with $\varphi_0 : \mathbb{R}^n \to \mathbb{R}$ as in (1.3). In fact, we allow $\varphi_0 : \mathbb{R}^n \to \mathbb{R}^\ell$. The surface area element $d\sigma(y)$ equals $\Sigma(y)\, dy$. Our first major result is that, with $K^\#$ given by

$$K^\# f(x) = \mathrm{PV}\int_{\mathbb{R}^n} k(\varphi(x), D\varphi(x)(x - y))\, f(y)\,\Sigma(y)\, dy, \quad x \in \mathbb{R}^n, \tag{1.17}$$

we have

$$K - K^\# \quad \text{compact on } \ L^p(B) \tag{1.18}$$

for $p \in (1, \infty)$, for any ball $B \subset \mathbb{R}^n$. Then, as we show, $K^{\#} f = p(x, D)(\Sigma f)$, with

$$p(x, D) \in \mathrm{OP}(L^\infty \cap \mathrm{vmo}) S^0_{\mathrm{cl}}(\mathbb{R}^n), \tag{1.19}$$

a class of pseudodifferential operators studied in [Taylor 2000] and shown to have a viable symbol calculus. Definitions and basic results are given in Appendix C of this paper. The proof of (1.18), given in Section 2, makes essential use of results in [Hofmann 1994] and further material in [Hofmann et al. 2010].

Since (1.16) and (1.17) are given in local graph coordinates, it is important to record how operators are related when represented in two different such coordinates and how a symbol can be associated to such an operator independently of the coordinate representation. These matters are handled in Section 3.

In connection with this, we mention work of Lewis, Salvaggi and Sisto [Lewis et al. 1993], providing such an analysis on $\mathscr{C}^1$ manifolds. In particular, (1.18) (for $\varphi \in \mathscr{C}^1$) plays a central role there. In that work, the function $k(x, z)$ is required to be analytic in $z \in \mathbb{R}^{n+1} \setminus \{0\}$. The need for such analyticity arises from technical issues, which we can overcome here thanks to the advances in [Hofmann 1994; Hofmann et al. 2010]. One desirable effect of not requiring such analyticity is that our results readily allow for microlocalization. Though we do not pursue microlocal analysis on boundaries of $\mathrm{Lip} \cap \mathrm{vmo}_1$ domains here, we are pleased to advertise the potential to pursue such analysis.

The structure of the rest of this paper is as follows. Section 2 is devoted to a proof of the basic result (1.18). Section 3 builds on this to produce a symbol calculus, making essential use of results on operators of the form (1.19), recalled in Appendix C. Section 4 applies these results to some boundary problems for elliptic systems on $\mathrm{Lip} \cap \mathrm{vmo}_1$ domains. These include the Dirichlet problem for a general class of second-order, strongly elliptic systems and a class of oblique derivative problems. We also produce a general result on regular boundary problems for first-order elliptic systems, and show how this plays out for the Hodge–Dirac operator $d + \delta$ acting on differential forms.

A set of appendices deals with auxiliary results. Appendix A gives material used in Section 2A. Appendix B gives a detailed analysis of just how a principal value integral like (1.1) works for such domains as we consider here. Appendix C reviews material on the class of pseudodifferential operators (1.19). Appendix D reviews matters related to $\mathrm{BMO}(M)$ and $\mathrm{VMO}(M)$ when $M$ is a space of homogeneous type. Appendix E proves that a bounded domain $\Omega \subset \mathbb{R}^{n+1}$ is locally the upper-graph of a function satisfying (1.3) if and only if its outward unit normal belongs to $\mathrm{VMO}(\partial\Omega)$.

## 2. From layer potential operators to pseudodifferential operators

The primary goal of this section is to establish the compactness of the difference between a singular integral operator $K$ of layer potential type as in (1.1) and a related operator $K^{\#}$, which belongs to the class of pseudodifferential operators $\mathrm{OP}(L^\infty \cap \mathrm{vmo}) S^0_{\mathrm{cl}}$, a class that is reviewed in Appendix C. We proceed in stages.

**2A.** ***General local compactness results.*** Below, the principal value integrals $\mathrm{PV}\!\int$ are understood in the sense of removing small balls centered at the singularity and passing to the limit by letting their radii approach zero; for a more flexible view on this topic see the discussion in Appendix B. We begin by recalling the following local compactness result:

**Theorem 2.1.** *Assume $\varphi : \mathbb{R}^n \to \mathbb{R}$ and $\psi : \mathbb{R}^n \to \mathbb{R}^m$ are two locally integrable functions satisfying*

$$\nabla\varphi \in \mathrm{vmo}(\mathbb{R}^n), \quad D\psi \in \mathrm{bmo}(\mathbb{R}^n), \tag{2A.1}$$

*and set*

$$\Gamma(x, y) := \varphi(x) - \varphi(y) - \nabla\varphi(x)(x - y), \quad x,\ y \in \mathbb{R}^n. \tag{2A.2}$$

*Given $F : \mathbb{R}^m \to \mathbb{R}$ smooth (of a sufficiently large order $M = M(m, n) \in \mathbb{N}$), even on $\mathbb{R}^m$ and such that*

$$|F(w)| \leq C(1 + |w|)^{-1} \quad \text{for every } w \in \mathbb{R}^m \tag{2A.3}$$

$$\text{and} \quad \partial^\alpha F \in L^1(\mathbb{R}^m) \quad \text{whenever } |\alpha| \leq M, \tag{2A.4}$$

*consider the principal value integral operator*

$$Tf(x) := \mathrm{PV}\int_{\mathbb{R}^n} |x - y|^{-(n+1)} F\left(\frac{\psi(x) - \psi(y)}{|x - y|}\right)\Gamma(x, y) f(y)\, dy, \quad x \in \mathbb{R}^n, \tag{2A.5}$$

*and the associated maximal operator*

$$T_* f(x) := \sup_{\varepsilon>0}\left|\int_{\substack{y\in\mathbb{R}^n\\ |x-y|>\varepsilon}} |x - y|^{-(n+1)} F\left(\frac{\psi(x) - \psi(y)}{|x - y|}\right)\Gamma(x, y) f(y)\, dy\right|, \quad x \in \mathbb{R}^n. \tag{2A.6}$$

*Then for each $p \in (1, \infty)$ there exists $C_{n,p} \in (0, \infty)$ such that*

$$\|T_* f\|_{L^p(\mathbb{R}^n)} \leq C_{n,p}\left(\sum_{|\alpha|\leq M} \|\partial^\alpha F\|_{L^1(\mathbb{R}^m)} + \sup_{w\in\mathbb{R}^m}\big[(1 + |w|)|F(w)|\big]\right) \times \|\nabla\varphi\|_{\mathrm{BMO}(\mathbb{R}^n)}(1 + \|D\psi\|_{\mathrm{BMO}(\mathbb{R}^n)})^N \|f\|_{L^p(\mathbb{R}^n)} \tag{2A.7}$$

*for every $f \in L^p(\mathbb{R}^n)$. Also, with $B_R$ abbreviating $B(0, R) := \{x \in \mathbb{R}^n : |x| < R\}$, it follows that for each $R \in (0, \infty)$ and $p \in (1, \infty)$ the operator*

$$T : L^p(B_R) \longrightarrow L^p(B_R) \quad \text{is compact.} \tag{2A.8}$$

This result is given in [Hofmann et al. 2010, Theorem 4.34, p. 2725 and Theorem 4.35, p. 2726]. As noted there, the analysis behind it is from [Hofmann 1994]. Of course, there is a natural analogue of Theorem 2.1 when the function $\varphi$ is vector-valued (implied by the scalar case by working componentwise). Here, the goal is to prove the following version of Theorem 2.1:

**Theorem 2.2.** *Suppose $\varphi : \mathbb{R}^n \to \mathbb{R}$ and $\psi : \mathbb{R}^n \to \mathbb{R}^m$ are two locally integrable functions satisfying*

$$\nabla\varphi \in \mathrm{vmo}(\mathbb{R}^n), \quad D\psi \in L^\infty(\mathbb{R}^n), \tag{2A.9}$$

*and let the symbol $\Gamma(x, y)$ retain the same significance as in* (2A.2). *Given an even, real-valued function $F \in \mathscr{C}^M(\mathbb{R}^k)$ (for a sufficiently large $M \in \mathbb{N}$) along with some matrix-valued function*

$$A : \mathbb{R}^n \longrightarrow \mathbb{R}^{k\times m}, \quad A \in L^\infty(\mathbb{R}^n), \tag{2A.10}$$

*consider the principal value singular integral operator*

$$T_A f(x) := \mathrm{PV}\int_{\mathbb{R}^n} |x-y|^{-(n+1)} F\left(A(x)\frac{\psi(x)-\psi(y)}{|x-y|}\right)\Gamma(x,y) f(y)\,dy, \quad x \in \mathbb{R}^n. \tag{2A.11}$$

*Then for each* $R \in (0,\infty)$ *and* $p \in (1,\infty)$ *the operator*

$$T_A : L^p(B_R) \longrightarrow L^p(B_R) \quad \textit{is compact.} \tag{2A.12}$$

Once again, there is a natural analogue of Theorem 2.2 when the function $\varphi$ is vector-valued (implied by the scalar case by working componentwise).

*Proof of Theorem 2.2.* Fix a finite number

$$R_* > \|D\psi\|_{L^\infty(\mathbb{R}^n)} \tag{2A.13}$$

and abbreviate $B_* := \{w \in \mathbb{R}^m : |w| < R_*\}$. Also, select a real-valued function $\chi$ satisfying

$$\chi \in \mathscr{C}^\infty(\mathbb{R}^m), \quad \chi \text{ even in } \mathbb{R}^m, \quad \operatorname{supp}\chi \subseteq B_*, \quad \chi(z) = 1 \quad \text{whenever } |z| \le \|D\psi\|_{L^\infty(\mathbb{R}^n)}. \tag{2A.14}$$

To proceed, let $\{\vartheta_j\}_{j\in\mathbb{N}} \subset L^2(B_*)$ denote an orthonormal basis of $L^2(B_*)$ consisting of real-valued eigenfunctions of the Dirichlet Laplacian in $B_*$ (as discussed in Appendix A). For $x \in \mathbb{R}^n$, we can write in $L^2(B_*)$ and for a.e. $z \in B_*$,

$$F(A(x)z) = \sum_{j\in\mathbb{N}} b_j(x)\vartheta_j(z), \tag{2A.15}$$

where, for each $j \in \mathbb{N}$, we have set

$$b_j(x) := \int_{B_*} F(A(x)z)\vartheta_j(z)\,dz, \quad x \in \mathbb{R}^n. \tag{2A.16}$$

To estimate the $b_j$, fix $j \in \mathbb{N}$, $x \in \mathbb{R}^n$, and observe that for each $N \in \mathbb{N}$ we may write

$$\begin{aligned}
\lambda_j^N |b_j(x)| &= \left|\int_{B_*} F(A(x)z)((-\Delta)^N\vartheta_j)(z)\,dz\right| \\
&= \left|\int_{B_*} (-\Delta_z)^N[F(A(x)z)]\vartheta_j(z)\,dz\right| \\
&\le C_N \|A\|_{L^\infty(\mathbb{R}^n)}^{2N}\Big\{\sup_{\substack{|w|\le R_*\|A\|_{L^\infty(\mathbb{R}^n)} \\ |\alpha|=2N}} |(\partial^\alpha F)(w)|\Big\}\|\vartheta_j\|_{L^\infty(B_*)} \\
&\le C_{A,F,R_*,N}\, j^{1/2+2/n}
\end{aligned} \tag{2A.17}$$

by (A.9). In light of (A.8) this ultimately shows that for each $N \in \mathbb{N}$ there exists a constant $C_N \in (0,\infty)$ such that

$$\|b_j\|_{L^\infty(\mathbb{R}^n)} \le C_N j^{-N} \quad \text{for all } j \in \mathbb{N}. \tag{2A.18}$$

Moving on, we note that combining (2A.15) with its version written for $-z$ in place of $z$, and keeping in mind that $F$ is even, yields

$$F(A(x)z) = \sum_{j\in\mathbb{N}} b_j(x)\widetilde{\vartheta}_j(z), \tag{2A.19}$$

where, for each $j \in \mathbb{N}$, we have set

$$\widetilde{\vartheta}_j(z) := \tfrac{1}{2}(\vartheta_j(z) + \vartheta_j(-z)), \quad z \in B_*. \tag{2A.20}$$

In particular, for each $j \in \mathbb{N}$,

$$\widetilde{\vartheta}_j \in \mathscr{C}^\infty_{loc}(B_*) \text{ is even, vanishes on } \partial B_*, \text{ and satisfies } -\Delta\widetilde{\vartheta}_j = \lambda_j\widetilde{\vartheta}_j \text{ in } B_*. \tag{2A.21}$$

Multiplying both sides of (2A.19) with the cut-off function $\chi$ from (2A.14) then finally yields

$$\chi(z)F(A(x)z) = \sum_{j\in\mathbb{N}} b_j(x)F_j(z), \quad x \in \mathbb{R}^n,\ z \in \mathbb{R}^m, \tag{2A.22}$$

where, for each $j \in \mathbb{N}$, we have set

$$F_j(z) := \chi(z)\widetilde{\vartheta}_j(z), \quad z \in \mathbb{R}^m, \tag{2A.23}$$

naturally viewed as zero outside $B_*$. Hence, for each $j \in \mathbb{N}$,

$$F_j \in \mathscr{C}^\infty(\mathbb{R}^m) \quad \text{is an even function supported in } B_*, \tag{2A.24}$$

and (A.11) implies that for every multi-index $\alpha \in \mathbb{N}_0^m$ there exists a constant $C_{m,\alpha} \in (0,\infty)$ such that

$$\|\partial^\alpha F_j\|_{L^\infty(\mathbb{R}^m)} \le C_{m,\alpha}\, j^{1/2+2/n}. \tag{2A.25}$$

Since

$$z = \frac{\psi(x)-\psi(y)}{|x-y|} \quad\Longrightarrow\quad |z| \le \|D\psi\|_{L^\infty(\mathbb{R}^m)} \quad\Longrightarrow\quad \chi(z) = 1, \tag{2A.26}$$

we deduce from (2A.22) that

$$T_A f(x) = \sum_{j\in\mathbb{N}} b_j(x)T_j f(x), \tag{2A.27}$$

where, for each $j \in \mathbb{N}$, we have set

$$T_j f(x) := \mathrm{PV}\int_{\mathbb{R}^n} |x-y|^{-(n+1)} F_j\Big(\frac{\psi(x)-\psi(y)}{|x-y|}\Big)\Gamma(x,y) f(y)\,dy, \quad x \in \mathbb{R}^n. \tag{2A.28}$$

At this stage, Theorem 2.1 applies to each operator $T_j$. In concert, estimates (2A.7) and (2A.25) yield a polynomial bound in $j \in \mathbb{N}$ on the operator norms of $T_j$ on $L^p(\mathbb{R}^n)$. Then, in the context of the expansion (2A.27), the rapid decrease (2A.18) implies the desired compactness on $L^p(B_R)$ for $T_A$ for each $R \in (0,\infty)$ and $p \in (1,\infty)$. □

It is possible to prove Theorem 2.2 using the Fourier transform in place of spectral methods, based on Dirichlet eigenfunction decompositions. We shall do so below and, in the process, derive further information about the family of truncated operators (indexed by $\varepsilon > 0$)

$$T_{A,\varepsilon} f(x) := \int_{\{y\in\mathbb{R}^n:|x-y|>\varepsilon\}} |x-y|^{-(n+1)} F\left(A(x)\frac{\psi(x)-\psi(y)}{|x-y|}\right)\Gamma(x,y) f(y)\,dy, \tag{2A.29}$$

where $x \in \mathbb{R}^n$, including the pointwise a.e. existence of the associated principal value singular integral operator.

**Theorem 2.3.** *For each $\varepsilon > 0$ let $T_{A,\varepsilon}$ be as in* (2A.29), *where $\Gamma(x, y)$ is defined as in* (2A.2) *for a function $\varphi : \mathbb{R}^n \to \mathbb{R}$ satisfying $\nabla\varphi \in \mathrm{BMO}(\mathbb{R}^n)$, $A \in L^\infty(\mathbb{R}^n)$ is a $k\times m$ matrix-valued function, $\psi : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz, and $F \in \mathscr{C}^M(\mathbb{R}^k)$ is even.*

*Then, if $M = M(m,n) \in \mathbb{N}$ is large enough, there is a positive $M_0 < \infty$ such that, for $1 < p < \infty$,*

$$\sup_{\varepsilon>0} \|T_{A,\varepsilon} f\|_{L^p(\mathbb{R}^n)} \le \Big\|\sup_{\varepsilon>0}|T_{A,\varepsilon} f|\Big\|_{L^p(\mathbb{R}^n)} \le C_0(1+\|\nabla\psi\|_{L^\infty(\mathbb{R}^n)})^{M_0}\|\nabla\varphi\|_{\mathrm{BMO}(\mathbb{R}^n)}\|f\|_{L^p(\mathbb{R}^n)}, \tag{2A.30}$$

*where the constant $C_0$ depends on $\|A\|_\infty$, $p$, $n$, $m$, $k$ and $\|F\|_{\mathscr{C}^M(B(0,\|A\|_\infty R_*))}$ with*

$$R_* := 2(\|\nabla\psi\|_\infty + 1). \tag{2A.31}$$

*Moreover,*

$$\nabla\varphi \in \mathrm{VMO}(\mathbb{R}^n) \implies \lim_{\varepsilon\to 0^+} T_{A,\varepsilon} f(x) \quad \text{exists for a.e. } x \in \mathbb{R}^n \text{ for all } f \in L^p(\mathbb{R}^n). \tag{2A.32}$$

*In fact, a more general result of this nature holds. Specifically, if $B : \mathbb{R}^n \to \mathbb{R}^{m'}$ is a bi-Lipschitz function and if, for each, $\varepsilon > 0$ we set*

$$T_{A,B,\varepsilon} f(x) := \int_{\{y\in\mathbb{R}^n:|B(x)-B(y)|>\varepsilon\}} |x-y|^{-(n+1)} F\left(A(x)\frac{\psi(x)-\psi(y)}{|x-y|}\right)\Gamma(x,y) f(y)\,dy, \tag{2A.33}$$

*where $x \in \mathbb{R}^n$, then*

$$\nabla\varphi \in \mathrm{VMO}(\mathbb{R}^n) \implies \lim_{\varepsilon\to 0^+} T_{A,B,\varepsilon} f(x) \quad \text{exists for a.e. } x \in \mathbb{R}^n \text{ for all } f \in L^p(\mathbb{R}^n). \tag{2A.34}$$

We shall prove estimate (2A.30) by reducing it to the scalar-valued case $k = m = 1$, with $A \equiv 1$, which is Theorem 1.10 in [Hofmann 1994]. Given (2A.30), for $\varphi \in \mathrm{vmo}(\mathbb{R}^n)$ one then gets local compactness (as in the statement of Theorem 2.2: compare (2A.12)) of the associated principal value operator by the usual methods.

*Proof of Theorem 2.3.* For $z \in \mathbb{R}^m$, set $F_x(z) := F(A(x)z)$. Note that, since $A \in L^\infty$, we have that $F_x(\cdot) \in \mathscr{C}^M$ with

$$\sup_{0\le j\le M} \|\nabla^j F_x(\cdot)\|_{L^\infty(B)} \quad \text{controlled uniformly in } x \text{ for every ball } B \subset \mathbb{R}^m. \tag{2A.35}$$

Moreover, as before, we may suppose that

$$F_x(\cdot) \text{ is supported in the ball } B(0, R_*) \subset \mathbb{R}^m \text{ for every } x \in \mathbb{R}^n, \tag{2A.36}$$

where $R_*$ is as in (2A.31). For notational convenience, we normalize $F$ so that

$$\sup_{0\le j\le M}\|\nabla^j F(\cdot)\|_{L^\infty(B(0,\|A\|_\infty R_*))}=1. \tag{2A.37}$$

We may write

$$F_x(z)=c\int_{\mathbb{R}^m}\widehat{F}_x(\xi)\cos(z\cdot\xi)\,d\xi, \tag{2A.38}$$

where $\widehat{F}_x$ is the Fourier transform of $F_x$, and we observe that, by standard estimates for the Fourier transform and our normalization of $F$ from (2A.37),

$$\operatorname*{ess\,sup}_{x\in\mathbb{R}^n}|\widehat{F}_x(\xi)|\le CR_*^m(1+|\xi|)^{-M}. \tag{2A.39}$$

Let $\eta\in\mathscr{C}_0^\infty(-2,2)$ be an even function with $\eta\equiv 1$ on $[-1,1]$ and, for $\xi\in\mathbb{R}^m$, $t\in\mathbb{R}$, set

$$E_\xi(t):=\cos(t)\eta\bigg(\frac{t}{(1+|\xi|)R_*}\bigg). \tag{2A.40}$$

Observe that, for $z\in B(0,R_*)\subset\mathbb{R}^m$, we may replace $\cos(z\cdot\xi)$ by $E_\xi(z\cdot\xi)$ in (2A.38). In concert with (2A.29) and (2A.38), this permits us to write

$$\begin{aligned}T_{A,\varepsilon}f(x)&=\int_{\{y\in\mathbb{R}^n:|x-y|>\varepsilon\}}|x-y|^{-(n+1)}F\bigg(A(x)\frac{\psi(x)-\psi(y)}{|x-y|}\bigg)\Gamma(x,y)f(y)\,dy\\&=c\int_{\mathbb{R}^m}\widehat{F}_x(\xi)\bigg\{\int_{\{y\in\mathbb{R}^n:|x-y|>\varepsilon\}}|x-y|^{-(n+1)}E_\xi\bigg(\xi\cdot\frac{\psi(x)-\psi(y)}{|x-y|}\bigg)\Gamma(x,y)f(y)\,dy\bigg\}\,d\xi\\&=c\int_{\mathbb{R}^m}(1+|\xi|)^{M-N}\widehat{F}_x(\xi)T_{\xi,\varepsilon}f(x)\,d\xi,\end{aligned} \tag{2A.41}$$

where

$$T_{\xi,\varepsilon}f(x):=\int_{\{y\in\mathbb{R}^n:|x-y|>\varepsilon\}}|x-y|^{-(n+1)}\widetilde{E}_\xi\bigg(\xi\cdot\frac{\psi(x)-\psi(y)}{|x-y|}\bigg)\Gamma(x,y)f(y)\,dy \tag{2A.42}$$

and, with $N$ a large number to be chosen later,

$$\widetilde{E}_\xi(t):=(1+|\xi|)^{N-M}E_\xi(t)\quad\text{for all } t\in\mathbb{R}. \tag{2A.43}$$

In turn, from (2A.41) and (2A.39) we deduce that

$$\Big\|\sup_{\varepsilon>0}|T_{A,\varepsilon}f|\Big\|_{L^p(\mathbb{R}^n)}\le CR_*^m\int_{\mathbb{R}^m}(1+|\xi|)^{-N}\Big\|\sup_{\varepsilon>0}|T_{\xi,\varepsilon}f|\Big\|_{L^p(\mathbb{R}^n)}\,d\xi, \tag{2A.44}$$

We now set

$$N:=M-2 \tag{2A.45}$$

and note that this choice ensures that, for all nonnegative integers $j$,

$$\bigg|\Big(\frac{d}{dt}\Big)^j\widetilde{E}_\xi(t)\bigg|\le C_j(1+|\xi|)^{-2}\bigg(\frac{1}{1+|t|/((1+|\xi|)R_*)}\bigg)^2\le C_jR_*^2(1+|t|)^{-2},$$

where the constant $C_j$ may depend on $j$ but is independent of $\xi$. By [Hofmann 1994, Theorem 1.10, p. 470] applied to the scalar-valued Lipschitz function $\xi \cdot \psi$, we then have that, for some $M_1 < \infty$,

$$\Big\| \sup_{\varepsilon>0} |T_{\xi,\varepsilon} f(x)| \Big\|_{L^p_x(\mathbb{R}^n)} \le C R_*^2 (1+|\xi| R_*)^{M_1} \|\nabla\varphi\|_{\mathrm{BMO}(\mathbb{R}^n)} \|f\|_{L^p(\mathbb{R}^n)}. \tag{2A.46}$$

Plugging the latter estimate into (2A.44) and finally choosing

$$M := M_1 + m + 3, \tag{2A.47}$$

we obtain (2A.30) thanks to (2A.45).

Finally, it remains to consider the issue of the existence of the limits in (2A.32) and (2A.34). We treat in detail the former, since the argument for the latter is similar, granted our results in Appendix B. To justify (2A.32), make the standing assumption that

$$\nabla\varphi \in \mathrm{VMO}(\mathbb{R}^n) \tag{2A.48}$$

and recall from (2A.41), (2A.45) that

$$T_{A,\varepsilon} f(x) = c \int_{\mathbb{R}^m} (1+|\xi|)^2 \widehat{F}_x(\xi) T_{\xi,\varepsilon} f(x)\, d\xi, \tag{2A.49}$$

where $T_{\xi,\varepsilon} f(x)$ is as in (2A.42). To proceed, observe that, for each $f \in L^p(\mathbb{R}^n)$,

$$\sup_{\varepsilon>0} \Big| (1+|\xi|)^2 \widehat{F}_x(\xi) T_{\xi,\varepsilon} f(x) \Big| \in L^1_\xi(\mathbb{R}^m) \quad \text{for a.e. fixed } x \in \mathbb{R}^n. \tag{2A.50}$$

To see that this is the case, use Minkowski's inequality along with (2A.39) and (2A.46) to estimate

$$\begin{aligned}
&\left\{ \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^m} \sup_{\varepsilon>0} |(1+|\xi|)^2 \widehat{F}_x(\xi) T_{\xi,\varepsilon} f(x)|\, d\xi \right)^p dx \right\}^{\frac{1}{p}} \\
&\qquad \le \int_{\mathbb{R}^m} \Big\| \sup_{\varepsilon>0} |(1+|\xi|)^2 \widehat{F}_x(\xi) T_{\xi,\varepsilon} f(x)| \Big\|_{L^p_x(\mathbb{R}^n)} d\xi \\
&\qquad \le \int_{\mathbb{R}^m} (1+|\xi|)^2 [\operatorname*{ess\,sup}_{x\in\mathbb{R}^n} |\widehat{F}_x(\xi)|] \Big\| \sup_{\varepsilon>0} |T_{\xi,\varepsilon} f(x)| \Big\|_{L^p_x(\mathbb{R}^n)} d\xi \\
&\qquad \le C R_*^{m+2} \|\nabla\varphi\|_{\mathrm{BMO}(\mathbb{R}^n)} \|f\|_{L^p(\mathbb{R}^n)} \int_{\mathbb{R}^m} (1+|\xi|)^{2-M} (1+|\xi| R_*)^{M_1}\, d\xi < +\infty,
\end{aligned} \tag{2A.51}$$

thanks to (2A.47). With (2A.51) in hand, the claim in (2A.50) readily follows. Next, granted (2A.48), we claim that for each fixed function $f \in L^p(\mathbb{R}^n)$ the following holds:

$$\text{for each fixed } \xi \in \mathbb{R}^m, \quad \lim_{\varepsilon\to 0^+} T_{\xi,\varepsilon} f(x) \quad \text{exists for a.e. } x \in \mathbb{R}^n. \tag{2A.52}$$

Given that we have already established (2A.30), this may be justified along the lines of the proof of Theorem 5.11, pp. 500–501 in [Hofmann 1994], based on Proposition B.2 and keeping in mind that VMO functions may be approximated in the BMO norm by continuous functions with compact support, which, in turn, are uniformly approximable by functions in $\mathscr{C}^\infty_0$.

In concert with the uniform integrability property (2A.50), the existence of the limit in (2A.52) makes it possible to use Lebesgue's dominated convergence theorem in order to write that, for a.e. $x \in \mathbb{R}^n$,

$$\begin{aligned}\lim_{\varepsilon\to 0^+} T_{A,\varepsilon} f(x) &= c \lim_{\varepsilon\to 0^+} \int_{\mathbb{R}^m} (1+|\xi|)^2 \widehat{F}_x(\xi) T_{\xi,\varepsilon} f(x)\, d\xi \\ &= c \int_{\mathbb{R}^m} (1+|\xi|)^2 \widehat{F}_x(\xi) \lim_{\varepsilon\to 0^+} T_{\xi,\varepsilon} f(x)\, d\xi. \end{aligned} \tag{2A.53}$$

This proves the claim in (2A.32) and finishes the proof of the theorem. □

**2B.** ***The local compactness of the remainder.*** Let $\varphi : \mathbb{R}^n \to \mathbb{R}^{n+\ell}$ be a Lipschitz map of "graph" type, i.e., assume that

$$\varphi(x) = (x, \varphi_0(x)) \quad \text{for all } x \in \mathbb{R}^n, \tag{2B.1}$$

for some

$$\varphi_0 : \mathbb{R}^n \longrightarrow \mathbb{R}^\ell \quad \text{Lipschitz.} \tag{2B.2}$$

Note that this implies

$$|\varphi(x) - \varphi(y)| \geq |x-y| \quad \text{for all } x,\ y \in \mathbb{R}^n. \tag{2B.3}$$

Let

$$\begin{gathered} k : \mathbb{R}^{n+\ell} \setminus \{0\} \to \mathbb{R} \text{ be a smooth function, positive, homogeneous of degree } -n \\ \text{and satisfying } k(-w) = -k(w) \text{ for all } w \in \mathbb{R}^{n+\ell} \setminus \{0\}. \end{gathered} \tag{2B.4}$$

Then

$$\begin{aligned} Kf(x) &:= \mathrm{PV}\int_{\mathbb{R}^n} k(\varphi(x)-\varphi(y)) f(y)\, dy \\ &= \mathrm{PV}\int_{\mathbb{R}^n} |x-y|^{-n} k\Big(\frac{\varphi(x)-\varphi(y)}{|x-y|}\Big) f(y)\, dy, \quad x \in \mathbb{R}^n, \end{aligned} \tag{2B.5}$$

defines a bounded operator on $L^p(\mathbb{R}^n)$ for each $p \in (1,\infty)$. We aim to establish a finer structure when $\varphi \in \mathscr{C}^1(\mathbb{R}^n)$ or, more generally, when the Jacobian $D\varphi$ of $\varphi$ satisfies

$$D\varphi \in L^\infty(\mathbb{R}^n) \cap \mathrm{vmo}(\mathbb{R}^n). \tag{2B.6}$$

Namely, we set

$$R := K - K_0, \tag{2B.7}$$

with

$$K_0 f(x) := \mathrm{PV}\int_{\mathbb{R}^n} k(D\varphi(x)(x-y)) f(y)\, dy, \quad x \in \mathbb{R}^n. \tag{2B.8}$$

Note that (2B.3) implies $|D\varphi(x)z| \geq |z|$ for all $z \in \mathbb{R}^n$. We have

$$\begin{aligned} \varphi \in \mathscr{C}^1(\mathbb{R}^n) \quad &\Longrightarrow \quad K_0 \in \mathrm{OP}\mathscr{C}^0 S^0_{\mathrm{cl}}, \\ D\varphi \in L^\infty(\mathbb{R}^n) \cap \mathrm{vmo}(\mathbb{R}^n) \quad &\Longrightarrow \quad K_0 \in \mathrm{OP}(L^\infty \cap \mathrm{vmo}) S^0_{\mathrm{cl}}. \end{aligned} \tag{2B.9}$$

The latter class is studied in [Taylor 2000, Chapter 1, §11] and, for the reader's convenience, useful background material on this topic is presented in Appendix C. See Theorem 2.6 for a derivation of the second part of (2B.9) in a more general setting. As for the “remainder” $R$ in (2B.7), we have

$$Rf(x) = \mathrm{PV}\int_{\mathbb{R}^n} r(x,y) f(y)\,dy, \quad x \in \mathbb{R}^n, \tag{2B.10}$$

where

$$r(x,y) := k(\varphi(x)-\varphi(y)) - k(D\varphi(x)(x-y)) = \int_0^1 r_\tau(x,y)\,d\tau, \tag{2B.11}$$

with

$$\begin{aligned} r_\tau(x,y) &:= (\nabla k)(\varphi(x)-\varphi(y)+\tau\Gamma(x,y))\cdot\Gamma(x,y), \\ \Gamma(x,y) &:= \varphi(x)-\varphi(y)-D\varphi(x)(x-y). \end{aligned} \tag{2B.12}$$

The following is our first major result:

**Theorem 2.4.** *Let $\varphi$ be as in* (2B.1)–(2B.2), *suppose $k$ is as in* (2B.4) *and define $R$ as in* (2B.7), *where $K$, $K_0$ are as in* (2B.5) *and* (2B.8), *respectively. Finally, assume that* (2B.6) *holds. Then, for each ball $B \subset \mathbb{R}^n$ and $p \in (1,\infty)$, the operator*

$$R : L^p(B) \longrightarrow L^p(B) \quad \textit{is compact.} \tag{2B.13}$$

In the case when $\varphi \in \mathscr{C}^1(\mathbb{R}^n)$ and $D\varphi$ has a modulus of continuity satisfying a Dini condition, the compactness result (2B.13) is straightforward. See [Taylor 2000, Chapter 3, §4].

*Proof of Theorem 2.4.* Note that

$$R = \int_0^1 R_\tau\,d\tau, \tag{2B.14}$$

interpreted as a Bochner integral, with

$$R_\tau f(x) := \mathrm{PV}\int_{\mathbb{R}^n} r_\tau(x,y) f(y)\,dy, \quad x \in \mathbb{R}^n, \tag{2B.15}$$

and the integral kernel $r_\tau(x,y)$ as in (2B.12). Given this, and bearing in mind that the collection of compact operators on $L^p(B)$ is a closed linear subspace of $\mathscr{L}(L^p(B), L^p(B))$, it suffices to show that each operator $R_\tau$ has the compactness property (2B.13).

With this goal in mind, for each $\tau \in [0,1]$ observe that the operator $R_\tau$ has the form

$$R_\tau f(x) = \mathrm{PV}\int_{\mathbb{R}^n} |x-y|^{-(n+1)} F\left(\frac{D\varphi(x)(x-y)+\tau\Gamma(x,y)}{|x-y|}\right)\Gamma(x,y) f(y)\,dy \tag{2B.16}$$

with $\Gamma(x,y)$ as in (2B.12) and $F := \nabla k$. Note that the argument of $F$ in (2B.23) is

$$D\varphi(x)(x-y)+\tau\Gamma(x,y) = (x-y, D\varphi_0(x)(x-y)+\tau\Gamma_0(x,y)), \tag{2B.17}$$

with $\varphi_0$ as in (2B.1)–(2B.2) and $\Gamma_0(x,y)$ as in (2B.12), but with $\varphi$ replaced by $\varphi_0$. In particular, there exists a constant $C \in (1,\infty)$ such that

$$1 \le \frac{|D\varphi(x)(x-y)+\tau\Gamma(x,y)|}{|x-y|} \le C \tag{2B.18}$$

for all $x,\ y \in \mathbb{R}^n$ and all $\tau \in [0,1]$. As such, we can alter the function $F(w)$ at will off the set $\{w \in \mathbb{R}^{n+\ell} : 1 \le |w| \le C\}$ and arrange that

$$F \in \mathscr{C}_0^\infty(\mathbb{R}^{n+\ell}) \tag{2B.19}$$

while keeping $F$ even.

Moving on, observe that another way of looking at the argument of $F$ in (2B.23) is to write

$$\begin{aligned} D\varphi(x)(x-y)+\tau\Gamma(x,y) &= \tau(\varphi(x)-\varphi(y))+(1-\tau)D\varphi(x)(x-y) \\ &= [\tau\varphi(x)+(1-\tau)D\varphi(x)x]-[\tau\varphi(y)+(1-\tau)D\varphi(x)y] \\ &= A_\tau(x)(\psi(x)-\psi(y)), \end{aligned} \tag{2B.20}$$

with

$$A_\tau(x) := \begin{pmatrix} \tau I & (1-\tau)D\varphi(x) \end{pmatrix} \tag{2B.21}$$

and

$$\psi(x) := \begin{pmatrix} \varphi(x) \\ x \end{pmatrix}, \quad \psi : \mathbb{R}^n \longrightarrow \mathbb{R}^{2n+\ell}. \tag{2B.22}$$

The bottom line is that for each $\tau \in [0,1]$ we have

$$R_\tau f(x) = \mathrm{PV}\int_{\mathbb{R}^n} |x-y|^{-(n+1)} F\left(A_\tau(x)\frac{\psi(x)-\psi(y)}{|x-y|}\right)\Gamma(x,y)f(y)\,dy, \quad x \in \mathbb{R}^n, \tag{2B.23}$$

where $A_\tau$, $\psi$ are as in (2B.21)–(2B.22) and we can assume $F$ is even and satisfies (2B.19). Granted this, Theorem 2.2 applies and yields that each $R_\tau$ has the compactness property (2B.13). □

**2C.** ***A variable coefficient version of the local compactness theorem.*** Here the goal is to work out a variable coefficient version of Theorem 2.4 by treating the following class of operators. Let $k$ be in $\mathscr{C}^\infty(\mathbb{R}^{n+\ell} \times (\mathbb{R}^{n+\ell} \setminus 0))$. Suppose $k(w,z)$ is odd in $z$ and homogeneous of degree $-n$ in $z$. In addition, assume bounds

$$|D_w^\alpha D_z^\beta k(w,z)| \le C_{\alpha\beta}|z|^{-n-|\beta|}. \tag{2C.1}$$

We take $\varphi : \mathbb{R}^n \to \mathbb{R}^{n+\ell}$ as in (2B.1)–(2B.2), (2B.6), and consider

$$Kf(x) := \mathrm{PV}\int_{\mathbb{R}^n} k(\varphi(x), \varphi(x)-\varphi(y))\,f(y)\,dy, \quad x \in \mathbb{R}^n. \tag{2C.2}$$

To analyze this type of singular integral operator with variable coefficient kernel, it is convenient to expand

$$k(w,z) = \sum_j a_j(w)\Omega_{n,j}(z), \tag{2C.3}$$

where, starting with orthonormal, real-valued, spherical harmonics $\Omega_j$ on $S^{n-1}$, we have set

$$\Omega_{n,j}(z) := \Omega_j\left(\frac{z}{|z|}\right)|z|^{-n}, \quad z \in \mathbb{R}^n \setminus \{0\}, \tag{2C.4}$$

and where the coefficient functions $a_j$ are given by

$$a_j(w) := \int_{S^{n-1}} k(w,z)\Omega_j(z)\,dz. \tag{2C.5}$$

We can arrange that all the functions $\Omega_{n,j}(z)$ in (2C.3) are odd. There is a polynomial bound in $j$ on the $\mathscr{C}^m$ norm of $\Omega_{n,j}|_{S^{n-1}}$ for each $m \in \mathbb{N}$, and the coefficients $a_j$ are rapidly decreasing in $\mathscr{C}^m$ norm for each $m \in \mathbb{N}$. We have

$$K = \sum_j K_j, \tag{2C.6}$$

where, for each $j$,

$$K_j f(x) := a_j(\varphi(x))\,\mathrm{PV}\int_{\mathbb{R}^n} \Omega_{n,j}(\varphi(x)-\varphi(y))\,f(y)\,dy, \quad x \in \mathbb{R}^n. \tag{2C.7}$$

The series (2C.6) converges rapidly in $L^p$-operator norm for each $p \in (1,\infty)$.

Let us compare $K$ with $K^\#$, defined as

$$K^\# f(x) := \mathrm{PV}\int_{\mathbb{R}^n} k(\varphi(x), D\varphi(x)(x-y))\,f(y)\,dy, \quad x \in \mathbb{R}^n. \tag{2C.8}$$

This time (2C.3) yields

$$K^\# = \sum_j K_j^\#, \tag{2C.9}$$

with $K_j^\#$ given by

$$K_j^\# f(x) := a_j(\varphi(x))\,\mathrm{PV}\int_{\mathbb{R}^n} \Omega_{n,j}(D\varphi(x)(x-y))\,f(y)\,dy, \quad x \in \mathbb{R}^n. \tag{2C.10}$$

We claim that the series (2C.9) is rapidly convergent in $L^p$-operator norm for each $p \in (1,\infty)$. Indeed, Theorem 2.4 directly implies that, for each $j$,

$$K_j - K_j^\# \quad \text{is compact on } L^p(B) \tag{2C.11}$$

for each ball $B \subset \mathbb{R}^n$ and each $p \in (1,\infty)$. The operator norm convergence of (2C.6) and (2C.9) then yield the following variable coefficient counterpart to Theorem 2.4:

**Theorem 2.5.** *Given $K$ as in* (2C.3) *and $K^\#$ as in* (2C.8),

$$K - K^\# \quad \textit{is compact on } L^p(B). \tag{2C.12}$$

Moving on, we propose to further analyze (2C.8) and show that (again, see the discussion in Appendix C for relevant definitions)

$$K^\# \in \mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}. \tag{2C.13}$$

To this end, it is convenient to write

$$k(w, Az) = \sum_j b_j(w, A)\Omega_{n,j}(z) \tag{2C.14}$$

for $A : \mathbb{R}^n \to \mathbb{R}^{n+\ell}$ of the form

$$A = \begin{pmatrix} I \\ A_0 \end{pmatrix}, \tag{2C.15}$$

with

$$b_j(w, A) := \int_{S^{n-1}} k(w, Az)\Omega_j(z)\,d\sigma(z). \tag{2C.16}$$

Again, we can arrange that only odd functions $\Omega_{n,j}$ arise in (2C.14). As $A_0$ varies over a compact subset of $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^\ell)$, the space of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^\ell$, we have uniform rapid decay of $b_j(w, A)$ and each of its derivatives. We have the following conclusion:

**Theorem 2.6.** *The operator $K^\#$ defined by* (2C.8) *satisfies*

$$K^\# f(x) = \sum_j b_j(\varphi(x), D\varphi(x))\, \mathrm{PV}\int_{\mathbb{R}^n} \Omega_{n,j}(x-y) f(y)\, dy, \quad x \in \mathbb{R}^n; \tag{2C.17}$$

*hence*

$$K^\# f(x) = p(x, D) f(x), \quad x \in \mathbb{R}^n \tag{2C.18}$$

*with*

$$p(x,\xi) := \sum_j b_j(\varphi(x), D\varphi(x)) \widehat{\Omega}_{n,j}(\xi). \tag{2C.19}$$

*Consequently*,

$$p \in (L^\infty \cap \mathrm{vmo}) S^0_{\mathrm{cl}} \tag{2C.20}$$

*and* (2C.13) *follows.*

## 3. Symbol calculus

Our goals here are to associate symbols to the operators studied in Section 2 and to examine how these operators behave under coordinate changes.

**3A.** ***Principal symbols.*** Let $\Omega \subset \mathbb{R}^{n+1}$ be a bounded $\mathrm{Lip} \cap \mathrm{vmo}_1$ domain, so $\partial\Omega$ is locally a graph of the form (2B.1)–(2B.2), (2B.6) with $\ell = 1$. Let $\partial^*\Omega$ denote the subset of $\partial\Omega$ of the form $\varphi(x)$ such that $x$ is an $L^p$-Lebesgue point of $D\varphi$ with $p > n$ (so in particular $\varphi$ is differentiable at $x$). Then we set

$$T_{\varphi(x)}\partial^*\Omega := \{D\varphi(x)v : v \in \mathbb{R}^n\} \quad \text{whenever } \varphi(x) \in \partial^*\Omega. \tag{3A.1}$$

In this fashion, we can talk about the tangent bundle and cotangent bundle over $\partial^*\Omega$,

$$T\partial^*\Omega \quad \text{and} \quad T^*\partial^*\Omega, \tag{3A.2}$$

where, in the latter case, the fiber $T^*_{\varphi(x)}\partial^*\Omega$ is the dual space to (3A.1).

Let $k(w, z)$ be smooth on $\mathbb{R}^{n+1} \times (\mathbb{R}^{n+1} \setminus 0)$, odd in $z$ and homogeneous of degree $-n$ in $z$. Consider

$$Kf(x) := \mathrm{PV}\int_{\partial\Omega} k(x, x-y) f(y)\, d\sigma(y), \quad K : L^p(\partial\Omega) \to L^p(\partial\Omega), \quad p \in (1,\infty). \tag{3A.3}$$

In the local coordinate system described above,

$$Kf(x) = \mathrm{PV}\int_{\mathbb{O}} k(\varphi(x), \varphi(x) - \varphi(y)) f(y) \Sigma(y)\, dy \tag{3A.4}$$

with $\mathbb{O} \subset \mathbb{R}^n$ and $d\sigma(y) = \Sigma(y)\, dy$. Note that $\Sigma \in L^\infty \cap \mathrm{vmo}$. As we have seen in Section 2C,

$$K = p(x, D) \quad \text{mod compact} \tag{3A.5}$$

with $p(x,\xi) \in (L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$ odd and homogeneous of degree 0 in $\xi$. We want to associate to $K$ a principal symbol $\sigma_K$ defined on $T^*\partial^*\Omega$. We propose

$$\sigma_K(\varphi(x),\xi) := p(x, D\varphi(x)^{\mathrm{T}}\xi) \tag{3A.6}$$

for $x \in \mathbb{O}$, $\varphi(x) \in \partial^*\Omega$, with $p$ as in (3A.5). If $\partial\Omega$ is smooth, this coincides with the classical transformation formula for the symbol of a pseudodifferential operator. Now $K = K^\#$ mod compact, with $K^\#$ given by (2C.8) with a factor of $\Sigma(y)$ thrown in. This factor can be changed to $\Sigma(x)$ mod compact, so we can take

$$p(x,D)f(x) = \mathrm{PV}\int k(\varphi(x), D\varphi(x)(x-y))\Sigma(x) f(y)\,dy. \tag{3A.7}$$

The standard formula connecting a pseudodifferential operator and its symbol yields

$$p(x,\zeta) = \int_{\mathbb{R}^n} k(\varphi(x), D\varphi(x)z)e^{-iz\cdot\zeta}\Sigma(x)\,dz, \tag{3A.8}$$

so (compare (3B.22)–(3B.23))

$$\begin{aligned} p(x, D\varphi(x)^{\mathrm{T}}\xi) &= \int_{\mathbb{R}^n} k(\varphi(x), D\varphi(x)z)e^{-iD\varphi(x)z\cdot\xi}\Sigma(x)\,dz \\ &= \int_{T_{\varphi(x)}\partial^*\Omega} k(\varphi(x), z^0)e^{-iz^0\cdot\xi}\,dz^0, \end{aligned} \tag{3A.9}$$

since the area element of $\partial\Omega$ at $w \in \partial^*\Omega$ coincides with that of $T_w\partial^*\Omega$. Hence,

$$\sigma_K(w,\xi) = \int_{T_w\partial^*\Omega} k(w,z^0)e^{-iz^0\cdot\xi}\,dz^0, \quad w \in \partial^*\Omega. \tag{3A.10}$$

This last formula is independent of the choice of local coordinates on $\partial\Omega$. If $\partial\Omega$ is smooth, (3A.10) is the standard formula. We note that $T^*_w\partial^*\Omega$ inherits an inner product, and hence a volume form, as a linear subspace of $\mathbb{R}^{n+1}$, and $dz^0 = \Sigma(x)\,dz$ when $w = \varphi(x)$.

Suppose $K$ is an $\ell \times \ell$ system of singular integral operators. We say $K$ is *elliptic on* $\partial\Omega$ if there exists a constant $C > 0$ such that

$$\|\sigma_K(w,\xi)v\| \geq C\|v\| \quad \text{for all } v \in \mathbb{C}^\ell \text{ and } \sigma\text{-a.e. } w \in \partial^*\Omega. \tag{3A.11}$$

In such a case, by (3A.6), the operator $p(x,D) \in \mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$ associated to $K$ in a local graph coordinate system is elliptic, i.e., its symbol $p(x,\xi)$ satisfies the analogue of (3A.11). We can thus prove the following:

**Theorem 3.1.** *Let $\Omega \subset \mathbb{R}^{n+1}$ be a bounded* $\mathrm{Lip} \cap \mathrm{vmo}_1$ *domain. If $K$ is an $\ell \times \ell$ elliptic system of singular integral operators of the form* (3A.3) *and satisfies the ellipticity condition* (3A.11), *then*

$$K : L^p(\partial\Omega) \longrightarrow L^p(\partial\Omega) \quad \textit{is Fredholm for all } p \in (1,\infty). \tag{3A.12}$$

*Moreover, the index of $K$ in* (3A.12) *is independent of $p \in (1,\infty)$, and we have the regularity result*

$$1 < p < q < \infty \quad \textit{and} \quad f \in L^p(\partial\Omega),\ Kf \in L^q(\partial\Omega) \quad \Longrightarrow \quad f \in L^q(\partial\Omega). \tag{3A.13}$$

*Proof.* Let $\{\mathcal{O}_j\}_j$ be an open cover of $\partial\Omega$ on which we have graph coordinates. (We also identify each $\mathcal{O}_j$ with an open subset of $\mathbb{R}^n$.) Let $\{\psi_j\}_j$ be a Lipschitz partition of unity on $\partial\Omega$ subordinate to this cover. Let $\varphi_j \in \mathrm{Lip}(\mathcal{O}_j)$ have compact support and satisfy $\varphi_j \equiv 1$ on a neighborhood of $\operatorname{supp}\psi_j$. Then

$$K = \sum_j KM_{\psi_j} = \sum_j M_{\varphi_j} KM_{\psi_j} \quad \text{mod compacts,} \tag{3A.14}$$

where, generally speaking, $M_\psi f := \psi f$. Now we have (see (3A.5))

$$M_{\varphi_j} KM_{\psi_j} = M_{\varphi_j} p_j(x, D) M_{\psi_j} \quad \text{mod compacts,} \tag{3A.15}$$

with $p_j(x, D) \in \mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$ elliptic. We have a parametrix $e_j(x, D) \in \mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$, satisfying

$$M_{\varphi_i} e_i(x, D) M_{\psi_i} M_{\varphi_j} KM_{\psi_j} = M_{\psi_i\psi_j} \quad \text{mod compacts.} \tag{3A.16}$$

Set

$$E := \sum_i M_{\varphi_i} e_i(x, D) M_{\psi_i}. \tag{3A.17}$$

Then

$$\begin{aligned} EK &= \sum_{i,j} M_{\varphi_i} e_i(x, D) M_{\psi_i} M_{\varphi_j} KM_{\psi_j} \quad \text{mod compacts} \\ &= \sum_{i,j} M_{\psi_i\psi_j} \quad \text{mod compacts} \\ &= I \quad \text{mod compacts.} \end{aligned} \tag{3A.18}$$

Similarly, $E$ is a right Fredholm inverse of $K$, and we have (3A.12).

Going further, for each $p \in (1, \infty)$ let $\iota_p(K)$ denote the index of $K$ on $L^p(\partial\Omega)$. Then, if $1 < p < q < \infty$ and $\mathcal{N}_p$ denotes the null space of $K$ on $L^p(\partial\Omega)$, and $\mathcal{N}'_p$ that of $K^*$ on $L^{p'}(\partial\Omega)$, we have

$$\mathcal{N}_q \subset \mathcal{N}_p, \quad \mathcal{N}'_p \subset \mathcal{N}'_q, \quad \text{hence} \quad \iota_p(K) \geq \iota_q(K). \tag{3A.19}$$

The same type of argument applies to $E$, yielding $\iota_p(E) \geq \iota_q(E)$, hence

$$\iota_p(K) = \iota_q(K), \tag{3A.20}$$

as wanted. Note that, together with (3A.19), this actually forces

$$\mathcal{N}_q = \mathcal{N}_p \quad \text{and} \quad \mathcal{N}'_p = \mathcal{N}'_q. \tag{3A.21}$$

Finally, for (3A.13), if $f \in L^p(\partial\Omega)$ and $Kf = g \in L^q(\partial\Omega)$, then $g$ annihilates $\mathcal{N}'_p$. Since $\mathcal{N}'_q = \mathcal{N}'_p$, $g$ annihilates $\mathcal{N}'_q$, so $g = K\tilde{f}$ for some $\tilde{f} \in L^q(\partial\Omega)$. Given $p < q$, we have $f - \tilde{f} \in \mathcal{N}_p$. Hence $f - \tilde{f} \in \mathcal{N}_q$, and thus $f \in L^q(\partial\Omega)$, as asserted in (3A.13). □

**3B.** ***Transformations of operators under coordinate changes.*** Let $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ be a bi-Lipschitz map, so there exist $a, b \in (0, \infty)$ such that

$$a|x-y| \le |\varphi(x)-\varphi(y)| \le b|x-y| \quad \text{for all } x,\ y \in \mathbb{R}^n. \tag{3B.1}$$

In addition, we assume

$$D\varphi \in \text{vmo}(\mathbb{R}^n). \tag{3B.2}$$

Given

$$k \in \mathscr{C}^\infty(\mathbb{R}^n \setminus 0) \quad \text{homogeneous of degree } -n, \quad k(-z) = -k(z), \tag{3B.3}$$

we set

$$Kf(x) := \text{PV}\int_{\mathbb{R}^n} k(x-y)\, f(y)\, dy, \quad x \in \mathbb{R}^n. \tag{3B.4}$$

Let us also set

$$K_\varphi f(x) := \text{PV}\int_{\mathbb{R}^n} k(\varphi(x)-\varphi(y))\, f(y)\, dy, \quad x \in \mathbb{R}^n. \tag{3B.5}$$

As in the past, we let $M_\chi$ denote the operator of pointwise multiplication by $\chi$.

**Definition 3.2.** Say that $\varphi$ is in $\mathfrak{T}(\mathbb{R}^n)$ provided that (3B.1)–(3B.2) hold and, in addition, whenever (3B.3) holds, the singular integral operator $K_\varphi$ associated with $\varphi$ as in (3B.5) may be decomposed as

$$K_\varphi f(x) = \text{PV}\int_{\mathbb{R}^n} k(D\varphi(x)(x-y))\, f(y)\, dy + R_\varphi f(x), \quad x \in \mathbb{R}^n, \tag{3B.6}$$

for a remainder with the property that for each cut-off function $\chi \in \mathscr{C}_0^\infty(\mathbb{R}^n)$ one has

$$M_\chi R_\varphi M_\chi : L^p(\mathbb{R}^n) \longrightarrow L^p(\mathbb{R}^n) \quad \text{compact for all } p \in (1, \infty). \tag{3B.7}$$

By Theorem 2.6, the principal value integral on the right-hand side of (3B.6) defines an operator

$$\widetilde{K}_\varphi \in \text{OP}(L^\infty \cap \text{vmo}) S^0_{\text{cl}}, \tag{3B.8}$$

which is bounded on $L^p(\mathbb{R}^n)$ for each $p \in (1, \infty)$.

The following is a variant of Theorem 2.4, proven by the same sort of arguments.

**Theorem 3.3.** *Assume $\varphi$ satisfies* (3B.1)–(3B.2). *Assume also that there exists $\kappa > 0$ such that, for all $\tau \in [0, 1]$,*

$$|\tau[\varphi(x)-\varphi(y)] + (1-\tau) D\varphi(x)(x-y)| \ge \kappa|x-y| \quad \textit{for all } x,\ y \in \mathbb{R}^n. \tag{3B.9}$$

*Then $\varphi \in \mathfrak{T}(\mathbb{R}^n)$.*

*In fact, given a function $\chi \in \mathscr{C}_0^\infty(\mathbb{R}^n)$, one has* (3B.7) *provided the estimate in* (3B.9) *holds for all points $x, y \in \operatorname{supp} \chi$.*

Note the similarity of (3B.9) and (2B.18). In this connection, if $\Sigma \subset \mathbb{R}^{n+\ell}$ is an $n$-dimensional graph over $\mathbb{R}^n$, as introduced in Section 2B, and if it is also represented as a graph over a nearby $n$-dimensional linear space $V$, then one gets a bi-Lipschitz map from $\mathbb{R}^n$ to $V \equiv \mathbb{R}^n$ satisfying (3B.9). In such a way, one can represent $\Sigma$ as a $\text{Lip} \cap \text{vmo}_1$ *manifold*, whose transition maps satisfy the conditions of Theorem 3.3. See the next section for more on this.

We proceed to a variable coefficient version of (3B.3)–(3B.7). Take $k$ measurable on $\mathbb{R}^n \times \mathbb{R}^n$, satisfying

$$k(x,z) \quad \text{homogeneous of degree } -n \text{ in } z, \quad k(x,-z) = -k(x,z). \tag{3B.10}$$

Assume $k(x,z)$ is smooth in $z \neq 0$ and that for each multiindex $\alpha$ there exists a finite constant $C_\alpha > 0$ such that

$$\|\partial_z^\alpha k(\cdot\,,z)\|_{L^\infty \cap \mathrm{vmo}} \le C_\alpha |z|^{-n-|\alpha|}, \tag{3B.11}$$

where, for $f \in L^\infty(\mathbb{R}^n)$,

$$\|f\|_{L^\infty\cap\mathrm{vmo}} := \begin{cases} \|f\|_{L^\infty} & \text{if } f \in \mathrm{vmo}, \\ \infty & \text{if } f \notin \mathrm{vmo}. \end{cases} \tag{3B.12}$$

Then we can write

$$k(x,z) = \sum_{j\ge 0} k_j(x)|z|^{-n}\Omega_j\Big(\frac{z}{|z|}\Big), \tag{3B.13}$$

where $\{\Omega_j\}_j$ is an orthonormal set of spherical harmonics on $S^{n-1}$, all odd, and for each $j \in \mathbb{N}$ we have

$$\|k_j\|_{L^\infty\cap\mathrm{vmo}} \le C_N \langle j\rangle^{-N} \quad \text{for every } N \in \mathbb{N}. \tag{3B.14}$$

In place of (3B.4)–(3B.6), we take

$$Kf(x) := \mathrm{PV}\int_{\mathbb{R}^n} k(x, x-y) f(y)\,dy, \quad x \in \mathbb{R}^n, \tag{3B.15}$$

$$K_\varphi f(x) := \mathrm{PV}\int_{\mathbb{R}^n} k(\varphi(x), \varphi(x)-\varphi(y)) f(y)\,dy, \quad x \in \mathbb{R}^n, \tag{3B.16}$$

and write

$$K_\varphi f(x) = \mathrm{PV}\int_{\mathbb{R}^n} k(\varphi(x), D\varphi(x)(x-y)) f(y)\,dy + R_\varphi f(x), \quad x \in \mathbb{R}^n. \tag{3B.17}$$

Using (3B.13)–(3B.14), we can write these as rapidly convergent series, and deduce that

$$\varphi \in \mathfrak{T}(\mathbb{R}^n) \implies M_\chi R_\varphi M_\chi : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n) \quad \text{compact for all } p \in (1,\infty) \tag{3B.18}$$

whenever $\chi \in \mathscr{C}_0^\infty(\mathbb{R}^n)$. Implementing this for (3B.16) involves using the following result:

**Lemma 3.4.** *The function spaces* $\mathrm{bmo}(\mathbb{R}^n)$ *and* $\mathrm{vmo}(\mathbb{R}^n)$ *are invariant under* $u \mapsto u \circ \varphi$, *provided* $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ *is a bi-Lipschitz map.*

*Proof.* This has the same proof as Proposition D.5 (see also [Taylor 2009, Proposition 3.3; Bourdaud et al. 2002, Theorem 2, p. 516]). □

As in (3B.8), the integral on the right-hand side of (3B.17) defines an operator

$$\widetilde{K}_\varphi \in \mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}. \tag{3B.19}$$

We use these results to analyze how an operator $P = p(x,D) \in \mathrm{OP}(L^\infty\cap\mathrm{vmo})S^0_{\mathrm{cl}}$ transforms under a map $\varphi \in \mathfrak{T}(\mathbb{R}^n)$. In more detail, given $P : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$, set

$$P_\varphi g(x) := Pf(\varphi(x)), \quad f \in L^p(\mathbb{R}^n),\ g(x) = f(\varphi(x)). \tag{3B.20}$$

Our hypothesis (3B.1) implies $\|g\|_{L^p} \approx \|f\|_{L^p}$, so $P_\varphi : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$. We claim that $P_\varphi \in \mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$, at least modulo an operator with the compactness property (3B.18). Furthermore, we obtain a formula for its principal symbol.

We take $p(x,\xi)$ to be homogeneous of degree 0 in $\xi$. To start, we assume

$$p(x,\xi) = -p(x,-\xi). \tag{3B.21}$$

Now

$$Pf(x) = \mathrm{PV}\int_{\mathbb{R}^n} k(x, x-y) f(y)\, dy, \quad x \in \mathbb{R}^n, \tag{3B.22}$$

with

$$k(x,z) = (2\pi)^{-n} \int_{\mathbb{R}^n} p(x,\xi) e^{iz\cdot\xi}\, d\xi, \tag{3B.23}$$

so

$$p(x,\xi) = \int_{\mathbb{R}^n} k(x,z) e^{-iz\cdot\xi}\, dz. \tag{3B.24}$$

Note that

$$p(x,\xi) = \sum_{j\ge 0} p_j(x)\Omega_j\Big(\frac{\xi}{|\xi|}\Big), \tag{3B.25}$$

with $\{\Omega_j\}_j$ as in (3B.13) (again, all odd) and

$$\|p_j\|_{L^\infty\cap\mathrm{vmo}} \le C_N \langle j\rangle^{-N} \quad \text{for all } N \in \mathbb{N}. \tag{3B.26}$$

It follows that $k(x,z)$ satisfies (3B.10)–(3B.11). Hence, (3B.15)–(3B.19) apply. Consequently, with $J_\varphi(y) := |\det D\varphi(y)|$,

$$\begin{aligned} P_\varphi g(x) &= Pf(\varphi(x)) && (3B.27)\\ &= \mathrm{PV}\int_{\mathbb{R}^n} k(\varphi(x), \varphi(x) - y') f(y')\, dy' && (3B.28)\\ &= \mathrm{PV}\int_{\mathbb{R}^n} k(\varphi(x), \varphi(x) - \varphi(y)) f(\varphi(y)) J_\varphi(y)\, dy && (3B.29)\\ &= \mathrm{PV}\int_{\mathbb{R}^n} k(\varphi(x), \varphi(x) - \varphi(y)) g(y) J_\varphi(y)\, dy. && (3B.30) \end{aligned}$$

Applying (3B.15)–(3B.18), we have

$$P_\varphi g(x) = \mathrm{PV}\int_{\mathbb{R}^n} k(\varphi(x), D\varphi(x)(x-y)) g(y) J_\varphi(y)\, dy + R_{1\varphi}, \tag{3B.31}$$

where $R_{1\varphi}$ has the compactness property (3B.18). Also, $J_\varphi \in L^\infty \cap \mathrm{vmo}$, so we can use the commutator estimate from [Coifman et al. 1976] to replace $J_\varphi(y)$ by $J_\varphi(x)$ in (3B.31), replacing $R_{1\varphi}$ by $R_{2\varphi}$, also satisfying (3B.18). Consequently, we have

$$P_\varphi g(x) = (2\pi)^{-n} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} p_\varphi(x,\xi) e^{i(x-y)\cdot\xi} g(y)\, dy\, d\xi + R_{2\varphi}, \tag{3B.32}$$

and

$$(2\pi)^{-n} \int_{\mathbb{R}^n} p_\varphi(x,\xi') e^{iz\cdot\xi'}\, d\xi' = J_\varphi(x) k(\varphi(x), D\varphi(x) z). \tag{3B.33}$$

Taking $\xi' = D\varphi(x)^{\mathrm{T}}\xi$ gives $d\xi' = J_\varphi(x)\,d\xi$. We have cancellation of the factors $J_\varphi(x)$, hence

$$(2\pi)^{-n}\int_{\mathbb{R}^n} p_\varphi(x, D\varphi(x)^{\mathrm{T}}\xi)e^{i\nabla\varphi(x)z\cdot\xi}\,d\xi = k(\varphi(x), D\varphi(x)z). \tag{3B.34}$$

Hence, with

$$\sigma(x,\xi) = p_\varphi(x, D\varphi(x)^{\mathrm{T}}\xi), \quad z' = D\varphi(x)z, \tag{3B.35}$$

we have

$$(2\pi)^{-n}\int_{\mathbb{R}^n} \sigma(x,\xi)e^{iz'\cdot\xi}\,d\xi = k(\varphi(x), z'), \tag{3B.36}$$

so

$$\sigma(x,\xi) = \int_{\mathbb{R}^n} k(\varphi(x), z')e^{-iz'\cdot\xi}\,dz'. \tag{3B.37}$$

Comparison with (3B.24) yields the formula

$$p_\varphi(x, D\varphi(x)^{\mathrm{T}}\xi) = p(\varphi(x),\xi). \tag{3B.38}$$

This has been derived for $p(x,\xi)$ satisfying (3B.21). We now address the general case.

**Theorem 3.5.** *Assume $\varphi \in \mathfrak{T}(\mathbb{R}^n)$. Given $P \in \mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$ with principal symbol $p(x,\xi)$ and $P_\varphi$ defined by* (3B.20), *one can decompose*

$$P_\varphi = p_\varphi(x, D) + R_\varphi \tag{3B.39}$$

*with $R_\varphi$ satisfying* (3B.18) *and $p_\varphi(x, D) \in \mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$ satisfying* (3B.38).

*Proof.* We have this when $p(x,\xi)$ satisfies (3B.21). It remains to treat the case $p(x,-\xi) = p(x,\xi)$. For this, we can write

$$p(x, D) = \sum_{j=1}^n q_j(x, D)s_j(x, D), \quad \text{where} \quad s_j(x,\xi) = \frac{\xi_j}{|\xi|}, \quad q_j(x,\xi) = p(x,\xi)\frac{\xi_j}{|\xi|}. \tag{3B.40}$$

The previous analysis holds for the factors $q_j(x, D)$ and $s_j(x, D)$, and our conclusion follows by basic operator calculus for $\mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$. □

**3C.** ***Admissible coordinate changes on a* Lip ∩ vmo$_1$ *surface.*** Let $\varphi : \mathbb{R}^n \to \mathbb{R}^{n+\ell}$ have the form $\varphi(x) = (x, \varphi_0(x))$ with $D\varphi_0(x) \in L^\infty(\mathbb{R}^n) \cap \mathrm{vmo}(\mathbb{R}^n)$, as in Section 2B. Thus $\varphi$ maps $\mathbb{R}^n$ onto an $n$-dimensional surface $\Sigma$. Let $V \subset \mathbb{R}^{n+\ell}$ be an $n$-dimensional linear space. If $V$ is not too far from $\mathbb{R}^n$ (depending on $\|D\varphi_0\|_{L^\infty}$), then $\Sigma$ is also a graph over $V$ and we have the coordinate change map

$$\psi : \mathbb{R}^n \longrightarrow V, \quad \psi(x) = Q\varphi(x), \tag{3C.1}$$

where $Q : \mathbb{R}^{n+\ell} \to V$ is the orthogonal projection. Consequently,

$$\psi(x) = Q\begin{pmatrix} x \\ \varphi_0(x)\end{pmatrix}, \quad D\psi(x)v = Q\begin{pmatrix} v \\ D\varphi_0(x)v\end{pmatrix}. \tag{3C.2}$$

Consequently,

$$\tau[\psi(x) - \psi(y)] + (1-\tau)D\psi(x)(x-y) = Q\begin{pmatrix} x-y \\ \tau[\varphi_0(x)-\varphi_0(y)]+(1-\tau)D\varphi_0(x)(x-y)\end{pmatrix}. \tag{3C.3}$$

Recall that the condition for Theorem 3.3 to apply is that (3C.3) has norm at least $\kappa|x-y|$ for some $\kappa > 0$, for $x,\ y \in \mathbb{R}^n$, $\tau \in [0,1]$. We see that the norm of (3C.3) is at least

$$\|Q(x-y)\| - \gamma(x,y), \tag{3C.4}$$

where, with $Q_0$ denoting the orthogonal projection of $\mathbb{R}^{n+\ell}$ onto $\mathbb{R}^n$,

$$\begin{aligned}\gamma(x,y) &= \|Q(I-Q_0)(\tau[\varphi_0(x)-\varphi_0(y)] + (1-\tau)D\varphi_0(x)(x-y))\| \\ &\le \|D\varphi_0\|_{L^\infty}\|Q(I-Q_0)\| \cdot |x-y|.\end{aligned} \tag{3C.5}$$

Since $Q(x-y) = (x-y) + (I-Q)Q_0(x-y)$, we deduce that the norm of (3C.3) is at least

$$\big(1 - \|(I-Q)Q_0\| - \|(I-Q_0)Q\| \cdot \|D\varphi_0\|_{L^\infty}\big)|x-y|. \tag{3C.6}$$

Consequently, Theorem 3.3 applies as long as

$$\|(I-Q)Q_0\| + \|(I-Q_0)Q\| \cdot \|D\varphi_0\|_{L^\infty} < 1. \tag{3C.7}$$

This in turn holds provided

$$\|Q-Q_0\| < (1+\|D\varphi_0\|_{L^\infty})^{-1}. \tag{3C.8}$$

We have the following conclusion:

**Proposition 3.6.** *Let $\psi : \mathbb{R}^n \to V$ be as constructed in* (3C.1). *Assume* (3C.8) *holds, where $Q$ and $Q_0$ are the orthogonal projections of $\mathbb{R}^{n+\ell}$ onto $V$ and $\mathbb{R}^n$, respectively. Take a linear isomorphism $J : V \to \mathbb{R}^n$. Then $J \circ \psi$ belongs to $\mathfrak{T}(\mathbb{R}^n)$.*

**3D.** ***Remark on double layer potentials.*** Assume that a kernel

$$E : \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{R}, \text{ which is a smooth function, positive homogeneous of degree } -(n+1) \text{ and satisfying } E(-X) = E(X) \text{ for all } X \in \mathbb{R}^{n+1}\setminus\{0\}, \tag{3D.1}$$

has been given. Also, let $\Omega \subset \mathbb{R}^{n+1}$ be a bounded $\mathrm{Lip} \cap \mathrm{vmo}_1$ domain and consider the singular integral operator

$$Kf(X) := \mathrm{PV}\int_{\partial\Omega} \langle \nu(X), X-Y\rangle E(X-Y) f(Y)\, d\sigma(Y), \quad X \in \partial\Omega, \tag{3D.2}$$

where $\nu$ and $\sigma$ are, respectively, the outward unit normal and surface measure on $\partial\Omega$. To study this, focus on a local version of (3D.2) of the following sort. Let

$$\varphi_0 : \mathbb{O} \longrightarrow \mathbb{R} \quad \text{Lipschitz with } \nabla\varphi_0 \in \mathrm{vmo}, \tag{3D.3}$$

where $\mathbb{O} \subset \mathbb{R}^n$ is open, be such that its graph is contained in $\partial\Omega$ and define the Lipschitz map $\varphi : \mathbb{O} \to \mathbb{R}^{n+1}$ by setting

$$\varphi(x) := (x, \varphi_0(x)) \quad \text{for all } x \in \mathbb{O}. \tag{3D.4}$$

Then, in these local coordinates, $K$ takes the form

$$K_\varphi f(x) = \mathrm{PV}\int_{\mathbb{O}} \langle (\nabla\varphi_0(x), -1), \varphi(x)-\varphi(y)\rangle E(\varphi(x)-\varphi(y)) f(y)\, dy. \tag{3D.5}$$

Its "sharp" form, obtained by replacing $\varphi(x)-\varphi(y)$ with $D\varphi(x)(x-y)$, is then

$$\begin{aligned} K^{\#}_{\varphi} f(x) &:= \mathrm{PV}\int_{\mathbb{O}} \langle(\nabla\varphi_0(x),-1), D\varphi(x)(x-y)\rangle E(D\varphi(x)(x-y))\, f(y)\, dy \\ &= \mathrm{PV}\int_{\mathbb{O}} \langle D\varphi(x)^{\mathrm{T}}(\nabla\varphi_0(x),-1), x-y\rangle E(D\varphi(x)(x-y))\, f(y)\, dy \\ &= 0, \end{aligned} \tag{3D.6}$$

since

$$D\varphi(x)=\begin{pmatrix} I_{n\times n} \\ \nabla\varphi_0(x)\end{pmatrix} \implies D\varphi(x)^{\mathrm{T}}(\nabla\varphi_0(x),-1)=\begin{pmatrix} I_{n\times n} & \nabla\varphi_0(x)\end{pmatrix}\begin{pmatrix}\nabla\varphi_0(x)^{\mathrm{T}} \\ -1\end{pmatrix}=0. \tag{3D.7}$$

In concert with our local compactness result, according to which $K_\varphi - K^{\#}_\varphi$ is compact on $L^p$ for each $p\in(1,\infty)$, this ultimately gives that

$$\begin{gathered}\text{if } \Omega\subset\mathbb{R}^{n+1} \text{ is a bounded Lip} \cap \mathrm{vmo}_1 \text{ domain and } E \text{ is as in (3D.1)} \\ \text{then } K \text{ from (3D.2) is compact on } L^p(\partial\Omega), \text{ for each } p\in(1,\infty). \end{gathered} \tag{3D.8}$$

Of course, the above result contains as a particular case the fact (which is a key result in the work of Fabes, Jodeit and Rivière [Fabes et al. 1978]) that the principal value, harmonic, double layer operator

$$Kf(X) := \lim_{\varepsilon\to 0^+} \frac{1}{\omega_n}\int_{\substack{Y\in\partial\Omega \\ |X-Y|>\varepsilon}} \frac{\langle \nu(Y), Y-X\rangle}{|X-Y|^{n+1}} f(Y)\, d\sigma(Y), \quad X\in\partial\Omega, \tag{3D.9}$$

is compact on $L^p(\partial\Omega)$ for each $p\in(1,\infty)$ if $\Omega\subset\mathbb{R}^{n+1}$ is a bounded $\mathscr{C}^1$ domain.

**3E.** ***Cauchy integrals and their symbols.*** Given $\ell\in\mathbb{N}$, let $M(\ell,\mathbb{C})$ denote the collection of $\ell\times\ell$ matrices with complex entries. Let $\mathfrak{D}$ be a first-order elliptic $\ell\times\ell$ system of differential operators on $\mathbb{R}^{n+1}$,

$$\mathfrak{D}u(x)=\sum_j A_j\partial_j u, \quad A_j\in M(\ell,\mathbb{C}). \tag{3E.1}$$

Thus $\sigma_{\mathfrak{D}}(\zeta)=i\sum_j A_j\zeta_j$ is invertible for each nonzero $\zeta\in\mathbb{R}^{n+1}$ and $\mathfrak{D}$ has a fundamental solution

$$k(z)=(2\pi)^{-(n+1)}\int_{\mathbb{R}^{n+1}} E(\zeta)e^{iz\cdot\zeta}\, d\zeta, \quad E(\zeta)=\sigma_{\mathfrak{D}}(\zeta)^{-1}, \tag{3E.2}$$

odd and homogeneous of degree $-n$ in $z$. If $\Omega\subset\mathbb{R}^{n+1}$ is a bounded UR (uniformly rectifiable) domain, we can form

$$\mathscr{B}f(x)=\int_{\partial\Omega} k(x-y)\, f(y)\, d\sigma(y), \quad x\in\Omega, \tag{3E.3}$$

with nontangential limits (see (4A.3))

$$\big(\mathscr{B}f\big|^{\mathrm{n.t.}}_{\partial\Omega}\big)(z) := \lim_{\Gamma_\kappa(x)\ni z\to x} \mathscr{B}f(z)=\frac{1}{2i}\sigma_{\mathfrak{D}}(\nu(x))^{-1}f(x)+Bf(x) \tag{3E.4}$$

for $\sigma$-a.e. $x\in\partial\Omega$, where $\Gamma_\kappa(x)\subset\Omega$ is a region of nontangential approach to $x\in\partial\Omega$ (see (4A.2)) and

$$Bf(x) := \mathrm{PV}\int_{\partial\Omega} k(x-y)\, f(y)\, d\sigma(y), \quad x\in\partial\Omega. \tag{3E.5}$$

One is hence motivated to consider the "Cauchy integral"

$$\mathscr{C}_{\mathfrak{D}} f(x) = i \int_{\partial\Omega} k(x-y)\sigma_{\mathfrak{D}}(\nu(y)) f(y)\, d\sigma(y), \quad x \in \Omega, \tag{3E.6}$$

with nontangential limits

$$\mathscr{C}_{\mathfrak{D}} f\big|^{\text{n.t.}}_{\partial\Omega}(x) = \tfrac{1}{2} f(x) + C_{\mathfrak{D}} f(x) \tag{3E.7}$$

for $\sigma$-a.e. $x \in \partial\Omega$, where

$$C_{\mathfrak{D}} f(x) := i \,\mathrm{PV}\!\int_{\partial\Omega} k(x-y)\sigma_{\mathfrak{D}}(\nu(y)) f(y)\, d\sigma(y), \quad x \in \partial\Omega. \tag{3E.8}$$

As shown in [Mitrea et al. 2015], a reproducing formula yields

$$P_{\mathfrak{D}} = \tfrac{1}{2} I + C_{\mathfrak{D}} \quad \Longrightarrow \quad P_{\mathfrak{D}}^2 = P_{\mathfrak{D}}. \tag{3E.9}$$

They study this in the setting of UR domains (and also for variable coefficient situations, which for simplicity we do not take up here in detail). The operator $P_{\mathfrak{D}}$ is a Calderón projector.

Here, we take $\Omega$ to be a $\mathrm{Lip} \cap \mathrm{vmo}_1$ domain and analyze the principal symbol of $P_{\mathfrak{D}}$ as a projection-valued function on $T^*\partial^*\Omega \setminus 0$. To start, we recall from (3A.10) that, for $B$ in (3E.5),

$$\sigma_B(w,\xi) = \int_{T_w \partial^*\Omega} k(z^0) e^{-iz^0\cdot\xi}\, dz^0, \quad w \in \partial^*\Omega. \tag{3E.10}$$

Plugging in (3E.2) and using basic Fourier analysis, we obtain

$$\sigma_B(w,\xi) = \frac{1}{2\pi} \,\mathrm{PV}\!\int_{-\infty}^{\infty} E(\xi + s\nu(w))\, ds. \tag{3E.11}$$

We then have

$$\sigma_{C_{\mathfrak{D}}}(w,\xi) = \frac{i}{2\pi} \,\mathrm{PV}\!\int_{-\infty}^{\infty} \sigma_{\mathfrak{D}}(\xi + is\nu(w))^{-1} \sigma_{\mathfrak{D}}(\nu(w))\, ds. \tag{3E.12}$$

Now $\sigma_{\mathfrak{D}}(\xi + s\nu(w)) = \sigma_{\mathfrak{D}}(\xi) + s\sigma_{\mathfrak{D}}(\nu(w))$, so

$$\sigma_{\mathfrak{D}}(\xi + s\nu(w))^{-1} \sigma_{\mathfrak{D}}(\nu(w)) = (M(w,\xi) + sI)^{-1}, \tag{3E.13}$$

with

$$M(w,\xi) = \sigma_{\mathfrak{D}}(\nu(w))^{-1} \sigma_{\mathfrak{D}}(\xi). \tag{3E.14}$$

The invertibility of $\sigma_{\mathfrak{D}}(\xi + s\nu(w))$ and of $\sigma_{\mathfrak{D}}(\nu(w))$ imply that

$$\operatorname{Spec} M(w,\xi) \cap \mathbb{R} = \varnothing. \tag{3E.15}$$

We have

$$\sigma_{C_{\mathfrak{D}}}(w,\xi) = \frac{i}{2\pi} \,\mathrm{PV}\!\int_{-\infty}^{\infty} (sI + M(w,\xi))^{-1}\, ds. \tag{3E.16}$$

**Lemma 3.7.** *Assume $A \in M(\ell,\mathbb{C})$ and $\operatorname{Spec} A \cap \mathbb{R} = \varnothing$. Then*

$$\frac{1}{2\pi i} \int_{-\infty}^{\infty} (s-A)^{-1} e^{i\varepsilon s}\, ds = \begin{cases} e^{i\varepsilon A} P_+(A) & \text{if } \varepsilon > 0, \\ -e^{i\varepsilon A} P_-(A) & \text{if } \varepsilon < 0, \end{cases} \tag{3E.17}$$

*where* $P_+(A)$ *is the projection of* $\mathbb{C}^\ell$ *onto the linear span of the generalized eigenvectors of* $A$ *associated to eigenvalues in* Spec $A$ *with positive imaginary part annihilating those associated to eigenvectors with negative imaginary part, and* $P_-(A) = I - P_+(A)$*. Hence*

$$\frac{1}{2\pi i}\,\mathrm{PV}\int_{-\infty}^{\infty}(s-A)^{-1}\,ds = P_+(A) - \tfrac{1}{2}I. \tag{3E.18}$$

*Proof.* If $\varepsilon > 0$, the left-hand side of (3E.17) is equal to

$$\lim_{R\to\infty}\frac{1}{2\pi i}\int_{\partial D_R^+}(s-A)^{-1}\,ds, \tag{3E.19}$$

where $D_R := \{s \in \mathbb{C} : |s| < R\}$ and $D_R^+ := D_R \cap \{s \in \mathbb{C} : \operatorname{Im} s > 0\}$. This path integral stabilizes when $R > \|A\|$ and the desired conclusion in this case follows from the Riesz functional calculus. The treatment of the case when $\varepsilon < 0$ is similar. Then (3E.18) follows readily from (3E.17). □

We apply Lemma 3.7 to (3E.16) with $A := -M(w, \xi)$. Making use of the identity $P_+(-M) = P_-(M)$, we have the following conclusion:

**Proposition 3.8.** *The operator* $C_{\mathscr{D}}$ *and the associated Calderón projector, derived from the Cauchy integral* (3E.6) *via* (3E.7)–(3E.9), *have symbols given by*

$$\sigma_{C_{\mathscr{D}}}(w,\xi) = -\big(P_-(M(w,\xi)) - \tfrac{1}{2}I\big) = \tfrac{1}{2}I - P_-(M(w,\xi)) \tag{3E.20}$$

*and*

$$\sigma_{P_{\mathscr{D}}}(w,\xi) = P_+(M(w,\xi)) \tag{3E.21}$$

*respectively, with* $M(w,\xi)$ *as in* (3E.14) *and* $P_+(A)$ *as described in Lemma 3.7.*

**Remark 3.9.** Extensions of the results in this section to variable coefficient operators (acting between vector bundles) and to domains on manifolds can be worked out using the formalism developed in [Mitrea et al. 2015; ≥ 2015].

## 4. Applications to elliptic boundary problems

Here we apply the results of Sections 2–3 to several classes of elliptic boundary problems, including the Dirichlet problem for general strongly elliptic, second-order systems and general regular boundary problems for first-order elliptic systems of differential operators.

**4A.** ***Single layers and boundary problems for elliptic systems.*** Let $M$ be a smooth, compact, $(n+1)$-dimensional manifold equipped with a Riemannian metric tensor

$$g = \sum_{j,k} g_{jk}\,dx_j \otimes dx_k \quad \text{with } g_{jk} \in \mathscr{C}^2. \tag{4A.1}$$

Also, consider a $\mathrm{Lip} \cap \mathrm{vmo}_1$ domain $\Omega \subset M$ (see the discussion in the last part of Appendix E). Having some fixed $\kappa \in (0,\infty)$, for each $x \in \partial\Omega$ define the nontangential approach region with vertex at $x$ by setting

$$\Gamma_\kappa(x) := \{y \in \Omega : \mathrm{dist}(x,y) < (1+\kappa)\,\mathrm{dist}(y,\partial\Omega)\}. \tag{4A.2}$$

Next, given an arbitrary $u : \Omega \to \mathbb{C}$, define its nontangential maximal function and its pointwise nontangential boundary trace at $x \in \partial\Omega$, respectively, as

$$(\mathcal{N}_\kappa u)(x) := \sup\{|u(y)| : y \in \Gamma_\kappa(x)\}, \quad \big(u\big|^{\text{n.t.}}_{\partial\Omega}\big)(x) := \lim_{\Gamma_\kappa(x)\ni y\to x} u(y) \tag{4A.3}$$

whenever the limit exists. The parameter $\kappa$ plays a somewhat secondary role in the proceedings, since for any $\kappa_1, \kappa_2 \in (0,\infty)$ and $p \in (0,\infty)$ there exists $C = C(\kappa_1,\kappa_2, p) \in (1,\infty)$ with the property that

$$C^{-1}\|\mathcal{N}_{\kappa_1} u\|_{L^p(\partial\Omega)} \le \|\mathcal{N}_{\kappa_2} u\|_{L^p(\partial\Omega)} \le C\|\mathcal{N}_{\kappa_1} u\|_{L^p(\partial\Omega)} \tag{4A.4}$$

for each $u : \Omega \to \mathbb{C}$. Given this, we will simplify notation and write $\mathcal{N}$ in place of $\mathcal{N}_\kappa$.

Moving on, let $L$ be a second-order, strongly elliptic, $k \times k$ system of differential operators on $M$. Assume that, locally,

$$Lu = \sum_{i,j} \partial_j A^{ij}(x)\partial_j u + \sum_j B^j(x)\partial_j u + V(x)u, \tag{4A.5}$$

where

$$A^{ij} \in \mathscr{C}^2, \quad B^j \in \mathscr{C}^1, \quad V \in L^\infty. \tag{4A.6}$$

Also, suppose

$$L : H^{1,p}(M) \longrightarrow H^{-1,p}(M) \quad \text{is an isomorphism for } 1 < p < \infty. \tag{4A.7}$$

We want to solve the Dirichlet boundary problem

$$Lu = 0 \quad \text{on } \Omega, \quad u\big|^{\text{n.t.}}_{\partial\Omega} = f \in L^p(\partial\Omega), \quad \mathcal{N}u \in L^p(\partial\Omega) \tag{4A.8}$$

via the layer potential method. To this end, let $E$ denote the Schwartz kernel of $L^{-1}$, so that

$$L^{-1}v(x) = \int_M E(x,y)v(y)\, d\,\mathrm{Vol}(y), \quad x \in M, \tag{4A.9}$$

where $d\,\mathrm{Vol}$ stands for the volume element on $M$. Then, with $\sigma$ denoting the surface measure on $\partial\Omega$, define the single layer potential operator and its boundary version by

$$\mathcal{S}g(x) := \int_{\partial\Omega} E(x,y)g(y)\, d\sigma(y), \quad x \in M \setminus \partial\Omega, \qquad \text{and} \qquad Sg := \mathcal{S}g\big|^{\text{n.t.}}_{\partial\Omega} \quad \text{on } \partial\Omega. \tag{4A.10}$$

We want to solve (4A.8) in the form

$$u = \mathcal{S}g, \quad \text{where } g \text{ is chosen so that } Sg = f. \tag{4A.11}$$

As such, if $H^{s,p}(\partial\Omega)$ with $1 < p < \infty$ and $-1 \le s \le 1$ denotes the $L^p$-based scale of Sobolev spaces of fractional order $s$ on $\partial\Omega$, we would like to show

$$S : H^{-1,p}(\partial\Omega) \longrightarrow L^p(\partial\Omega) \quad \text{is Fredholm of index 0.} \tag{4A.12}$$

Since the adjoint of $S$ is the single layer associated with $L^*$ (which continues to be a second-order, strongly elliptic, $k \times k$ system of differential operators on $M$), this is further equivalent (with $q := p'$ the

Hölder conjugate exponent of $p$) to the condition that

$$S : L^q(\partial\Omega) \longrightarrow H^{1,q}(\partial\Omega) \quad \text{is Fredholm of index } 0. \tag{4A.13}$$

Such a result was established for $q$ close to 2, when $\Omega$ is a Lipschitz domain, in Chapter 3 of [Mitrea et al. 2001]. The argument made use of a Rellich-type identity. In the scalar case the result was established (in the setting of regular SKT domains) in [Hofmann et al. 2010, Section 6.4], and applied in Section 7.1 of that paper to the Dirichlet problem. If $\partial\Omega$ is smooth, it is standard that $S$ is in $\mathrm{OP}\, S^{-1}(\partial\Omega)$ and it is strongly elliptic, from which (4A.12) and (4A.13) follow. Here is what we propose:

**Proposition 4.1.** *Let* $\Omega$ *be a* $\mathrm{Lip} \cap \mathrm{vmo}_1$ *domain and let* $L$ *be a second-order*, *strongly elliptic*, $k \times k$ *system of differential operators on* $M$ *as in* (4A.5)–(4A.6) *and satisfying* (4A.7). *Then* (4A.12) *holds for all* $p \in (1, \infty)$ *and* (4A.13) *holds for all* $q \in (1, \infty)$.

*Proof.* We start with the proof of (4A.13). Pick $L^\infty \cap \mathrm{vmo}$ vector fields $X_j$, $1 \le j \le N$, tangent to $\partial\Omega$, such that

$$\sum_{j=1}^{N} |X_j(x)| \ge A > 0 \quad \text{for a.e. } x \in \partial\Omega. \tag{4A.14}$$

Then let $\nabla_T f := \{X_j f : 1 \le j \le N\}$. We have $\nabla_T S : L^q(\partial\Omega) \to L^q(\partial\Omega)$ for all $q \in (1, \infty)$. Theorem 2.4 (or rather its standard "variable coefficient" extension) implies

$$\nabla_T S = k_0(x, D) + R, \quad k_0(x, D) \in \mathrm{OP}(L^\infty \cap \mathrm{vmo}) S^0_{\mathrm{cl}} \tag{4A.15}$$

with $R$ compact on $L^q(\partial\Omega)$. At this point we make the following:

**Claim.** *We have the (overdetermined) ellipticity property*

$$\|k_0(x, \xi)v\| \ge A_0 \|v\|, \quad A_0 > 0. \tag{4A.16}$$

Assuming for now this claim (whose proof will be provided later), we obtain that

$$k_0^*(x, D) k_0(x, D) \in \mathrm{OP}(L^\infty \cap \mathrm{vmo}) S^0_{\mathrm{cl}} \quad \text{mod compacts} \tag{4A.17}$$

is a (determined) elliptic operator, so it has a parametrix $Q \in \mathrm{OP}(L^\infty \cap \mathrm{vmo}) S^0_{\mathrm{cl}}$ (see Appendix C). Hence,

$$Q k_0^*(x, D) \nabla_T S = I + R_1, \quad \text{with } R_1 \text{ compact on } L^q(\partial\Omega). \tag{4A.18}$$

This implies that

$$S : L^q(\partial\Omega) \longrightarrow H^{-1,q}(\partial\Omega) \quad \text{is semi-Fredholm;} \tag{4A.19}$$

namely, it has closed range and finite-dimensional null space.

To complete the argument, we take a continuous family $L_\tau$, $\tau \in [0, 1]$, of second-order, strongly elliptic operators on $M$ such that $L_1 = L$ and $L_0$ is *scalar*. This gives a norm-continuous family

$$S_\tau : L^q(\partial\Omega) \longrightarrow H^{1,q}(\partial\Omega), \quad \text{all semi-Fredholm.} \tag{4A.20}$$

We know that $S_0$ is Fredholm of index 0. Hence, so are all the operators $S_\tau$ in (4A.20). This gives (4A.13), which, by duality, also yields (4A.12).

Now we return to the proof of the claim made in (4A.16). That is, we shall establish the (overdetermined) ellipticity of $k_0(x,D) \in \mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$ arising in (4A.15) (which is equal modulo a compact operator to $\nabla_T S$). To begin, we discuss the smooth case. If $\partial\Omega$ is smooth and $L$ is strongly elliptic of second order with smooth coefficients, then actually $S$ is in $\mathrm{OPS}^{-1}(\partial\Omega)$ and this operator is strongly elliptic. In fact, given $(x,\xi) \in T^*\partial\Omega \setminus 0$, and with $\nu \in T^*_x\partial\Omega$ the outward unit conormal to $\partial\Omega$, we have

$$\sigma_S(x,\xi) = C_n \int_{-\infty}^{+\infty} \sigma_E(x,\xi+t\nu)\,dt = C_n \int_{-\infty}^{+\infty} \sigma_L(x,\xi+t\nu)^{-1}\,dt. \tag{4A.21}$$

This is seen as in [Taylor 1996, (11.11)–(11.12) in Chapter 7], where we take $m=-2$, $x_n=0$. Strong ellipticity of $S$ then follows from (4A.21), keeping in mind the strong ellipticity of $L$. Specifically, $\sigma_S(x,\xi)$ is positive homogeneous of degree $-1$ in $\xi$ and the integrals in (4A.21) are absolutely convergent since $|\sigma_L(x,\xi+t\nu(x))^{-1}| \le C(|\xi|^2+t^2)^{-1}$. Thus, for any section $\eta$ and any $0 \ne \xi \in T^*_x\partial\Omega \subset T^*_xM$, we may estimate

$$\begin{aligned}\langle -\sigma_S(x,\xi)\eta,\eta\rangle_x = C_n \int_{-\infty}^{+\infty} \langle -\sigma_L(x,\xi+t\nu(x))^{-1}\eta,\eta\rangle\,dt &\ge C|\eta|^2 \int_{-\infty}^{+\infty} (|\xi|^2+t^2)^{-1}\,dt \\ &\ge C|\eta|^2|\xi|^{-1}\end{aligned} \tag{4A.22}$$

for some $C>0$. This yields the strong ellipticity of $S$. Next, since $\sigma_{X_jS} = \sigma_{X_j}\sigma_S$, the ellipticity of $\nabla_T S$ is an immediate consequence of what we have just proved and (4A.14).

To tackle the case when $\Omega$ is a $\mathrm{Lip} \cap \mathrm{vmo}_1$ domain, we take local graph coordinates $\varphi(x) = (x,\varphi_0(x))$ and arrange that the vector fields $\{X_j\}_{1\le j\le N}$ include those associated with coordinate differentiation. The integral kernel $E(x,y)$ has the form

$$E(x,y) = E_0(x,x-y) + r(x,y), \tag{4A.23}$$

where $E_0(x,z)$ is smooth on $\{z \ne 0\}$ and homogeneous of degree $-(n-1)$ in $z$ (note that $\dim\partial\Omega = n$) and $r(x,y)$ has lower order. See the analysis in [Mitrea et al. 2001]. Locally, the operator $S$ has the form

$$Sg(x) = \int_{\mathbb{R}^n} E_0(\varphi(x),\varphi(x)-\varphi(y))g(y)\Sigma(y)\,dy + Rg(x), \quad x \in \mathbb{R}^n, \tag{4A.24}$$

where $d\sigma(y) = \Sigma(y)\,dy$ and $R$ denotes the integral operator with kernel $r(x,y)$. Hence, for each $j \in \{1,\dots,n\}$,

$$\partial_j Sg(x) = \mathrm{PV}\!\int_{\mathbb{R}^n} \partial_j\varphi(x)\cdot\nabla_2 E_0(\varphi(x),\varphi(x)-\varphi(y))g(y)\Sigma(y)\,dy + R_jg(x), \quad x \in \mathbb{R}^n, \tag{4A.25}$$

where here and below $R_j$ will denote (perhaps different) operators that are compact on $L^p$ for $1<p<\infty$. Theorem 2.4 (or rather its natural "variable coefficient" extension from Section 2C) gives

$$\partial_j Sg(x) = \mathrm{PV}\!\int_{\mathbb{R}^n} \partial_j\varphi(x)\cdot\nabla_2 E_0(\varphi(x),D\varphi(x)(x-y))g(y)\Sigma(y)\,dy + R_jg(x), \quad x \in \mathbb{R}^n; \tag{4A.26}$$

that is,

$$\partial_j Sg(x) = T_j(x, D)(\Sigma g)(x) + R_j g(x), \quad x \in \mathbb{R}^n, \tag{4A.27}$$

where $T_j(x, D)(\Sigma g)(x)$ is given by the principal value integral in (4A.26). We therefore have that $T_j(x, D)$ is in $\mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$, with symbol

$$T_j(x, \xi) = \int_{\mathbb{R}^n} e^{-iz\cdot\xi} \partial_j \varphi(x) \cdot \nabla_2 E_0(\varphi(x), D\varphi(x)z)\, dz. \tag{4A.28}$$

Given that $L$ is a $k \times k$ system, $T_j(x, \xi)$ is a $k \times k$ matrix, i.e., $T_j(x, \xi) \in M(k, \mathbb{C})$ for $\xi \neq 0$ and a.e. $x$. We need to show that there exists $C > 0$ such that, for all $\xi \neq 0$ and $v \in \mathbb{C}^k$,

$$\sum_j \|T_j(x, \xi)v\| \geq C\|v\| \quad \text{for a.e. } x. \tag{4A.29}$$

Recall that $\varphi$ has the form (2B.1), so $D\varphi(x) : \mathbb{R}^n \to \mathbb{R}^{n+1}$ has the form

$$D\varphi(x) = \begin{pmatrix} I \\ D\varphi_0(x) \end{pmatrix}, \quad D\varphi_0(x) : \mathbb{R}^n \to \mathbb{R} \tag{4A.30}$$

for a.e. $x \in \mathbb{R}^n$. Freezing coefficients at a point where $\varphi$ is differentiable, we can rephrase our task as follows: Let $L_0(\zeta)$ be a matrix in $M(k, \mathbb{C})$ whose entries are homogeneous polynomials of degree 2 in $\zeta \in \mathbb{R}^{n+1}$ and which is positive definite for each $\zeta \neq 0$. For $\zeta \neq 0$ set $E_0(\zeta) := L_0(\zeta)^{-1}$. In addition, consider a linear mapping $A : \mathbb{R}^n \to \mathbb{R}^{n+1}$ of the form

$$A = \begin{pmatrix} I \\ A_0 \end{pmatrix}, \quad A_0 : \mathbb{R}^n \to \mathbb{R}. \tag{4A.31}$$

Let $A_0$ run over a compact set in $\mathscr{L}(\mathbb{R}^n, \mathbb{R})$. Also let $L_0$ and $E_0 = L_0^{-1}$ run over compact sets of symbols. Take

$$T_j(\xi) := \int_{\mathbb{R}^n} e^{-iz\cdot\xi} Ae_j \cdot \nabla E_0(Az)\, dz, \tag{4A.32}$$

where $\{e_j\}_{1\leq j\leq n}$ denotes the standard orthonormal basis of $\mathbb{R}^n$. We need to prove that there exists a finite constant $C > 0$ such that, for all $v \in \mathbb{C}^k$ and $\xi \neq 0$,

$$\sum_j \|T_j(\xi)v\| \geq C\|v\|, \tag{4A.33}$$

uniformly in $A_0$, $L_0$, $E_0$. This is equivalent to the ellipticity of $\nabla_T S$ if $\varphi(x) = Ax$, so $\partial\Omega$ is a hyperplane in $\mathbb{R}^{n+1}$. In this case, the previous analysis applies, since $S \in \mathrm{OPS}^{-1}(\partial\Omega)$ is strongly elliptic, and (4A.33) follows.

This finishes the proof of the claim in (4A.16), which, in turn, finishes the proof of Proposition 4.1. □

We next note a regularity result, under the assumption that $\Omega$ is a $\mathrm{Lip} \cap \mathrm{vmo}_1$ domain. Let us temporarily denote

$$S_{s,p} = S : H^{s,p}(\partial\Omega) \longrightarrow H^{s+1,p}(\partial\Omega), \quad s \in \{0, -1\}, \tag{4A.34}$$

with adjoint

$$S^*_{s,p} = S^* : H^{-1-s,q}(\partial\Omega) \longrightarrow H^{-s,q}(\partial\Omega), \quad q = p'. \tag{4A.35}$$

Clearly the null spaces $\mathrm{Ker}(S_{s,p})$ and $\mathrm{Ker}(S^*_{s,p})$ of these operators satisfy

$$\mathrm{Ker}(S_{0,p}) \subset \mathrm{Ker}(S_{-1,p}), \quad \mathrm{Ker}(S^*_{-1,p}) \subset \mathrm{Ker}(S^*_{0,p}), \tag{4A.36}$$

so the vanishing index property established in Proposition 4.1 forces

$$\mathrm{Ker}(S_{0,p}) = \mathrm{Ker}(S_{-1,p}) \quad \text{and} \quad \mathrm{Ker}(S^*_{-1,p}) = \mathrm{Ker}(S^*_{0,p}). \tag{4A.37}$$

Also,

$$1 < p < \tilde{p} < \infty \quad \Longrightarrow \quad \mathrm{Ker}(S_{0,\tilde{p}}) = \mathrm{Ker}(S_{0,p}), \quad \mathrm{Ker}(S^*_{0,p}) = \mathrm{Ker}(S^*_{0,\tilde{p}}) \tag{4A.38}$$

and, again, the aforementioned vanishing index property implies

$$\mathrm{Ker}(S_{0,p}) = \mathrm{Ker}(S_{0,\tilde{p}}). \tag{4A.39}$$

Collectively, (4A.37) and (4A.39) prove the following regularity result:

**Proposition 4.2.** *Assume that* $\Omega$ *is a* $\mathrm{Lip} \cap \mathrm{vmo}_1$ *domain in* $M$ *and suppose* $L$ *is a second-order, strongly elliptic system of differential operators on* $M$ *as in* (4A.5)–(4A.6) *and satisfying* (4A.7). *Then, given* $f \in H^{-1,p}(\partial\Omega)$ *for some* $p \in (1,\infty)$, *one has*

$$Sf = 0 \quad \Longrightarrow \quad f \in \bigcap_{1<q<\infty} L^q(\partial\Omega). \tag{4A.40}$$

Recall that standard Lipschitz theory (see [Mitrea et al. 2001]) gives

$$f \in L^p(\partial\Omega) \text{ with } p \in (1,\infty) \text{ and } u := \mathscr{S} f \quad \Longrightarrow \quad \begin{cases} Lu = 0 \quad \text{on } M \setminus \partial\Omega, \\ \mathcal{N}u,\ \mathcal{N}(\nabla u) \in L^p(\partial\Omega), \\ u\big|^{\text{n.t.}}_{\partial\Omega} = Sf, \end{cases} \tag{4A.41}$$

and

$$f \in H^{-1,p}(\partial\Omega) \text{ with } p \in (1,\infty) \text{ and } u := \mathscr{S} f \quad \Longrightarrow \quad \begin{cases} Lu = 0 \quad \text{on } M \setminus \partial\Omega, \\ \mathcal{N}u \in L^p(\partial\Omega), \\ u\big|^{\text{n.t.}}_{\partial\Omega} = Sf. \end{cases} \tag{4A.42}$$

In addition, we single out the following additional properties. Let $H^{s,p}(\Omega)$, with $s \in \mathbb{R}$ and $p \in (1,\infty)$ stand for the $L^p$-based Sobolev space of fractional smoothness $s$ in $\Omega$. Also, let $\mathrm{Tr}: H^{1,2}(\Omega) \to H^{\frac{1}{2},2}(\partial\Omega)$ denote the boundary trace operator in the sense of Sobolev spaces, and set $H^{1,2}_0(\Omega) := \mathrm{Ker}\,\mathrm{Tr}$. Then

$$f \in L^2(\partial\Omega) \quad \Longrightarrow \quad u := \mathscr{S} f \in H^1(\Omega), \quad \mathrm{Tr}\, u = u\big|^{\text{n.t.}}_{\partial\Omega} = Sf. \tag{4A.43}$$

These considerations are relevant in the context of the following well-posedness result:

**Theorem 4.3.** *Suppose* $\Omega \subset M$ *is a* $\mathrm{Lip} \cap \mathrm{vmo}_1$ *domain and suppose* $L$ *is a second-order, strongly elliptic system of differential operators on* $M$ *as in* (4A.5)–(4A.6) *and satisfying* (4A.7). *Set*

$$\Omega_+ := \Omega, \quad \Omega_- := M \setminus \overline{\Omega} \tag{4A.44}$$

*and assume that the following nondegeneracy conditions hold*:

$$\begin{aligned} &u \in H_0^{1,2}(\Omega_+), \quad Lu = 0 \ \textit{in}\ \Omega_+ \quad \Longrightarrow \quad u = 0 \quad \textit{in}\ \Omega_+, \\ &u \in H_0^{1,2}(\Omega_-), \quad Lu = 0 \ \textit{in}\ \Omega_- \quad \Longrightarrow \quad u = 0 \quad \textit{in}\ \Omega_-. \end{aligned} \tag{4A.45}$$

*Then*

$$\begin{aligned} S : H^{-1,p}(\partial\Omega) \longrightarrow L^p(\partial\Omega) \quad &\textit{is invertible for each}\ \ p \in (1,\infty), \\ S : L^p(\partial\Omega) \longrightarrow H^{1,p}(\partial\Omega) \quad &\textit{is invertible for each}\ \ p \in (1,\infty). \end{aligned} \tag{4A.46}$$

*In particular*, *the Dirichlet problem*

$$Lu = 0 \quad \textit{in}\ \Omega, \quad u\big|_{\partial\Omega}^{\text{n.t.}} = f \in L^p(\partial\Omega), \quad \mathcal{N}u \in L^p(\partial\Omega) \tag{4A.47}$$

*is well posed and its unique solution is given by* $u = \mathcal{S}(S^{-1}f)$, *where* $S^{-1}f \in H^{-1,p}(\partial\Omega)$.

*Furthermore*, *the regularity problem*

$$Lu = 0 \quad \textit{in}\ \Omega, \quad u\big|_{\partial\Omega}^{\text{n.t.}} = f \in H^{1,p}(\partial\Omega), \quad \mathcal{N}u,\ \mathcal{N}(\nabla u) \in L^p(\partial\Omega), \tag{4A.48}$$

*is well posed and its unique solution is given by* $u = \mathcal{S}(S^{-1}f)$, *where* $S^{-1}f \in L^p(\partial\Omega)$.

It is worth pointing out that the nondegeneracy conditions in (4A.45) hold, in particular, if the system in question is of the form

$$L = \mathfrak{D}^*\mathfrak{D}, \tag{4A.49}$$

where

$$\mathfrak{D} \text{ is a first-order system with the unique continuation property,} \tag{4A.50}$$

in the sense that, if $u \in H^{1,2}(M)$ is such that $\mathfrak{D}u = 0$ on $M$ and $u$ vanishes on some nonempty open subset of $M$, then necessarily $u = 0$ everywhere on $M$. As a consequence, Theorem 4.3 applies to the Laplace–Beltrami operator on a Riemannian manifold, in which scenario the present well-posedness results complement those in [Mitrea and Taylor 1999].

*Proof of Theorem 4.3.* First, we shall show that

$$f \in L^2(\partial\Omega) \quad \text{and} \quad Sf = 0 \quad \Longrightarrow \quad f = 0. \tag{4A.51}$$

Suppose $f$ is as in the left-hand side of (4A.51) and set $u := \mathcal{S}f$ in $M \setminus \partial\Omega$. In light of (4A.43), the hypothesis (4A.45) then yields $u = 0$ both in $\Omega_+$ and in $\Omega_-$. Recall that $L$ is as in (4A.5)–(4A.6) and set (with $\nu = (\nu_i)_i$ denoting the outward unit conormal to $\Omega$)

$$\Xi_\pm f := \sum_{i,j} \nu_i A^{ij}(\partial_j \mathcal{S}f)\big|_{\partial\Omega_\pm}^{\text{n.t.}}. \tag{4A.52}$$

Then, on the one hand, the jump formulas from [Mitrea et al. 2001, Theorem 2.9, p. 21] yield

$$\Xi_\pm f = \left(\mp\tfrac{1}{2}I + K^*\right)f, \tag{4A.53}$$

where $K^*$ is a principal value singular integral operator on $\partial\Omega$ and $I$ is the identity. As such, we have the jump relation

$$f = \Xi_- f - \Xi_+ f. \tag{4A.54}$$

On the other hand, clearly $u = \mathscr{S}f = 0$ on $\Omega_+ \cup \Omega_-$ implies $\Xi_\pm f = 0$. We conclude that $f = 0$, finishing the proof of (4A.51).

In turn, (4A.51), Proposition 4.2, and Proposition 4.1 imply that, for each $p \in (1,\infty)$, the operator $S$ is an isomorphism in (4A.12) and (4A.13). This proves the claims in (4A.46). With these in hand, the fact that the Dirichlet and regularity boundary value problems (4A.47)–(4A.48) may be solved in the form $u = \mathscr{S}(S^{-1}f)$ follows from (4A.41)–(4A.42).

Turning to the uniqueness part, it suffices to show that any solution $u$ of the homogeneous version of the Dirichlet problem (4A.47) vanishes identically in $\Omega$. To this end, we introduce the Green function

$$G(x,y) := \Gamma(x,y) - \mathscr{S}\big[S^{-1}(E(x,\cdot)|_{\partial\Omega})\big](y), \quad (x,y) \in \Omega\times\Omega\setminus\mathrm{diag}, \tag{4A.55}$$

where the intervening single layer potential operators are associated with $L^*$. For each fixed $x \in \Omega$, the function $E(x,\cdot)|_{\partial\Omega}$ belongs to $H^{1,q}(\partial\Omega)$ for any $q \in (1,\infty)$. Thus, on account of (4A.46) we see that $G(x,y)$ is well defined. To proceed, consider a sequence of Lipschitz subdomains $\Omega_j$ of $\Omega$ such that $\Omega_j \nearrow \Omega$ as $j \to \infty$ as in [Mitrea and Taylor 1999, Appendix A]; in particular, their Lipschitz character is controlled uniformly in $j$. Let $G_j$ stand for the Green function corresponding to $\Omega_j$. By construction, $G_j(x,\cdot)|_{\partial\Omega_j} = 0$ and we claim that, for each $q \in (1,\infty)$, there exists a constant $C_q \in (0,\infty)$ with the property that

$$\sup_{j\in\mathbb{N}} \|\mathscr{N}_j(\nabla_2 G_j(x,\cdot))\|_{L^q(\partial\Omega_j)} \le C_q. \tag{4A.56}$$

This follows from the fact that if $S_j$ denotes the single layer constructed in relation to $\partial\Omega_j$ then, for each $q \in (1,\infty)$, the operator norm of $S_j^{-1} : H^{1,q}(\partial\Omega_j) \to L^q(\partial\Omega_j)$ is uniformly bounded in $j$. In turn, this is seen from (4A.18) and reasoning by contradiction.

For each $j \in \mathbb{N}$ let $\sigma_j$ denote the surface measure on $\partial\Omega_j$. Integrations by parts against these Green functions give that, if $u$ solves the homogeneous version of the Dirichlet problem (4A.47) and if $x \in \Omega$ is an arbitrary fixed point, then for $j \in \mathbb{N}$ sufficiently large we have

$$\begin{aligned} |u(x)| &= \left|\int_{\Omega_j} \langle (L^* G_j(x,\cdot))(y), u(y)\rangle \, d\,\mathrm{Vol}(y)\right| \\ &= \int_{\partial\Omega_j} O\big(|u|\cdot|\nabla_2 G_j(x,\cdot)|\big)\, d\sigma_j \\ &\le C\|u\|_{L^p(\partial\Omega_j)}, \end{aligned} \tag{4A.57}$$

where the last step utilizes Hölder's inequality and (4A.56). Because $\|u\|_{L^p(\partial\Omega_j)} \to 0$ by Lebesgue's dominated convergence theorem (and the manner in which $\Omega_j \nearrow \Omega$ as $j \to \infty$), we ultimately obtain $u(x) = 0$. Given that $x \in \Omega$ was arbitrary, the desired uniqueness statement follows.

Note that for the proof of uniqueness we could have avoided using the approximating family $\Omega_j \nearrow \Omega$ and, instead, worked directly with the Green function for $L^*$ constructed as in (4A.55), by reasoning as in the proof of [Hofmann et al. 2010, Theorem 7.2, p. 2831] as carried out in Step 3 on pp. 2832–2837. □

In the last part of this section we discuss the Poisson problem for strongly elliptic systems with data in Sobolev–Besov spaces in Lipschitz domains with normal in vmo. Throughout, retain the setting of

Theorem 4.3. For starters, from (4A.46) and complex interpolation we deduce, with the help of [Fabes et al. 1998, Lemma 8.4], that

$$S : H^{s-1,p}(\partial\Omega) \longrightarrow H^{s,p}(\partial\Omega) \quad \text{is invertible for each } p \in (1,\infty) \text{ and } s \in [0,1]. \tag{4A.58}$$

With $B_s^{p,q}(\partial\Omega)$ for $p, q \in (0,\infty]$ and $0 \neq s \in (-1,1)$ denoting the scale of Besov spaces on $\partial\Omega$, real interpolation then also gives that

$$S : B_{s-1}^{p,q}(\partial\Omega) \longrightarrow B_s^{p,q}(\partial\Omega) \quad \text{is invertible for } p \in (1,\infty),\ q \in (0,\infty] \text{ and } s \in (0,1). \tag{4A.59}$$

Furthermore, the action of the single layer potential operator $\mathcal{S}$ on Sobolev–Besov spaces on Lipschitz domains has been studied in [Mitrea and Taylor 2000]. The emphasis there is on the Hodge–Laplacian but the approach (which utilizes size estimates for the integral kernel and its derivatives) is general enough to work in the present setting. Indeed, the mapping properties from [Mitrea and Taylor 2000, Lemmas 7.2–7.3] are directly applicable here. They imply that if $B_s^{p,q}(\Omega)$ for $p, q \in (0,\infty]$ and $s \in \mathbb{R}$ stands for the scale of Besov spaces in $\Omega$, the single layer operator induces well-defined and bounded linear mappings in the following contexts:

$$\mathcal{S} : B_{-s}^{p,p}(\partial\Omega) \longrightarrow B_{1+\frac{1}{p}-s}^{p,p}(\Omega) \qquad \text{for } 1 \le p \le \infty \text{ and } 0 < s < 1, \tag{4A.60}$$

$$\mathcal{S} : B_{-s}^{p,p}(\partial\Omega) \longrightarrow H^{1+\frac{1}{p}-s,p}(\Omega) \qquad \text{for } 1 < p < \infty \text{ and } 0 < s < 1, \tag{4A.61}$$

$$\mathcal{S} : H^{-s,p}(\partial\Omega) \longrightarrow B_{1-s+\frac{1}{p}}^{p,\max\{p,2\}}(\Omega) \qquad \text{for } 1 < p < \infty \text{ and } 0 \le s \le 1. \tag{4A.62}$$

**Theorem 4.4.** *Suppose* $\Omega \subset M$ *is a* $\mathrm{Lip} \cap \mathrm{vmo}_1$ *domain and suppose* $L$ *is a second-order, strongly elliptic system of differential operators on* $M$ *as in* (4A.5)–(4A.6) *and satisfying* (4A.7) *and* (4A.45). *In addition, assume that* $L^*$, *the adjoint of* $L$, *also satisfies the nondegeneracy conditions in* (4A.45).

*Then, for any* $p \in (1,\infty)$ *and any* $s \in (0,1)$, *the Poisson problem with a Dirichlet boundary condition,*

$$\begin{cases} Lu = f \in H^{s+\frac{1}{p}-2,p}(\Omega), \\ \mathrm{Tr}\, u = g \in B_s^{p,p}(\partial\Omega), \\ u \in H^{s+\frac{1}{p},p}(\Omega), \end{cases} \tag{4A.63}$$

*has a unique solution.*

*Proof.* Extend the given $f \in H^{s+\frac{1}{p}-2,p}(\Omega)$ to some $\tilde{f} \in H^{s+\frac{1}{p}-2,p}(M)$, then consider

$$v := (L^{-1}\tilde{f})|_{\Omega} \in H^{s+\frac{1}{p},p}(\Omega). \tag{4A.64}$$

In particular, $h := \mathrm{Tr}\, v \in B_s^{p,p}(\partial\Omega)$ and a solution $u$ of the boundary value problem (4A.63) is given by

$$u := v - \mathcal{S}(S^{-1}(h-g)) \quad \text{in } \Omega, \tag{4A.65}$$

with $S^{-1}$ the inverse of the operator in (4A.59) (with $q = p$) and $\mathcal{S}$ considered as in (4A.61).

There remains to prove uniqueness. The existence result just established may be interpreted (taking $g = 0$) as the statement that

$$L : H_0^{s+\frac{1}{p},p}(\Omega) \longrightarrow H^{s+\frac{1}{p}-2,p}(\Omega) \quad \text{is surjective for each } p \in (1,\infty) \text{ and } s \in (0,1) \tag{4A.66}$$

in the class of operators $L$ described in the statement. Since the class in question is stable under taking adjoints, writing (4A.66) for $L^*$ then taking adjoints yields (after adjusting notation) that

$$L : H_0^{s+\frac{1}{p},p}(\Omega) \longrightarrow H^{s+\frac{1}{p}-2,p}(\Omega) \quad \text{is injective for each } p \in (1,\infty) \text{ and } s \in (0,1). \tag{4A.67}$$

With this in hand, the fact that any null solution of (4A.63) necessarily vanishes identically in $\Omega$ readily follows. This completes the proof of the theorem. □

**4B.** ***Oblique derivative problems.*** To start, let $\Omega \subset \mathbb{R}^n$ be a bounded, regular SKT domain, so its unit normal field $\nu$ belongs to $\mathrm{vmo}(\partial\Omega)$. We have tangential vector fields

$$\partial_{\tau_{jk}} = \nu_k \partial_j - \nu_j \partial_k, \quad 1 \le j,k \le n \tag{4B.1}$$

(see [Hofmann et al. 2010, Section 3.6]).

Let $\xi_{jk}$, $1 \le j, k \le n$, be real-valued functions on $\partial\Omega$ and define the tangential vector field

$$X := \sum_{j,k=1}^{n} \xi_{jk} \partial_{\tau_{jk}}. \tag{4B.2}$$

Assume that for each $j, k \in \{1, \dots, n\}$ we have

$$\xi_{jk}\nu_j,\ \xi_{jk}\nu_k \in \mathrm{vmo}(\partial\Omega) \cap L^\infty(\partial\Omega). \tag{4B.3}$$

Given $p \in (1,\infty)$, the goal here is to study the oblique derivative problem

$$\Delta u = 0 \quad \text{on } \Omega, \quad (\partial_\nu + X)u = f \quad \text{on } \partial\Omega, \quad \mathcal{N}u,\ \mathcal{N}(\nabla u) \in L^p(\partial\Omega), \tag{4B.4}$$

where $f \in L^p(\partial\Omega)$ is given. Above, $\partial_\nu u$ and $Xu$ are understood, respectively, as

$$\partial_\nu u := \sum_{j=1}^{n} \nu_j \big((\partial_j u)\big|_{\partial\Omega}^{\text{n.t.}}\big) \quad \text{and} \quad Xu := \sum_{j,k=1}^{n} \xi_{jk} \partial_{\tau_{jk}} \big(u\big|_{\partial\Omega}^{\text{n.t.}}\big). \tag{4B.5}$$

We look for a solution of (4B.4) in the form

$$u := \mathcal{S}g \quad \text{in } \Omega, \tag{4B.6}$$

where $g \in L^p(\partial\Omega)$ is yet to be determined and $\mathcal{S}$ is the harmonic single layer potential operator associated with $\Omega$. That is,

$$\mathcal{S}g(x) := \int_{\partial\Omega} E(x-y)g(y)\, d\sigma(y), \quad x \in \Omega, \tag{4B.7}$$

with $E$ denoting the standard fundamental solution for the Laplacian in $\mathbb{R}^n$, i.e., for all $x \in \mathbb{R}^n \setminus \{0\}$,

$$E(x) := \begin{cases} |x|^{2-n}/(\omega_{n-1}(2-n)) & \text{if } n \ge 3, \\ \frac{1}{2\pi} \ln|x| & \text{if } n = 2, \end{cases} \tag{4B.8}$$

where $\omega_{n-1}$ is the surface measure of the unit sphere $S^{n-1}$ in $\mathbb{R}^n$. As shown in [Hofmann et al. 2010, Section 4],

$$\partial_\nu \mathcal{S}g\big|_{\partial\Omega}^{\text{n.t.}} = \big(-\tfrac{1}{2}I + K^*\big)g, \tag{4B.9}$$

where

$$K^*: L^p(\partial\Omega) \to L^p(\partial\Omega) \quad \text{is compact for every } p \in (1,\infty). \tag{4B.10}$$

Meanwhile,

$$X(\mathscr{S}g) = Cg := \sum_{j,k}(A_{jk}g - B_{jk}g) \quad \text{on } \partial\Omega, \tag{4B.11}$$

where

$$A_{jk}g(x) := \mathrm{PV}\int_{\partial\Omega} a_{jk}(x)\partial_j E(x-y)g(y)\,d\sigma(y), \quad x \in \partial\Omega, \tag{4B.12}$$

and

$$B_{jk}g(x) := \mathrm{PV}\int_{\partial\Omega} b_{jk}(x)\partial_k E(x-y)g(y)\,d\sigma(y), \quad x \in \partial\Omega, \tag{4B.13}$$

with

$$a_{jk}(x) := \xi_{jk}(x)\nu_k(x), \quad b_{jk}(x) := \xi_{jk}(x)\nu_j(x). \tag{4B.14}$$

The following provides a key to the study of (4B.4):

**Lemma 4.5.** *If $\Omega \subset \mathbb{R}^n$ is a bounded, regular* SKT *domain and* (4B.3) *holds, then*

$$A_{jk} + A_{jk}^* \quad \textit{and} \quad B_{jk} + B_{jk}^* \quad \textit{are compact on } L^p(\partial\Omega) \textit{ for all } p \in (1,\infty). \tag{4B.15}$$

*Proof.* For each $j \in \{1,\dots,n\}$,

$$F_j g(x) := \mathrm{PV}\int_{\partial\Omega} \partial_j E(x-y)g(y)\,d\sigma(y), \quad x \in \partial\Omega, \tag{4B.16}$$

defines an operator of Calderón–Zygmund type that is bounded on $L^p(\partial\Omega)$ for all $p \in (1,\infty)$, since $\Omega$ is a UR domain. Then

$$A_{jk} + A_{jk}^* = [a_{jk}, F_j], \quad B_{jk} + B_{jk}^* = [b_{jk}, F_k], \tag{4B.17}$$

so (4B.15) follows from a general commutator estimate of Coifman–Rochberg–Weiss-type (see [Hofmann et al. 2010, Section 2.4]), since $a_{jk}, b_{jk} \in \mathrm{vmo}(\partial\Omega)$. □

In light of (4B.9) and (4B.11), solving the oblique derivative boundary value problem (4B.4) via the single layer representation (4B.6) is equivalent to finding a function $g \in L^p(\partial\Omega)$ satisfying

$$\left(-\tfrac{1}{2}I + C + K^*\right)g = f. \tag{4B.18}$$

In this regard, the following Fredholmness result is particularly relevant.

**Proposition 4.6.** *If $\Omega$ is bounded, regular* SKT *domain in $\mathbb{R}^n$ and if* (4B.3) *holds, then*

$$-\tfrac{1}{2}I + C + K^* : L^2(\partial\Omega) \to L^2(\partial\Omega) \quad \textit{is Fredholm of index } 0. \tag{4B.19}$$

*Proof.* By Lemma 4.5, we can write $C + K^* = \widetilde{C} + K_2$, where

$$\widetilde{C}^* := -\widetilde{C} \quad \text{and} \quad K_2 \text{ is a compact operator on } L^p(\partial\Omega) \text{ for all } p \in (1,\infty) \tag{4B.20}$$

Then, for $g \in L^2(\partial\Omega)$,

$$\Re\left(\left(-\tfrac{1}{2}I + \widetilde{C}\right)g, g\right) = -\tfrac{1}{2}\|g\|^2_{L^2(\partial\Omega)}, \tag{4B.21}$$

which, in turn, shows that

$$-\tfrac{1}{2}I + \widetilde{C} \quad \text{is invertible on } L^2(\partial\Omega). \tag{4B.22}$$

Since the operator in (4B.19) is a compact perturbation of this, the desired conclusion follows. □

**Corollary 4.7.** *In the setting of Proposition 4.6, there exists* $\varepsilon > 0$ *such that*

$$-\tfrac{1}{2}I + C + K^* : L^p(\partial\Omega) \to L^p(\partial\Omega) \quad \textit{is Fredholm of index } 0 \tag{4B.23}$$

*whenever* $|p-2| < \varepsilon$.

*Proof.* For $p$ close to 2, that

$$-\tfrac{1}{2}I + \widetilde{C} : L^p(\partial\Omega) \to L^p(\partial\Omega) \quad \text{is invertible} \tag{4B.24}$$

follows from (4B.22) and the stability results in [Šneĭberg 1974] (see also [Kalton and Mitrea 1998]). Meanwhile, the operator in (4B.23) is a compact perturbation of that in (4B.24) for all $p \in (1, \infty)$. □

In the context of Corollary 4.7, one wonders whether (4B.23) holds for all $p \in (1,\infty)$. We show that it does hold if $\Omega$ is a bounded $\mathrm{Lip} \cap \mathrm{vmo}_1$ domain in $\mathbb{R}^n$:

**Proposition 4.8.** *If* $\Omega$ *is a bounded* $\mathrm{Lip} \cap \mathrm{vmo}_1$ *domain in* $\mathbb{R}^n$ *and if* (4B.3) *holds, then the Fredholmness result* (4B.23) *is true for all* $p \in (1,\infty)$.

*Proof.* For starters, we note that, since (4B.3) and (4B.14) imply that $a_{jk}$, $b_{jk} \in \mathrm{vmo}(\partial\Omega)$, it follows from Lemma E.1 that $a_{jk} \circ \phi$, $b_{jk} \circ \phi \in \mathrm{vmo}(U)$ whenever $\phi : U \to \partial\Omega$ is a coordinate chart for $\partial\Omega$ (in the sense of Definition E.3). Keeping this in mind it follows that, in the present setting, the operator $C$ defined by (4B.11) belongs to $\mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$, and (4B.20) implies that its principal symbol is purely imaginary. Hence, for each $s \in \mathbb{R}$, $F_s := -\frac{1}{2}I + sC$ is an elliptic operator in $\mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$. Thus, these operators $F_s$ are all Fredholm on $L^p(\partial\Omega)$ and all have index independent of $s$. Clearly, $F_0$ has index zero, hence so does $F_1$, and the desired conclusion follows. □

We are now ready to state our main Fredholm solvability result for the oblique derivative problem. This builds on the earlier work of Calderón [1985]. Other extensions in the Euclidean setting are in [Kenig and Pipher 1988; Pipher 1987]; see also [Mitrea et al. ≥ 2015] for some recent refinements in the two-dimensional setting. For Lipschitz domains on manifolds see [Mitrea and Taylor 1999].

**Theorem 4.9.** *Let* $\Omega$ *is a bounded* $\mathrm{Lip} \cap \mathrm{vmo}_1$ *domain in* $\mathbb{R}^n$ *with outward unit normal* $\nu$. *Assume that* (4B.3) *holds and define the tangential vector field* $X$ *as in* (4B.2). *Finally, fix* $p \in (1,\infty)$.

*Then, for any boundary datum* $f \in L^p(\partial\Omega)$ *satisfying finitely many* (*necessary*) *linear conditions, the oblique derivative problem* (4B.4) *has a solution. Moreover, such a solution is unique modulo a finite-dimensional linear space, whose dimension coincides with the number of linearly independent constraints required for the boundary data.*

*Hence, the oblique derivative problem* (4B.4) *is Fredholm solvable with index zero.*

*Proof.* Fatou results in Lipschitz domains give that

$$\Delta u = 0 \text{ on } \Omega \text{ and } \mathcal{N}u,\ \mathcal{N}(\nabla u) \in L^p(\partial\Omega) \quad \Longrightarrow \quad u\big|^{\text{n.t.}}_{\partial\Omega} \text{ exists and } u\big|^{\text{n.t.}}_{\partial\Omega} \in H^{1,p}(\partial\Omega). \tag{4B.25}$$

Going further, from (4B.25) and the well-posedness of the $L^p$ regularity problem for the Laplacian in bounded Lip $\cap$ vmo$_1$ domains established in Theorem 4.3, it follows that

$$u = 0 \text{ on } \Omega \text{ and } \mathcal{N}u,\ \mathcal{N}(\nabla u) \in L^p(\partial\Omega) \quad \Longrightarrow \quad u = \mathcal{S}g \text{ in } \Omega \text{ for some (unique) } g \in L^p(\partial\Omega). \tag{4B.26}$$

In turn, from (4B.26) we deduce that, if the boundary datum $f \in L^p(\partial\Omega)$ is such that the oblique derivative problem (4B.4) has a solution $u$, then there exists a (unique) function $g \in L^p(\partial\Omega)$ with the property that

$$f = (\partial_\nu + X)u = (\partial_\nu + X)(\mathcal{S}g) = \left(-\tfrac{1}{2}I + C + K^*\right)g. \tag{4B.27}$$

This analysis shows that the oblique derivative problem (4B.4) is solvable precisely for boundary data $f$ belonging to the image of the operator $-\frac{1}{2}I + C + K^*$ on $L^p(\partial\Omega)$. By Proposition 4.8, this is a closed subspace of $L^p(\partial\Omega)$ of finite codimension. The above analysis also shows that the space of null solutions for the oblique derivative problem (4B.4) is isomorphic to the kernel of the operator $-\frac{1}{2}I + C + K^*$ on $L^p(\partial\Omega)$. Again, by Proposition 4.8, this is a finite-dimensional subspace of $L^p(\partial\Omega)$. Moreover, since the operator in question has index zero, we conclude that the number of (necessary) linear conditions which the boundary data must satisfy coincides with the dimension of the space of null solutions. Hence, the problem in question is Fredholm solvable with index zero. □

**4C. *Regular boundary problems for first-order elliptic systems.*** Suppose $\Omega \subset M$ a Lip $\cap$ vmo$_1$ domain and let $\mathfrak{D}$ be a first-order elliptic differential operator on $M$. It is permissible that $\mathfrak{D}$ acts on sections of a vector bundle $E \to M$. In local coordinates, assume that

$$\mathfrak{D}u(x) = \sum_j A_j(x)\partial_j u(x) + B(x)u(x), \quad \text{where } A_j \in \mathscr{C}^2,\ B \in \mathscr{C}^1. \tag{4C.1}$$

As in Section 3E (see especially Remark 3.9), we associate to $\mathfrak{D}$ a Cauchy integral $\mathscr{C}_{\mathfrak{D}}$ and a projection $P_{\mathfrak{D}}$, which is an element of $\mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$ in local graph coordinates.

When $\Omega$ is smooth, there is a well-established theory of regular boundary problems associated to $\mathfrak{D}$ (though sometimes regular boundary conditions do not exist). We want to investigate the situation where $\Omega \subset M$ is a Lip $\cap$ vmo$_1$ domain.

Let $F \to \partial^*\Omega$ be an $L^\infty \cap$ vmo vector bundle of rank $k$, so $F$ is locally trivializable to $\mathbb{C}^k \times \mathbb{O}$ with transition matrices in $L^\infty \cap$ vmo. Let

$$B : L^p(\partial\Omega, E) \longrightarrow L^p(\partial\Omega, F) \tag{4C.2}$$

be an operator that, in local graph coordinates and local trivializations of $E$ and $F$, satisfies

$$B \in \mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}. \tag{4C.3}$$

We can use analogues of (3A.6)–(3A.10) to define

$$\sigma_B(x,\xi) : E_x \longrightarrow F_x \tag{4C.4}$$

for almost all $(x,\xi) \in T^*\partial^*\Omega \setminus 0$. Extending the setup used when $\partial\Omega$ is smooth, we propose the following criterion for regularity:

$$\sigma_B(x,\xi) : \sigma_{P_{\mathcal{D}}}(x,\xi)E_x \longrightarrow F_x \quad \text{is an isomorphism for a.e. } (x,\xi) \in T^*\partial^*\Omega \setminus 0 \tag{4C.5}$$

and there exists $C > 0$ such that, for almost all $(x,\xi) \in T^*\partial^*\Omega \setminus 0$,

$$v \in E_x,\ \sigma_{P_{\mathcal{D}}} v = v \quad \Longrightarrow \quad \|\sigma_B(x,\xi)v\| \geq C\,\|v\|. \tag{4C.6}$$

Note that (4C.5)–(4C.6) is equivalent to (4C.6) alone plus

$$\dim \sigma_{P_{\mathcal{D}}}(x,\xi)E_x = \dim F_x. \tag{4C.7}$$

Also, $\sigma_{P_{\mathcal{D}}}(x,-\xi) = I - \sigma_{P_{\mathcal{D}}}(x,\xi)$, so, if $\dim \partial\Omega \geq 2$, the left-hand side of (4C.7) is equal to $\frac{1}{2}\dim E_x$.

Here is our basic Fredholm result:

**Proposition 4.10.** *Assume $\Omega \subset M$ is a* $\mathrm{Lip} \cap \mathrm{vmo}_1$ *domain and suppose $\mathcal{D} : E \to E$ is a first-order elliptic differential operator as in* (4C.1). *Under the hypotheses* (4C.5)–(4C.6), *the operator*

$$B : P_{\mathcal{D}}L^p(\partial\Omega, E) \longrightarrow L^p(\partial\Omega, F) \quad \textit{is Fredholm} \tag{4C.8}$$

*for each $p \in (1,\infty)$.*

*Proof.* The hypotheses imply that $\sigma_{BP_{\mathcal{D}}}(x,\xi) : E_x \to F_x$ is surjective for almost every $(x,\xi)$, and furthermore

$$BP_{\mathcal{D}}P_{\mathcal{D}}^*B^* \in \mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}} \quad \text{is elliptic.} \tag{4C.9}$$

Hence $B$ has a right Fredholm inverse, so $B$ in (4C.8) has closed range of finite codimension. Also,

$$f \in P_{\mathcal{D}}L^p(\partial\Omega, E), \quad Bf = 0 \tag{4C.10}$$

is equivalent to

$$\begin{pmatrix} B \\ I - P_{\mathcal{D}} \end{pmatrix} f = 0, \quad f \in L^p(\partial\Omega, E), \tag{4C.11}$$

and the operator on the left-hand side of (4C.11) (call it $Q$) is an element of $\mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$ (mod compacts) with symbol $\sigma_Q(x,\xi)$ injective, and furthermore

$$Q^*Q \in \mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}} \quad \text{is elliptic.} \tag{4C.12}$$

Thus, $Q$ has a left Fredholm inverse, so its null space in $L^p(\partial\Omega, E)$ is finite dimensional. This proves (4C.8). □

**Theorem 4.11.** *Under the hypotheses of Proposition 4.10, the boundary problem*

$$\begin{cases} \mathcal{D}u = 0 \quad \textit{on } \Omega, \\ \mathcal{N}u \in L^p(\partial\Omega), \\ Bu = f \in L^p(\partial\Omega, F), \end{cases} \tag{4C.13}$$

*is Fredholm solvable for each $p \in (1,\infty)$.*

*Proof.* To restate the result, consider

$$\mathscr{H}^p(\Omega,\mathfrak{D}) := \{u \in \mathscr{C}^1(\Omega, E) : \mathfrak{D}u = 0 \text{ on } \Omega,\ \mathcal{N}u \in L^p(\partial\Omega)\}. \tag{4C.14}$$

In [Mitrea et al. ≥ 2015], a Fatou-type lemma is established showing that each $u \in \mathscr{H}^p(\Omega,\mathfrak{D})$ has a boundary trace provided $\Omega$ is a regular SKT domain. From there, results in [Mitrea et al. 2015, §3.1] (see also [Mitrea et al. ≥ 2015]) imply that the boundary trace yields an isomorphism

$$\tau : \mathscr{H}^p(\Omega,\mathfrak{D}) \xrightarrow{\sim} P_{\mathfrak{D}} L^p(\partial\Omega, E) \tag{4C.15}$$

for $p \in (1,\infty)$. The assertion of Theorem 4.11 is that, if $B$ satisfies the hypotheses of Proposition 4.10, then

$$B \circ \tau : \mathscr{H}^p(\Omega,\mathfrak{D}) \longrightarrow L^p(\partial\Omega, F) \quad \text{is Fredholm.} \tag{4C.16}$$

In light of (4C.15), the result (4C.16) is equivalent to (4C.8). □

As we have mentioned, sometimes $\mathfrak{D}$ has no boundary conditions of the form (4C.2)–(4C.4) satisfying the regularity condition (4C.5)–(4C.6). In Section 4D we shall give important examples (well known for smooth boundaries) of regular boundary conditions for $\mathfrak{D} = d + d^*$ acting on differential forms. Here, we record a simple example (also well known) of a first-order elliptic operator with no such regular boundary condition. Namely, we take a bounded $\Omega \subset \mathbb{R}^2$ (possibly with smooth boundary) and set

$$\mathfrak{D} = \frac{\partial}{\partial x_1} + i\frac{\partial}{\partial x_2} \tag{4C.17}$$

acting on complex-valued $u$, so $E_x = \mathbb{C}$. In this case, $\sigma_{\mathfrak{D}}(x,\xi)u = i(\xi_1 + i\xi_2)u$, or, if we identify $\xi = (\xi_1,\xi_2) \in \mathbb{R}^2$ with $\xi_1 + i\xi_2 \in \mathbb{C}$, $\sigma_{\mathfrak{D}}(x,\xi)u = i\xi u$; hence,

$$M(x,\xi) = \nu^{-1}\xi. \tag{4C.18}$$

Now $\xi$ runs over the orthogonal complement of $\nu$, i.e., over real multiples of $i\nu$. We have

$$M(x,i\nu) = i, \quad M(x,-i\nu) = -i, \tag{4C.19}$$

so

$$P_+(M(x,i\nu)) = I, \quad P_+(M(x,-i\nu)) = 0. \tag{4C.20}$$

Since the ranges have different dimensions, there is no way to achieve (4C.5) for both $\xi = i\nu$ and $\xi = -i\nu$.

Returning to the setting of Proposition 4.10 and Theorem 4.11, we see from (4C.9) that the operator $B$ in (4C.8) has a right Fredholm inverse that is an element of $\mathrm{OP}(L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$, and that this operator is independent of $p \in (1,\infty)$. Since $B$ in (4C.8) is Fredholm, this right Fredholm inverse is also a left Fredholm inverse for each $p \in (1,\infty)$. Call it

$$H : L^p(\partial\Omega, F) \longrightarrow P_{\mathfrak{D}} L^p(\partial\Omega, E). \tag{4C.21}$$

Using this observation, we can prove the following:

**Proposition 4.12.** *Under the hypotheses of Proposition 4.10, the index of $B$ in* (4C.8), *and hence the index of $B \circ \tau$ in* (4C.11), *is independent of $p$.*

*Proof.* Setting $V_p = P_{\mathscr{D}} L^p(\partial\Omega, E)$ and $W_p = L^p(\partial\Omega, F)$, our setup is

$$B : V_p \to W_p, \quad H : W_p \to V_p \quad \text{Fredholm inverses} \tag{4C.22}$$

for $p \in (1, \infty)$. Setting

$$\operatorname{Ker}_p B := \{f \in V_p : Bf = 0\}, \quad \operatorname{Coker}_p B := \{\varphi \in W_p' : B^*\varphi = 0\}, \tag{4C.23}$$

we have

$$\begin{aligned} 1 < p < q < \infty \quad &\Longrightarrow \quad \operatorname{Ker}_q B \subset \operatorname{Ker}_p B,\ \operatorname{Coker}_p B \subset \operatorname{Coker}_q B \\ &\Longrightarrow \quad \operatorname{index}_q B \le \operatorname{index}_p B. \end{aligned} \tag{4C.24}$$

The same argument gives

$$1 < p < q < \infty \quad \Longrightarrow \quad \operatorname{index}_q H \le \operatorname{index}_p H, \tag{4C.25}$$

and, since $\operatorname{index}_p B = -\operatorname{index}_p H$, we have

$$1 < p, q < \infty \quad \Longrightarrow \quad \operatorname{index}_p B = \operatorname{index}_q B, \tag{4C.26}$$

as desired. □

The results (4C.24)–(4C.26) also imply that

$$1 < p, q < \infty \quad \Longrightarrow \quad \operatorname{Ker}_p B = \operatorname{Ker}_q B. \tag{4C.27}$$

Let us set

$$\mathscr{H}^p_B(\Omega) := \{u \in \mathscr{H}^p(\Omega, \mathscr{D}) : Bu = 0 \text{ on } \partial\Omega\}. \tag{4C.28}$$

Then, the isomorphism (4C.15) gives

$$\tau : \mathscr{H}^p_B(\Omega) \xrightarrow{\sim} P_{\mathscr{D}} L^p(\partial\Omega, E) \cap \operatorname{Ker} B = \operatorname{Ker}_p B. \tag{4C.29}$$

Thus (4C.27) yields the following:

**Corollary 4.13.** *Under the hypotheses of Proposition 4.10, the space $\mathscr{H}^p_B(\Omega)$ defined in* (4C.28) *is independent of $p \in (1, \infty)$.*

**4D.** ***Absolute and relative boundary conditions for the Hodge–Dirac operator.*** Let $\Omega$ be a $\mathrm{Lip} \cap \mathrm{vmo}_1$ domain in a smooth Riemannian manifold $M$. Let $d$ denote the exterior derivative on $M$, denote by $\delta = d^*$ its adjoint, then define the Hodge–Dirac operator

$$\mathscr{D} := d + \delta \tag{4D.1}$$

acting on sections of

$$E := \Lambda^*_{\mathbb{C}} M. \tag{4D.2}$$

We take $F := \Lambda^*_{\mathbb{C}} \partial^* \Omega$ and

$$Bu := j^* u, \tag{4D.3}$$

the pull-back associated to $j : \partial^*\Omega \hookrightarrow M$. We claim that $(\mathscr{D}, B)$ given by (4D.1) and (4D.3) satisfy the regularity conditions (4C.5)–(4C.6), i.e.,

$$\sigma_B(x,\xi) : P_+(M(x,\xi))E_x \longrightarrow F_x \quad \text{isomorphically} \tag{4D.4}$$

for almost every $(x,\xi) \in T^*\partial^*\Omega \setminus 0$, with a uniform lower bound of the form

$$v \in F_x, \quad P_+(M(x,\xi))v = v \quad \Longrightarrow \quad \|\sigma_B(x,\xi)v\| \geq C\|v\|. \tag{4D.5}$$

Recall that $P_+(M(x,\xi))$ is the projection of $E_x$ onto the span of the generalized eigenvectors of $M(x,\xi)$ associated with eigenvalues with positive imaginary part, annihilating those associated with eigenvalues with negative imaginary part, where

$$M(x,\xi) = \sigma_{\mathscr{D}}(x,\nu)^{-1}\sigma_{\mathscr{D}}(x,\xi). \tag{4D.6}$$

Checking (4D.4)–(4D.5) is a purely algebraic problem, and to do this algebra it suffices to take the case

$$M := \mathbb{R}^{n+1}, \quad \Omega := \{x \in \mathbb{R}^{n+1} : x_{n+1} < 0\}. \tag{4D.7}$$

Let $\wedge$ and $\vee$ denote, respectively, the exterior and interior product of forms. The following calculation shows that we have symbols independent of $x$:

$$\sigma_{\mathscr{D}}(\xi)u = i\xi \wedge u - i\xi \vee u, \quad \sigma_B(\xi)u = j^*u = \nu \vee (\nu \wedge u). \tag{4D.8}$$

In addition, $\sigma_{\mathscr{D}}(\xi)^2 = |\xi|^2 I$ and, more generally, the anticommutator identity holds:

$$\sigma_{\mathscr{D}}(\xi)\sigma_{\mathscr{D}}(\eta) + \sigma_{\mathscr{D}}(\eta)\sigma_{\mathscr{D}}(\xi) = 2\langle \xi, \eta\rangle I. \tag{4D.9}$$

Consequently, $\sigma_{\mathscr{D}}(\nu)^{-1} = \sigma_{\mathscr{D}}(\nu)$ and, for $\xi \in T^*\partial\Omega \setminus 0$,

$$M(\xi) = \sigma_{\mathscr{D}}(\nu)\sigma_{\mathscr{D}}(\xi) = -\sigma_{\mathscr{D}}(\xi)\sigma_{\mathscr{D}}(\nu); \tag{4D.10}$$

hence

$$M(\xi)^2 = -|\xi|^2 I, \tag{4D.11}$$

so

$$\operatorname{Spec} M(\xi) = \{i|\xi|, -i|\xi|\}. \tag{4D.12}$$

Note that if $\xi$, $\eta$ belong to $T^*\partial\Omega = \mathbb{R}^n$ and have the same length, then $M(\xi)$ and $M(\eta)$ are conjugate if $n \geq 2$, since then one can pass from $\xi$ to $\eta$ by an element of $\mathrm{SO}(n)$. On the other hand, $M(-\xi) = -M(\xi)$. It follows that

$$\dim P_+(M(\xi)) = \tfrac{1}{2}\dim E_x = \dim F_x \tag{4D.13}$$

for all $\xi \neq 0$. For $n = 1$, this can be checked by a simple direct calculation.

Having this, all we need to show to establish (4D.4)–(4D.5) is that

$$v \in \Lambda^*_{\mathbb{C}}\mathbb{R}^{n+1}, \quad \xi \in \mathbb{R}^n, \quad |\xi| = 1, \quad M(\xi)v = iv, \quad j^*v = 0 \tag{4D.14}$$

implies

$$v = 0. \tag{4D.15}$$

Indeed, (4D.14) implies

$$\sigma_{\mathfrak{D}}(\xi)v = i\sigma_{\mathfrak{D}}(\nu)v = -\nu \wedge v + \nu \vee v; \tag{4D.16}$$

hence, since $j^*v = 0$ forces $\nu \wedge v = 0$, we obtain

$$\sigma_{\mathfrak{D}}(\xi)v = \nu \vee v. \tag{4D.17}$$

Now the right-hand side of (4D.17) belongs to $\Lambda^*_{\mathbb{C}}\mathbb{R}^n$. But, if $\nu \wedge v = 0$ and $\xi \in \mathbb{R}^n \setminus 0$, the left-hand side of (4D.17) cannot belong to $\Lambda^*_{\mathbb{C}}\mathbb{R}^n$ unless it is zero. This implies $\sigma_{\mathfrak{D}}(\xi)v = 0$, and hence (4D.15) follows.

A similar argument applies if we replace $B$ in (4D.3) by

$$Bu = \nu \vee u\big|^{\text{n.t.}}_{\partial\Omega}. \tag{4D.18}$$

Then we need to show that

$$v \in \Lambda^*_{\mathbb{C}}\mathbb{R}^{n+1}, \quad \xi \in \mathbb{R}^n, \quad |\xi| = 1, \quad M(\xi)v = iv, \quad \nu \vee v = 0 \tag{4D.19}$$

implies (4D.15). Indeed, (4D.19) implies

$$\sigma_{\mathfrak{D}}(\xi)v = -\nu \wedge v. \tag{4D.20}$$

If $\nu \vee v = 0$ and $\xi \in \mathbb{R}^n \setminus 0$, one cannot factor out a $\nu$ on the left-hand side of (4D.20) unless this term vanishes, so again we get (4D.15).

The boundary condition (4D.3) is called the relative boundary condition for $d + \delta$, and (4D.18) is called the absolute boundary condition for $d + \delta$. The arguments above establish the following:

**Proposition 4.14.** *The absolute boundary condition* (4D.18) *and the relative boundary condition* (4D.3) *are each regular boundary conditions for the elliptic operator $d + \delta$. Consequently, specializing* (4C.28), *the spaces*

$$\begin{aligned} \mathcal{H}_A(\Omega) &:= \{u \in \mathcal{H}^p(\Omega, d+\delta) : \nu \vee u\big|^{\text{n.t.}}_{\partial\Omega} = 0\}, \\ \mathcal{H}_R(\Omega) &:= \{u \in \mathcal{H}^p(\Omega, d+\delta) : \nu \wedge u\big|^{\text{n.t.}}_{\partial\Omega} = 0\}, \end{aligned} \tag{4D.21}$$

*where $p \in (1, \infty)$ and, as in* (4C.14),

$$\mathcal{H}^p(\Omega, d+\delta) := \{u \in \mathscr{C}^1(\Omega, \Lambda^*_{\mathbb{C}}) : (d+\delta)u = 0 \text{ in } \Omega,\ \mathcal{N}u \in L^p(\partial\Omega)\}, \tag{4D.22}$$

*are finite dimensional. Furthermore, by Corollary 4.13 the spaces in* (4D.21) *are independent of $p \in (1, \infty)$.*

Here, $\Lambda^*_{\mathbb{C}} := \bigoplus_{\ell=0}^n \Lambda^\ell_{\mathbb{C}}$, where $n := \dim \Omega$. We also set

$$\Lambda^o_{\mathbb{C}} := \bigoplus_{\ell \text{ odd}} \Lambda^\ell_{\mathbb{C}}, \quad \Lambda^e_{\mathbb{C}} := \bigoplus_{\ell \text{ even}} \Lambda^\ell_{\mathbb{C}}, \tag{4D.23}$$

$$\mathcal{H}^p_\sigma(\Omega, d+\delta) := \mathcal{H}^p(\Omega, d+\delta) \cap \mathscr{C}^0(\Omega, \Lambda^\sigma_{\mathbb{C}}), \quad \sigma = o \text{ or } e, \tag{4D.24}$$

$$\mathcal{H}^\sigma_b(\Omega) := \mathcal{H}_b(\Omega) \cap \mathscr{C}^0(\Omega, \Lambda^\sigma_{\mathbb{C}}), \quad b = A \text{ or } R,\ \sigma = o \text{ or } e. \tag{4D.25}$$

Note that

$$\begin{aligned} d+\delta &: \mathscr{C}^1(\Omega, \Lambda^o_{\mathbb{C}}) \longrightarrow \mathscr{C}^0(\Omega, \Lambda^e_{\mathbb{C}}), \\ d+\delta &: \mathscr{C}^1(\Omega, \Lambda^e_{\mathbb{C}}) \longrightarrow \mathscr{C}^0(\Omega, \Lambda^o_{\mathbb{C}}), \end{aligned} \tag{4D.26}$$

so

$$\mathscr{H}^p(\Omega, d+\delta) = \mathscr{H}^p_e(\Omega, d+\delta) \oplus \mathscr{H}^p_o(\Omega, d+\delta), \tag{4D.27}$$

$$\mathscr{H}_b(\Omega) = \mathscr{H}^e_b(\Omega) \oplus \mathscr{H}^o_b(\Omega), \quad b = A \text{ or } R. \tag{4D.28}$$

In this vein, we wish to note that if we also consider

$$\widetilde{\mathscr{H}}_A(\Omega) := \{u \in \mathscr{H}^p(\Omega, d \oplus \delta) : \nu \vee u\big|^{\text{n.t.}}_{\partial\Omega} = 0\}, \tag{4D.29}$$

$$\widetilde{\mathscr{H}}_R(\Omega) := \{u \in \mathscr{H}^p(\Omega, d \oplus \delta) : \nu \wedge u\big|^{\text{n.t.}}_{\partial\Omega} = 0\}, \tag{4D.30}$$

where

$$\mathscr{H}^p(\Omega, d \oplus \delta) := \{u \in \mathscr{C}^1(\Omega, \Lambda^*_{\mathbb{C}}) : du = \delta u = 0 \text{ on } \Omega, \mathcal{N}u \in L^p(\partial\Omega)\}, \tag{4D.31}$$

then from [Mitrea 2001, Theorem 6.1] it follows that

$$\widetilde{\mathscr{H}}_A(\Omega) = \mathscr{H}_A(\Omega) \quad \text{and} \quad \widetilde{\mathscr{H}}_R(\Omega) = \mathscr{H}_R(\Omega). \tag{4D.32}$$

In more detail, (4D.32) was demonstrated for $p$ close to 2 in [Mitrea 2001] in the setting of a general Lipschitz domain. However, the independence of $\mathscr{H}_A(\Omega)$ and $\mathscr{H}_R(\Omega)$ from $p$, plus the obvious inclusions $\widetilde{\mathscr{H}}_A(\Omega) \subset \mathscr{H}_A(\Omega)$ and $\widetilde{\mathscr{H}}_R(\Omega) \subset \mathscr{H}_R(\Omega)$, imply that $\widetilde{\mathscr{H}}_A(\Omega)$ and $\widetilde{\mathscr{H}}_R(\Omega)$ are also independent of $p$.

## Auxiliary results

We collect here a number of auxiliary results that are useful in the body of the paper.

**Appendix A.** ***Spectral theory for the Dirichlet Laplacian.*** Specifically, fix an arbitrary bounded open set $\mathbb{O} \subseteq \mathbb{R}^n$ and, for any given $p \in (1, \infty)$ and $k \in \mathbb{Z}$, denote by $W^{k,p}(\mathbb{O})$ the standard $L^p$-based Sobolev space of smoothness order $k$. Also, let $\mathring{W}^{k,p}(\mathbb{O})$ be the closure of $\mathscr{C}^\infty_0(\mathbb{O})$ in $W^{k,p}(\mathbb{O})$.

Let $\Delta_D$ be the realization of the Laplacian with (homogeneous) Dirichlet boundary condition as an unbounded linear operator in the context of the Hilbert space $L^2(\mathbb{O})$, with domain

$$\text{Dom}(\Delta_D) := \{u \in \mathring{W}^{1,2}(\mathbb{O}) : \Delta u \in L^2(\mathbb{O})\}. \tag{A.1}$$

Then $-\Delta_D$ is a nonnegative self-adjoint operator mapping $\text{Dom}(\Delta_D)$ isomorphically onto $L^2(\mathbb{O})$, and its inverse

$$G_D := (-\Delta_D)^{-1} : L^2(\mathbb{O}) \longrightarrow L^2(\mathbb{O}) \tag{A.2}$$

is self-adjoint, nonnegative and compact. In particular, $-\Delta_D$ has a pure point spectrum

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_j \leq \lambda_{j+1} \leq \cdots \tag{A.3}$$

listed according to their (finite) multiplicities. See, for example, [Dautray and Lions 1990, p. 82].

Let us temporarily write $\lambda_j(\mathbb{O})$ in place of $\lambda_j$ in order to emphasize the dependence on the underlying domain $\mathbb{O}$. The classical Rayleigh–Ritz min–max principle asserts (see, e.g., [Dautray and Lions 1990, Theorem 10, p. 102]) that, for each $j \in \mathbb{N}$,

$$\lambda_j(\mathbb{O}) = \min_{\substack{V_j \subseteq \mathring{W}^{1,2}(\mathbb{O}) \\ \dim V_j = j}} \max_{u \in V_j \setminus \{0\}} \frac{\int_{\mathbb{O}} |\nabla u|^2}{\int_{\mathbb{O}} |u|^2}. \tag{A.4}$$

Assume now that $\widetilde{\mathbb{O}}$ is a bounded, open subset of $\mathbb{R}^n$ such that $\mathbb{O} \subseteq \widetilde{\mathbb{O}}$. Given that extension by zero is a well-defined, norm-preserving mapping from $\mathring{W}^{1,2}(\mathbb{O})$ into $\mathring{W}^{1,2}(\widetilde{\mathbb{O}})$, it readily follows from (A.4) that the following domain monotonicity property holds:

$$\lambda_j(\mathbb{O}) \geq \lambda_j(\widetilde{\mathbb{O}}) \quad \text{for all } j \in \mathbb{N}. \tag{A.5}$$

In this vein, let us also mention that each $\lambda_j(\mathbb{O})$ is invariant with respect to translations and rotations of $\mathbb{O}$, and one has the scaling property

$$\lambda_j(c\mathbb{O}) = c^{-2}\lambda_j(\mathbb{O}) \quad \text{for all } c \in (0,\infty),\ j \in \mathbb{N}. \tag{A.6}$$

Finally, pick a complete set of normalized eigenfunctions $\{\vartheta_j\}_{j\in\mathbb{N}} \subset L^2(\mathbb{O})$ for $-\Delta_D$. Thus,

$$\vartheta_j \in \mathring{W}^{1,2}(\mathbb{O}), \quad \|\vartheta_j\|_{L^2(\mathbb{O})} = 1 \quad \text{and} \quad -\Delta\vartheta_j = \lambda_j\vartheta_j \qquad \text{for each } j \in \mathbb{N}. \tag{A.7}$$

**Lemma A.1.** *Let $\mathbb{O}$ be a bounded, open subset of $\mathbb{R}^n$.*

*Then there exist $c_1, c_2 \in (0,\infty)$ depending only on $n$ and $\mathbb{O}$ such that*

$$c_1 j^{2/n} \leq \lambda_j \leq c_2 j^{2/n} \quad \textit{for each } j \in \mathbb{N}. \tag{A.8}$$

*Also, there exists $C_{\mathbb{O},n} \in (0,\infty)$ with the property that*

$$\|\vartheta_j\|_{L^\infty(\mathbb{O})} \leq C_{\mathbb{O},n} j^{1/2+2/n} \quad \textit{for each } j \in \mathbb{N}. \tag{A.9}$$

*Moreover, for each $j \in \mathbb{N}$ one has*

$$\vartheta_j \in \mathscr{C}^\infty_{loc}(\mathbb{O}) \tag{A.10}$$

*and, for every compact subset $K$ of $\mathbb{O}$ and every multi-index $\alpha \in \mathbb{N}_0^n$, there exists a constant $C_{\mathbb{O},K,\alpha} \in (0,\infty)$ with the property that*

$$\|\partial^\alpha\vartheta_j\|_{L^\infty(K)} \leq C_{\mathbb{O},K,\alpha} j^{1/2+2/n}. \tag{A.11}$$

*Proof.* When $\mathbb{O}$ is the cube $(0,1)^n$ in $\mathbb{R}^n$, the pure point spectrum of the Dirichlet Laplacian is given by

$$\{\lambda_j((0,1)^n)\}_{j\in\mathbb{N}} = \{4\pi^2(k_1^2 + \cdots + k_n^2) : k_i \in \mathbb{N},\ 1 \leq i \leq n\}, \tag{A.12}$$

an identification that takes into account multiplicities. From this one can deduce Weyl's asymptotic formula

$$\lambda_j((0,1)^n) \approx \frac{4\pi^2 j^{2/n}}{\pi^{n/2}\Gamma(n/2+1)}, \tag{A.13}$$

valid for large values of $j \in \mathbb{N}$, and the estimates in (A.8) follow in this scenario from (A.13). The general situation when $\mathbb{O}$ is an arbitrary bounded open set in $\mathbb{R}^n$ may then be handled based on the special case just treated and the comments in (A.5)–(A.6).

The operator $G_D$ in (A.2) is an integral operator whose kernel is the negative of the Green function for $\mathbb{O}$, i.e.,

$$G_D u(x) = -\int_{\mathbb{O}} G(x, y) u(y)\, dy, \quad x \in \mathbb{O}, \tag{A.14}$$

for each $u \in L^2(\mathbb{O})$. Since (see [Grüter and Widman 1982]) we have

$$|G(x, y)| \le \frac{C_n}{|x-y|^{n-2}}, \quad x,\ y \in \mathbb{O}, \tag{A.15}$$

(assuming $n > 2$; the case $n = 2$, when a logarithm is involved, is treated analogously), it follows that $G_D$ behaves like a fractional integral operator of order 2; hence (see [Stein 1970]),

$$G_D : L^p(\mathbb{O}) \longrightarrow L^q(\mathbb{O}) \quad \text{linearly and boundedly if} \quad \begin{cases} q < \infty \ \text{ and } \ 1/q \ge 1/p - 2/n, \text{ or} \\ q = \infty \ \text{ and } \ p > n/2. \end{cases} \tag{A.16}$$

Iterating, it follows that

$$(G_D)^k : L^2(\mathbb{O}) \longrightarrow L^\infty(\mathbb{O}) \quad \text{boundedly if } \ k > n/4. \tag{A.17}$$

On the other hand, for each fixed $j \in \mathbb{N}$, from (A.7) we have $\vartheta_j = \lambda_j G_D \vartheta_j$, which, inductively, implies $\vartheta_j = \lambda_j^k (G_D)^k \vartheta_j$ for each $k \in \mathbb{N}$. Consequently, if $k := [n/4] + 1$ then $k \in \mathbb{N}$ satisfies $k \in (n/4, n/4 + 1]$; hence, we may estimate

$$\begin{aligned} \|\vartheta_j\|_{L^\infty(\mathbb{O})} &= \|\lambda_j^k (G_D)^k \vartheta_j\|_{L^\infty(\mathbb{O})} \\ &\le \|(G_D)^k\|_{\mathscr{L}(L^2(\mathbb{O}), L^\infty(\mathbb{O}))} \lambda_j^k \|\vartheta_j\|_{L^2(\mathbb{O})} \\ &\le C_{\mathbb{O},n}\, j^{2k/n} \le C_{\mathbb{O},n}\, j^{1/2+2/n} \end{aligned} \tag{A.18}$$

by (A.17), (A.7) and (A.8). This proves (A.9).

Finally, (A.10)–(A.11) follow from (A.7), (A.9) and elliptic regularity. □

**Appendix B.** ***Truncating singular integrals.*** If $U \subseteq \mathbb{R}^n$, call $\Phi : U \to \mathbb{R}^m$ bi-Lipschitz if there exist $M_1$, $M_2$ with $0 < M_1 \le M_2 < \infty$ such that

$$M_1 |x - y| \le |\Phi(x) - \Phi(y)| \le M_2 |x - y| \quad \text{for all } \ x,\ y \in U. \tag{B.1}$$

When $U$ is an open set, it is known from [Rademacher 1919] that necessarily $m \ge n$, $\Phi$ is an open mapping, the Jacobian matrix $D\Phi = (\partial_k \Phi_j)_{1 \le j \le m, 1 \le k \le n}$ exists a.e. in $U$, and

$$\operatorname{rank} D\Phi(x) = n \quad \text{for a.e. } \ x \in U. \tag{B.2}$$

**Lemma B.1.** *Let $A : \mathbb{R}^n \to \mathbb{R}^m$ and $B : \mathbb{R}^n \to \mathbb{R}^{m'}$ be functions satisfying*

$$|A(x) - A(y)| \le M|x - y| \quad \textit{and} \tag{B.3}$$

$$M^{-1}|x - y| \le |B(x) - B(y)| \le M|x - y| \quad \textit{for all } \ x,\ y \in \mathbb{R}^n \tag{B.4}$$

*for some positive constant $M$. Also let $F:\mathbb{R}^m \to \mathbb{R}$ be an odd function of class $\mathscr{C}^1$. Finally, fix a point $x \in \mathbb{R}^n$ where both $DA(x)$, $DB(x)$ exist*, rank $DB(x) = n$ *and, for each $\varepsilon > 0$, consider*

$$
\begin{aligned}
U(\varepsilon) &:= \{y \in \mathbb{R}^n : 1 > |x-y| > \varepsilon\},\\
V(\varepsilon) &:= \{y \in \mathbb{R}^n : |DB(x)(x-y)| > \varepsilon,\ |x-y| < 1\},\\
W(\varepsilon) &:= \{y \in \mathbb{R}^n : |B(x)-B(y)| > \varepsilon,\ |x-y| < 1\}.
\end{aligned}
\tag{B.5}
$$

*Then, whenever any of the three limits*

$$\lim_{\varepsilon \searrow 0} \int_{U(\varepsilon)} \frac{1}{|x-y|^n} F\left(\frac{A(x)-A(y)}{|x-y|}\right) dy, \tag{B.6}$$

$$\lim_{\varepsilon \searrow 0} \int_{V(\varepsilon)} \frac{1}{|x-y|^n} F\left(\frac{A(x)-A(y)}{|x-y|}\right) dy, \tag{B.7}$$

$$\lim_{\varepsilon \searrow 0} \int_{W(\varepsilon)} \frac{1}{|x-y|^n} F\left(\frac{A(x)-A(y)}{|x-y|}\right) dy \tag{B.8}$$

*exists (in $\mathbb{R}$), it follows that all exist and are equal.*

*Proof.* Without loss of generality we can take $x = 0$ and assume that $A(0) = 0$, $B(0) = 0$. As a consequence of this normalization and (B.3), we have

$$\frac{|A(y)|}{|y|} \le M \quad \text{for all } \ y \in \mathbb{R}^n \setminus \{0\}. \tag{B.9}$$

The fact that $DA(0)$, $DB(0)$ exist implies that we can find a function $\eta:(0,\infty) \to [0,\infty)$ with the property that $\eta(t) \searrow 0$ as $t \searrow 0$ and

$$|B(y) - DB(0)y| + |A(y) - DA(0)y| \le |y|\eta(|y|) \quad \text{for all } \ y \in \mathbb{R}^n. \tag{B.10}$$

In particular,

$$
\begin{aligned}
|A(y) + A(-y)| &= |(A(y) - DA(0)y) + (A(-y) - DA(0)(-y))|\\
&\le |A(y) - DA(0)y| + |A(-y) - DA(0)(-y)|\\
&\le 2|y|\eta(|y|) \quad \text{for all } \ y \in \mathbb{R}^n.
\end{aligned}
\tag{B.11}
$$

Recall that the matrix $DB(0)$ is assumed to have rank $n$. Hence, $\|DB(0)\| > 0$ and, letting

$$\Delta(\varepsilon) := \left\{y \in \mathbb{R}^n : \varepsilon \ge |y| \ge \varepsilon/\|DB(0)\|\right\} \tag{B.12}$$

for each $\varepsilon > 0$,

$$V(\varepsilon) \setminus U(\varepsilon) \subseteq \Delta(\varepsilon) \quad \text{for all } \ \varepsilon > 0. \tag{B.13}$$

Observing that $U(\varepsilon)$ and $V(\varepsilon)$ are symmetric with respect to the origin, employing the properties of $F$ and $\eta$, and keeping in mind (B.10), (B.13), (B.11) and (B.9), we may use the mean value theorem in order to estimate the absolute value of the difference of the limits in (B.6) and (B.7) by

$$\begin{aligned}
\lim_{\varepsilon\searrow 0}&\left|\int_{V(\varepsilon)\setminus U(\varepsilon)}\frac{1}{|y|^n}F\Big(\frac{A(y)}{|y|}\Big)\,dy\right| \\
&=\lim_{\varepsilon\searrow 0}\frac{1}{2}\left|\int_{V(\varepsilon)\setminus U(\varepsilon)}\frac{1}{|y|^n}\Big[F\Big(\frac{A(y)}{|y|}\Big)+F\Big(\frac{A(-y)}{|y|}\Big)\Big]\,dy\right| \qquad\text{(B.14)}\\
&=\lim_{\varepsilon\searrow 0}\frac{1}{2}\left|\int_{V(\varepsilon)\setminus U(\varepsilon)}\frac{1}{|y|^n}\Big[F\Big(\frac{A(y)}{|y|}\Big)-F\Big(-\frac{A(-y)}{|y|}\Big)\Big]\,dy\right| \\
&\le\Big[\sup_{|\xi|\le M}|\nabla F(\xi)|\Big]\lim_{\varepsilon\searrow 0}\int_{\Delta(\varepsilon)}\eta(|y|)|y|^{-n}\,dy \\
&\le C\lim_{\varepsilon\searrow 0}\eta(\varepsilon)=0. \qquad\text{(B.15)}
\end{aligned}$$

This proves that the limits in (B.6) and (B.7) exist simultaneously and are equal.

In order to prove the simultaneous existence and coincidence of the limits in (B.7) and (B.8), observe that for each $y\in V(\varepsilon)\setminus W(\varepsilon)$ we have $M^{-1}|y|\le|B(y)|\le\varepsilon$, so $|y|\le\varepsilon M$. That is,

$$y\in V(\varepsilon)\setminus W(\varepsilon)\quad\Longrightarrow\quad |y|\le\varepsilon M. \tag{B.16}$$

In turn, this forces

$$|(DB)(0)y|\le|(DB)(0)y-B(y)|+|B(y)|\le\varepsilon M\eta(\varepsilon M)+\varepsilon \tag{B.17}$$

and, further,

$$y\in V(\varepsilon)\setminus W(\varepsilon)\quad\Longrightarrow\quad \varepsilon<|(DB)(0)y|\le\varepsilon M\eta(\varepsilon M)+\varepsilon. \tag{B.18}$$

From (B.16) and (B.18) we may therefore conclude that

$$V(\varepsilon)\setminus W(\varepsilon)\subseteq Z[\varepsilon;M\eta(\varepsilon M)], \tag{B.19}$$

where, in general, we define

$$Z[\varepsilon;a]:=\{y\in\mathbb{R}^n:\varepsilon<|DB(0)y|\le\varepsilon a+\varepsilon\}\quad\text{for all }\ \varepsilon>0\ \text{ and }\ a>0. \tag{B.20}$$

Let $\mathcal{H}^k_N$ be the $k$-dimensional Hausdorff measure in $\mathbb{R}^N$. To estimate the $n$-dimensional Lebesgue measure of $Z[\varepsilon;a]$, note first that, for each $a>0$ fixed,

$$Z[\varepsilon;a]=\varepsilon Z[1;a]\quad\text{for all }\ \varepsilon>0. \tag{B.21}$$

On the other hand, if we set $H_n:=\{DB(0)y:y\in\mathbb{R}^n\}\subseteq\mathbb{R}^{m'}$ then, since $DB(0)$ is a rank-$n$ matrix, it follows that $H_n$ is an $n$-dimensional plane in $\mathbb{R}^{m'}$ and $DB(0):\mathbb{R}^n\to H_n$ is a linear isomorphism. As such, we obtain

$$\begin{aligned}
\mathcal{H}^n_n(Z[1;a])&=\mathcal{H}^n_n(\{y\in\mathbb{R}^n:1<|DB(0)y|\le a+1\})\\
&\le C\mathcal{H}^n_{m'}(\{z\in H_n:1<|z|\le a+1\}).
\end{aligned}\tag{B.22}$$

A moment's reflection shows that

$$\lim_{a\to 0^+}\mathcal{H}^n_{m'}(\{z\in H_n:1<|z|\le a+1\})=0. \tag{B.23}$$

From this, (B.21), (B.19) and the fact that $\eta(\varepsilon M) \to 0$ as $\varepsilon \to 0^+$, we may conclude that

$$\lim_{\varepsilon \to 0^+} \frac{\mathcal{H}^n_n(V(\varepsilon)\setminus W(\varepsilon))}{\varepsilon^n} = 0. \tag{B.24}$$

Since the expression $(1/|y|^n)F(A(y)/|y|)$ restricted to $V(\varepsilon)\setminus W(\varepsilon)$ is pointwise of the order $\varepsilon^{-n}$ in a uniform fashion, we deduce from (B.24) that

$$\lim_{\varepsilon \to 0^+} \int_{V(\varepsilon)\setminus W(\varepsilon)} \frac{1}{|y|^n} F\left(\frac{A(y)}{|y|}\right) dy = 0, \tag{B.25}$$

as desired.

Finally, an argument analogous to (B.18) gives that

$$\varepsilon - \varepsilon M \eta(\varepsilon M) < |(DB)(0)y| \le \varepsilon \quad \text{for all } \ y \in W(\varepsilon)\setminus V(\varepsilon). \tag{B.26}$$

Thus, for reasons similar to those discussed above, we also have

$$\lim_{\varepsilon \to 0^+} \int_{W(\varepsilon)\setminus V(\varepsilon)} \frac{1}{|y|^n} F\left(\frac{A(y)}{|y|}\right) dy = 0, \tag{B.27}$$

which completes the proof of the lemma. □

The main result in this appendix, pertaining to the manner in which singular integrals are truncated, reads as follows:

**Proposition B.2.** *Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be a Lipschitz function and assume that $F : \mathbb{R}^m \to \mathbb{R}$ is an odd function of class $\mathscr{C}^N$ for some sufficiently large integer $N = N(m)$. Also, suppose $B : \mathbb{R}^n \to \mathbb{R}^{m'}$ is a bi-Lipschitz function and pick $p \in (1,\infty)$. Then, for each fixed $f \in L^p(\mathbb{R}^n)$, the limit*

$$\lim_{\varepsilon \searrow 0} \int_{\{y\in\mathbb{R}^n : |B(x)-B(y)|>\varepsilon\}} \frac{1}{|x-y|^n} F\left(\frac{A(x)-A(y)}{|x-y|}\right) f(y)\, dy \tag{B.28}$$

*exists at a.e. point $x \in \mathbb{R}^n$. Moreover, this limit is independent of the choice of the function $B$, in the sense that for each given $f \in L^p(\mathbb{R}^n)$ the limit* (B.28) *is equal to*

$$\lim_{\varepsilon \searrow 0} \int_{\{y\in\mathbb{R}^n : |x-y|>\varepsilon\}} \frac{1}{|x-y|^n} F\left(\frac{A(x)-A(y)}{|x-y|}\right) f(y)\, dy \tag{B.29}$$

*for a.e. $x \in \mathbb{R}^n$.*

As a preamble, we deal with a simple technical result. In the sequel, we agree to let $\mathcal{M}$ stand for the usual Hardy–Littlewood maximal operator.

**Lemma B.3.** *Assume that*

$$C_1|x-y| \le \rho(x,y) \le C_2|x-y| \quad \textit{for all } \ x,\ y \in \mathbb{R}^n \tag{B.30}$$

*and*

$$|k(x,y)| \le \frac{C_0}{|x-y|^n} \quad \textit{for all } \ x,\ y \in \mathbb{R}^n \tag{B.31}$$

*for some finite positive constants $C_0, C_1, C_2$. Then*

$$\Delta(x) := \left| \int_{\substack{|x-y|>\varepsilon \\ y\in\mathbb{R}^n}} k(x,y) f(y) dy - \int_{\substack{\rho(x-y)>\varepsilon \\ y\in\mathbb{R}^n}} k(x,y) f(y)\, dy \right| \leq C_0 (C_1^{-n} + C_2^n) \mathcal{M} f(x) \tag{B.32}$$

*for all $x \in \mathbb{R}^n$.*

*Proof.* A direct size estimate gives

$$\Delta(x) \leq \int_{\substack{|x-y|>\varepsilon,\, \rho(x,y)<\varepsilon \\ y\in\mathbb{R}^n}} \frac{C_0}{|x-y|^n} |f(y)|\, dy + \int_{\substack{|x-y|<\varepsilon,\, \rho(x,y)>\varepsilon \\ y\in\mathbb{R}^n}} \frac{C_0}{|x-y|^n} |f(y)|\, dy =: I + II, \tag{B.33}$$

where the last equality defines $I, II$. We have:

$$I \leq \frac{C_0}{\varepsilon^n} \int_{C_1|x-y|<\varepsilon} |f(y)|\, dy \leq \frac{C_0}{C_1^n} \mathcal{M} f(x) \tag{B.34}$$

and

$$II \leq \frac{C_0 C_2^n}{\varepsilon^n} \int_{|x-y|<\varepsilon} |f(y)|\, dy \leq C_0 C_2^n \mathcal{M} f(x). \tag{B.35}$$

The desired conclusion follows. ☐

Below, we shall also make use of the following standard result:

**Lemma B.4.** *Let $\{T_\varepsilon\}_{\varepsilon>0}$ be a family of operators with the following properties:*

(1) *There exists a dense subset $\mathcal{V}$ of $L^p(\mathbb{R}^n)$ such that for any $f \in \mathcal{V}$ the limit $\lim_{\varepsilon\to 0^+} T_\varepsilon f(x)$ exists for almost every $x \in \mathbb{R}^n$.*

(2) *The maximal operator $T_* f(x) := \sup\{|T_\varepsilon f(x)| : \varepsilon > 0\}$ is bounded on $L^p(\mathbb{R}^n)$.*

*Then, the limit $\lim_{\varepsilon\to 0^+} T_\varepsilon f(x)$ exists for any $f \in L^p(\mathbb{R}^n)$ at almost any $x \in \mathbb{R}^n$, and the operator*

$$Tf(x) := \lim_{\varepsilon\to 0^+} T_\varepsilon f(x) \tag{B.36}$$

*is bounded on $L^p(\mathbb{R}^n)$.*

*Proof.* The boundedness of the operator $T$ is an immediate consequence of (2), once we prove the existence of the limit in (B.36). In this regard, having fixed $f \in L^p(\mathbb{R}^n)$, we aim to show that

$$\Big| \{x \in \mathbb{R}^n : \limsup_{\varepsilon\to 0^+} T_\varepsilon f(x) \neq \liminf_{\varepsilon\to 0^+} T_\varepsilon f(x)\} \Big| = 0. \tag{B.37}$$

Fix $\theta > 0$ and consider

$$S := \Big\{ x \in \mathbb{R}^n : \Big| \limsup_{\varepsilon\to 0^+} T_\varepsilon f(x) - \liminf_{\varepsilon\to 0^+} T_\varepsilon f(x) \Big| > \theta \Big\}. \tag{B.38}$$

Also, fix $\delta > 0$ and select $h \in \mathcal{V}$ such that $\|f - h\|_{L^p(\mathbb{R}^n)} < \delta$. Then

$$S \subset S_1 \cup S_2, \tag{B.39}$$

where

$$S_1 := \{x \in \mathbb{R}^n : \big|\limsup_{\varepsilon\to 0^+} T_\varepsilon f(x) - \lim_{\varepsilon\to 0^+} T_\varepsilon h(x)\big| > \tfrac{1}{2}\theta\},$$
$$S_2 := \{x \in \mathbb{R}^n : \big|\liminf_{\varepsilon\to 0^+} T_\varepsilon f(x) - \lim_{\varepsilon\to 0^+} T_\varepsilon h(x)\big| > \tfrac{1}{2}\theta\}. \tag{B.40}$$

Then the measure of the set $S_1$ can be estimated by

$$|S_1| \le \big|\{x \in \mathbb{R}^n : T_*(f-h)(x) > \theta/2\}\big| \le \left(\frac{2}{\theta}\right)^p \int_{\mathbb{R}^n} |T_*(f-h)(x)|^p \, dx$$
$$\le C\left(\frac{2}{\theta}\right)^p \|f-h\|^p_{L^p(\mathbb{R}^n)} \le C\left(\frac{2}{\theta}\right)^p \delta^p. \tag{B.41}$$

Since $\delta > 0$ was arbitrary, this proves that $|S_1| = 0$. The same consideration works for the set $S_2$; hence also $|S| = 0$ by (B.39). This concludes the proof of Lemma B.4. □

We are now ready to present:

*Proof of Proposition B.2.* For each bi-Lipschitz function $B$ defined in $\mathbb{R}^n$, consider the truncated singular integral operator

$$T_{B,\varepsilon} f(x) := \int_{\{y\in\mathbb{R}^n : |B(x)-B(y)|>\varepsilon\}} \frac{1}{|x-y|^n} F\left(\frac{A(x)-A(y)}{|x-y|}\right) f(y)\, dy, \quad x \in \mathbb{R}^n, \tag{B.42}$$

where $\varepsilon > 0$. The maximal operator associated with the family $\{T_{B,\varepsilon}\}_{\varepsilon>0}$ is defined as

$$T_{B,*} f(x) := \sup_{\varepsilon>0} |T_{B,\varepsilon} f(x)|, \quad x \in \mathbb{R}^n. \tag{B.43}$$

In particular, corresponding to the case when $B = I$, the identity on $\mathbb{R}^n$, we have

$$T_{I,\varepsilon} f(x) = \int_{\{y\in\mathbb{R}^n : |x-y|>\varepsilon\}} \frac{1}{|x-y|^n} F\left(\frac{A(x)-A(y)}{|x-y|}\right) f(y)\, dy, \quad x \in \mathbb{R}^n, \tag{B.44}$$

and

$$T_{I,*} f(x) = \sup_{\varepsilon>0} |T_{I,\varepsilon} f(x)|, \quad x \in \mathbb{R}^n. \tag{B.45}$$

We proceed is a number of steps.

**Step 1.** *Given* $p \in (1,\infty)$ *there exists a constant* $C \in (0,\infty)$ *with the property that, for each Lipschitz function* $A : \mathbb{R} \to \mathbb{R}$ *and for each* $\varepsilon > 0$, *the truncated Cauchy integral operator*

$$\mathscr{C}_{A,\varepsilon} f(x) := \int_{\{y\in\mathbb{R} : |x-y|>\varepsilon\}} \frac{f(y)}{x-y+i(A(x)-A(y))}\, dy, \quad x \in \mathbb{R}, \tag{B.46}$$

*satisfies*

$$\|\mathscr{C}_{A,\varepsilon} f\|_{L^p(\mathbb{R})} \le C(1 + \|A'\|_{L^\infty(\mathbb{R})}) \|f\|_{L^p(\mathbb{R})}. \tag{B.47}$$

This is the Coifman–McIntosh–Meyer theorem [Coifman et al. 1982]. An elegant proof is given by M. Melnikov and J. Verdera [1995].

**Step 2.** *Given $p \in (1,\infty)$ there exists a constant $C \in (0,\infty)$ with the property that, if $\beta \in (1,\infty)$ and if $B : \mathbb{R} \to \mathbb{R}$ is a Lipschitz function satisfying $\beta^{-1} < B'(x) < \beta$ for a.e. $x \in \mathbb{R}$, then for each $\varepsilon > 0$ and each $\eta \in [-1,1]$ the operator*

$$\widetilde{\mathscr{C}}_{B,\eta,\varepsilon} f(x) := \int_{\{y\in\mathbb{R}:|x-y|>\varepsilon\}} \frac{f(y)}{\eta(x-y)i + B(x) - B(y)}\,dy, \quad x \in \mathbb{R}, \tag{B.48}$$

*satisfies*

$$\|\widetilde{\mathscr{C}}_{B,\eta,\varepsilon} f\|_{L^p(\mathbb{R})} \le C\beta^4 \|f\|_{L^p(\mathbb{R})}. \tag{B.49}$$

To prove (B.49), changing variables $s := B(x)$ and $t := B(y)$ allows us to write

$$(\widetilde{\mathscr{C}}_{B,\eta,\varepsilon} f)(B^{-1}(s)) = \int_{|B^{-1}(s)-B^{-1}(t)|>\varepsilon} \frac{f(B^{-1}(t))[B'(B^{-1}(t))]^{-1}}{s-t+i\eta(B^{-1}(s)-B^{-1}(t))}\,dt. \tag{B.50}$$

Based on this and Lemma B.3, we then obtain the pointwise estimate

$$|(\widetilde{\mathscr{C}}_{B,\eta,\varepsilon} f)(B^{-1}(s))| \le |\mathscr{C}_{\eta B^{-1},\varepsilon}((f/B')\circ B^{-1})(s)| + C\beta^3 \mathscr{M} f(B^{-1}(s)) \tag{B.51}$$

for all $s \in \mathbb{R}$. Then (B.49) follows from (B.51) with the help of (B.47).

**Step 3.** *Suppose $F(z)$ is an analytic function in the open strip $\{z \in \mathbb{C} : |\mathrm{Im}\, z| < 2\}$. Let $A : \mathbb{R} \to \mathbb{R}$ be a Lipschitz function with $\|A'\|_{L^\infty(\mathbb{R})} \le M$. Then, for each $p \in (1,\infty)$ there exists a constant $C = C_p \in (0,\infty)$ such that, for each $\varepsilon > 0$, the operator*

$$K_{A,F,\varepsilon} f(x) := \int_{|x-y|>\varepsilon} \frac{1}{x-y} F\left(\frac{A(x)-A(y)}{x-y}\right) f(y)\,dy, \quad x \in \mathbb{R}, \tag{B.52}$$

*satisfies*

$$\|K_{A,F,\varepsilon} f\|_{L^p(\mathbb{R})} \le C(1+M^4)\sup\{|F(z)| : z \in \mathbb{C},\ |\mathrm{Im}\, z| < 2\}\|f\|_{L^p(\mathbb{R})}. \tag{B.53}$$

To justify (B.53), let $\gamma^1_\pm := \{\zeta = u \pm i : |u| \le 2M\}$, $\gamma^2_\pm := \{\zeta = \pm 2M + iv : |v| \le 1\}$, and set $\gamma := \gamma^1_+ \cup \gamma^2_+ \cup \gamma^1_- \cup \gamma^2_-$. Since $F$ is analytic for $z \in \mathbb{C}$ with $|\mathrm{Im}\, z| < 2$, Cauchy's reproducing formula yields

$$F(s) = \frac{1}{2\pi i}\int_\gamma \frac{F(\zeta)}{\zeta - s}\,d\zeta = \frac{1}{2\pi i}\int_{\gamma^1_+\cup\gamma^1_-} \frac{F(\zeta)}{\zeta - s}\,d\zeta + \frac{1}{2\pi i}\int_{\gamma^2_+\cup\gamma^2_-} \frac{F(\zeta)}{\zeta - s}\,d\zeta. \tag{B.54}$$

Accordingly,

$$\begin{aligned} K_{A,F,\varepsilon} f(x) &= \frac{1}{2\pi i}\int_{\gamma^1_+\cup\gamma^1_-} F(\zeta)\int_{|x-y|>\varepsilon} \frac{1}{x-y}\frac{f(y)}{\zeta - \frac{A(x)-A(y)}{x-y}}\,dy\,d\zeta \\ &\quad + \frac{1}{2\pi i}\int_{\gamma^2_+\cup\gamma^2_-} F(\zeta)\int_{|x-y|>\varepsilon} \frac{1}{x-y}\frac{f(y)}{\zeta - \frac{A(x)-A(y)}{x-y}}\,dy\,d\zeta \\ &= I_+ + I_- + II_+ + II_-, \end{aligned} \tag{B.55}$$

where

$$I_\pm := \mp\frac{1}{2\pi}\int_{\gamma^1_\pm} F(\zeta)\int_{|x-y|>\varepsilon} \frac{f(y)}{x-y+i[A^\pm_\zeta(x)-A^\pm_\zeta(y)]}\,dy\,d\zeta \tag{B.56}$$

with $A^{\pm}_{\zeta}(x) := \mp[A(x) - (\Re\,\zeta)x]$, and

$$II_{\pm} := \frac{1}{2\pi i}\int_{\gamma^2_{\pm}} F(\zeta)\int_{|x-y|>\varepsilon} \frac{f(y)}{(\operatorname{Im}\zeta)(x-y)i + [B^{\pm}(x) - B^{\pm}(y)]}\,dy\,d\zeta \tag{B.57}$$

with $B^{\pm}(x) := -[A(x) \mp 2Mx]$. At this point, the proof of (B.53) is concluded by invoking the results from Steps 1–2.

**Step 4.** *Suppose $F \in \mathscr{C}^N(\mathbb{R})$, $N \geq 6$, and assume that $A : \mathbb{R} \to \mathbb{R}$ is a Lipschitz function with $\|A'\|_{L^\infty(\mathbb{R})} \leq M$. Then, for each $p \in (1,\infty)$, there exists a constant $C = C_p \in (0,\infty)$ such that the operator* (B.52) *satisfies, for each $\varepsilon > 0$,*

$$\|K_{A,F,\varepsilon} f\|_{L^p(\mathbb{R})} \leq C(1 + M^4)\sup\{|F^{(k)}(x)| : |x| \leq M+1,\ 0 \leq k \leq 6\}\|f\|_{L^p(\mathbb{R})}. \tag{B.58}$$

In dealing with (B.58), there is no loss of generality in assuming that $F$ is supported in the interval $[-M-1, M+1]$. With "hat" denoting the Fourier transform we have

$$K_{A,F,\varepsilon} f(x) = \int_{\mathbb{R}} \widehat{F}(\xi)\left(\int_{\{y\in\mathbb{R}:|x-y|>\varepsilon\}} \frac{1}{x-y} e^{i\xi\frac{A(x)-A(y)}{x-y}} f(y)\,dy\right) d\xi. \tag{B.59}$$

Note that the inner integral above is precisely the truncated Cauchy operator (B.46) corresponding to the choice $F(z) := \exp(iz)$ and with $A$ replaced by $\xi A$. Consequently, (B.58) follows from (B.59) with the help of (B.53).

**Step 5.** *Suppose $F \in \mathscr{C}^N(\mathbb{R}^m)$, $N \geq m+5$, $F$ is odd, and assume that $A : \mathbb{R}^n \to \mathbb{R}^m$ is a Lipschitz function with $\|DA\|_{L^\infty(\mathbb{R}^n,\mathbb{R}^m)} \leq M$. Then, for each $p \in (1,\infty)$, there exists a constant $C = C_p \in (0,\infty)$ such that, for each $\varepsilon > 0$, the operator*

$$K_{A,F,\varepsilon} f(x) := \int_{|x-y|>\varepsilon} \frac{1}{|x-y|^n} F\left(\frac{A(x)-A(y)}{|x-y|}\right) f(y)\,dy, \quad x \in \mathbb{R}^n, \tag{B.60}$$

*satisfies*

$$\|K_{A,F,\varepsilon} f\|_{L^p(\mathbb{R}^n)} \leq C(1 + M^4)\sup\{|\partial^\alpha F(x)| : |x| \leq M+1,\ |\alpha| \leq m+5\}\|f\|_{L^p(\mathbb{R}^n)}. \tag{B.61}$$

In the case $n = 1$, since $F$ is odd we may write

$$\frac{1}{|x-y|} F\left(\frac{A(x)-A(y)}{|x-y|}\right) = \frac{1}{x-y} F\left(\frac{A(x)-A(y)}{x-y}\right), \tag{B.62}$$

so (B.61) follows from an argument similar to the one used in the treatment of Step 4, based on writing

$$K_{A,F,\varepsilon} f(x) = \int_{\mathbb{R}^m} \widehat{F}(\xi)\left(\int_{\{y\in\mathbb{R}:|x-y|>\varepsilon\}} \frac{1}{x-y} e^{i\langle\xi,\frac{A(x)-A(y)}{x-y}\rangle} f(y)\,dy\right) d\xi \tag{B.63}$$

and invoking the result established in Step 3. For $n > 1$ we can reduce the problem to the one-dimensional case by the classical method of rotation.

**Step 6.** *Retain the same assumptions as in Step* 5. *Then there is a constant $C$ such that*

$$|\{x \in \mathbb{R}^n : |K_{A,F,\varepsilon} f(x)| > \lambda\}| \leq \frac{C}{\lambda}\|f\|_{L^1(\mathbb{R}^n)} \tag{B.64}$$

*for every function* $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ *and every positive number* $\lambda$. *In particular,* $K_{A,F,\varepsilon}$ *extends to a bounded operator from* $L^1(\mathbb{R}^n)$ *into* $L^{1,\infty}(\mathbb{R}^n)$ *(where* $L^{1,\infty}(\mathbb{R}^n)$ *stands for the weak-*$L^1$ *space in* $\mathbb{R}^n$*).*

This follows from Step 5 (with $p = 2$) and the classical Calderón–Zygmund lemma.

**Step 7.** *Retain the same assumptions as in Step 5. There exists a finite constant* $C > 0$ *depending only on the dimension with the property that, for each fixed* $\varepsilon_0 > 0$, *the following Cotlar-type estimate holds:*

$$K^{(\varepsilon)}_{A,F,*} f(x) \le C\mathcal{M}f(x) + 2\mathcal{M}(K_{A,F,\varepsilon_0} f)(x) \quad \text{for all } \varepsilon > \varepsilon_0 \tag{B.65}$$

*for each* $f \in \mathrm{Lip}_{comp}(\mathbb{R}^n)$ *and each* $x \in \mathbb{R}^n$, *where*

$$K^{(\varepsilon)}_{A,F,*} f(x) := \sup_{\varepsilon'>\varepsilon} |K_{A,F,\varepsilon'} f(x)|. \tag{B.66}$$

Without loss of generality, it suffices to prove (B.65) for $x = 0$, so we focus on showing that

$$|K_{A,F,\varepsilon} f(0)| \le C\mathcal{M}f(0) + 2\mathcal{M}(K_{A,F,\varepsilon_0}) f(0) \quad \text{for all } \varepsilon > \varepsilon_0. \tag{B.67}$$

Then (B.67) implies (B.65) by suitably taking the supremum.

The first step is to observe that, for all $x \in \mathbb{R}^n$ and for all $\varepsilon > 0$,

$$|K_{A,F,\varepsilon} f(x') - K_{A,F,\varepsilon} f(x)| \le C\mathcal{M}f(0) \quad \text{provided } |x - x'| \le \varepsilon/2. \tag{B.68}$$

To see that this is the case, abbreviate $k(x, y) := F\big((A(x) - A(y))/|x - y|\big)/|x - y|^n$, then write

$$\begin{aligned} |K_{A,F,\varepsilon} f(x') - K_{A,F,\varepsilon} f(x)| &\le \left| \int_{|x-y|\ge\varepsilon} (k(x', y) - k(x, y)) f(y)\, dy \right| \\ &\quad + \left| \int_{|x'-y|\ge\varepsilon} k(x', y) f(y)\, dy - \int_{|x-y|\ge\varepsilon} k(x', y) f(y)\, dy \right| \\ &=: I + II. \end{aligned} \tag{B.69}$$

The term $II$ can be bounded by a multiple of $\mathcal{M}f(0)$ using an argument similar to that in Lemma B.3. The estimate for $I$ follows from the mean vale theorem, the nature of the kernel $k(x, y)$, and the standard inequality

$$\varepsilon \int_{|y|\ge\varepsilon} |y|^{-n-1} |f(y)|\, dy \le C\mathcal{M}f(0) \quad \text{for all } \varepsilon > 0. \tag{B.70}$$

Turning to the proof of (B.67) in earnest, fix $\varepsilon > \varepsilon_0 > 0$ then introduce $f_1 := f\chi_{B(0,\varepsilon)}$ and set $f_2 := f - f_1$. In particular, this entails

$$K_{A,F,\varepsilon} f(0) = K_{A,F,\varepsilon_0} f_2(0). \tag{B.71}$$

Then, for each $x \in B(0, \varepsilon/2)$, by (B.68) we have

$$|K_{A,F,\varepsilon_0} f_2(x) - K_{A,F,\varepsilon_0} f_2(0)| \le C\mathcal{M}f(0); \tag{B.72}$$

therefore,

$$|K_{A,F,\varepsilon_0} f_2(0)| \le |K_{A,F,\varepsilon_0} f(x)| + |K_{A,F,\varepsilon_0} f_1(x)| + C\mathcal{M}f(0) \quad \text{for a.e. } x \in B(0, \varepsilon/2). \tag{B.73}$$

We finish the proof by analyzing the weak-$L^1$ norms of the above functions. To this end, define

$$N(f) := \sup_{\lambda>0}\big[\lambda\mu(\{x \in B : |f(x)| > \lambda\})\big], \tag{B.74}$$

where $B := B(0,\varepsilon/2)$ and $\mu$ stands for the $n$-dimensional Lebesgue measure restricted to the ball $B$ of constant density $|B|^{-1}$. Observe that $f(x) = \alpha$ on $B$ implies $N(f) = \alpha$ for any constant $\alpha$, and that $N(f_1 + f_2 + f_3) \le 2N(f_1) + 4N(f_2) + 4N(f_3)$ for all functions $f_1$, $f_2$ and $f_3$. Then the estimate

$$|K_{A,F,\varepsilon} f(0)| = |K_{A,F,\varepsilon_0} f_2(0)| \le 2N(K_{A,F,\varepsilon_0} f) + 4N(K_{A,F,\varepsilon_0} f_1) + 4C\mathcal{M}f(0) \tag{B.75}$$

follows from (B.71), these observations and (B.73). It remains to note that the right-hand side above can be further bounded using Chebyshev's inequality, which yields $N(K_{A,F,\varepsilon_0} f) \le C\mathcal{M}(K_{A,F,\varepsilon_0} f)(0)$, and the weak-$L^1$ boundedness result from Step 6, which eventually gives $N(K_{A,F,\varepsilon_0} f_1) \le C\mathcal{M}f(0)$. From these, (B.67) follows.

**Step 8.** *Retain the same assumptions as in Step* 5 *and consider the maximal operator*

$$K_{A,F,*} f(x) := \sup_{\varepsilon>0} |K_{A,F,\varepsilon} f(x)|, \quad x \in \mathbb{R}^n. \tag{B.76}$$

*Then for each* $p \in (1,\infty)$ *there exists a constant* $C = C(F, A, m, n, p) \in (0,\infty)$ *with the property that*

$$\|K_{A,F,*} f\|_{L^p(\mathbb{R}^n)} \le C\|f\|_{L^p(\mathbb{R}^n)} \quad \textit{for all } f \in L^p(\mathbb{R}^n). \tag{B.77}$$

To see this, fix an arbitrary $f \in L^p(\mathbb{R}^n)$ and first observe from (B.66) that for each $x \in \mathbb{R}^n$ we have

$$K^{(\varepsilon)}_{A,F,*} f(x) \nearrow K_{A,F,*} f(x) \quad \text{as } \varepsilon \searrow 0. \tag{B.78}$$

Based on this, Lebesgue's monotone convergence theorem, (B.65), (B.61) and the boundedness of the Hardy–Littlewood maximal function, we obtain

$$\begin{aligned}\|K_{A,F,*} f\|_{L^p(\mathbb{R}^n)} &= \lim_{\varepsilon\to 0^+} \|K^{(\varepsilon)}_{A,F,*} f\|_{L^p(\mathbb{R}^n)}\\ &\le C \lim_{\varepsilon\to 0^+} (\|\mathcal{M}f\|_{L^p(\mathbb{R}^n)} + \|\mathcal{M}(K_{A,F,\varepsilon/2} f)\|_{L^p(\mathbb{R}^n)}) \le C\|f\|_{L^p(\mathbb{R}^n)},\end{aligned} \tag{B.79}$$

completing the proof of (B.77).

In terms of the maximal operator $T_{I,*}$ from (B.45), estimate (B.77) yields

$$\|T_{I,*} f\|_{L^p(\mathbb{R}^n)} \le C\|f\|_{L^p(\mathbb{R}^n)} \quad \text{for all } f \in L^p(\mathbb{R}^n). \tag{B.80}$$

In order to show the existence of the pointwise limit in (B.29), the strategy is to return to the various particular operators discussed in Steps 1–5 and show that, in each case, such a pointwise convergence holds for such operators acting on functions in $L^p$, almost everywhere in $\mathbb{R}^n$. In all cases, we shall make use of the abstract scheme described in Lemma B.4.

**Step 9.** *Pointwise convergence for the Cauchy operator* (B.46): *Let* $\mathcal{V} := (1 + iA')\,\mathrm{Lip}_{\mathrm{comp}}(\mathbb{R})$, *which is a dense subclass of* $L^p(\mathbb{R})$, $1 < p < \infty$, *since* $A$ *is real-valued and Lipschitz. We claim that*

$$\textit{for any } h \in \mathcal{V},\ \lim_{\varepsilon\to 0^+} \mathcal{C}_{A,\varepsilon} h(x) \textit{ exists for a.e. } x \in \mathbb{R}. \tag{B.81}$$

Indeed, if $h = (1 + iA')f$ with $f \in \mathrm{Lip}_{\mathrm{comp}}(\mathbb{R})$, then we can write

$$\begin{aligned}\mathscr{C}_{A,\varepsilon}h(x) = &\int_{1>|x-y|>\varepsilon} \frac{1+iA'(y)}{x-y+i(A(x)-A(y))}(f(y)-f(x))\,dy\\ &- f(x)\int_{1>|x-y|>\varepsilon} \frac{-(1+iA'(y))}{x-y+i(A(x)-A(y))}\,dy\\ &+ \int_{|x-y|>1} \frac{1+iA'(y)}{x-y+i(A(x)-A(y))} f(y)\,dy\\ =:&\ I + II + III. \end{aligned} \tag{B.82}$$

Using the fact that $f$ is a compactly supported Lipschitz function, it is immediate that $\lim_{\varepsilon\to 0^+} I$ and $\lim_{\varepsilon\to 0^+} III$ exist at every $x \in \mathbb{R}$. Furthermore, the fundamental theorem of calculus gives

$$II = -f(x)\ln\left(\frac{-1+i(A(x)-A(x+\varepsilon))/\varepsilon}{1+i(A(x)-A(x-\varepsilon))/\varepsilon}\right) \tag{B.83}$$

and the limit as $\varepsilon \to 0^+$ of the right-hand side exists for almost every $x \in \mathbb{R}$ since, by Rademacher's theorem, the Lipschitz function $A$ is a.e. differentiable. This concludes the proof of (B.81).

Finally, a combination of (B.81), Lemma B.4 and (a suitable version of) the maximal inequality (B.80) gives that for $f \in L^p(\mathbb{R})$ the limit $\lim_{\varepsilon\to 0^+} \mathscr{C}_{A,\varepsilon} f(x)$ exists for almost every $x \in \mathbb{R}$.

**Step 10.** *Pointwise convergence for the Cauchy operator* (B.48).

Using Step 9, (B.50) and Lemma B.1, it follows that, for each function $f \in L^p(\mathbb{R})$, the limit $\lim_{\varepsilon\to 0^+} \widetilde{\mathscr{C}}_{B,\eta,\varepsilon} f(x)$ exists for almost every $x \in \mathbb{R}$.

**Step 11.** *Pointwise convergence for the operator* (B.52). *Specifically, we claim that, if* $f \in L^p(\mathbb{R})$, *the limit* $\lim_{\varepsilon\to 0} K_{A,F,\varepsilon} f(x)$ *exists for almost every* $x \in \mathbb{R}$.

In order to prove this claim, fix $f \in L^p(\mathbb{R})$ and recall $I_\pm$, $II_\pm$ as defined in (B.55). The goal is to first show that $\lim_{\varepsilon\to 0} I_+$ exists for almost every $x \in \mathbb{R}$. To this end, for $x, \zeta \in \mathbb{R}$ set

$$F_\varepsilon^{\zeta,x} := F(\zeta)\int_{|x-y|>\varepsilon} \frac{f(y)}{x-y+i[A_\zeta^\pm(x)-A_\zeta^\pm(y)]}\,dy. \tag{B.84}$$

Then, employing Step 9 it follows that for each $\zeta \in \gamma_+^1$ the limit

$$\lim_{\varepsilon\to 0^+} F_\varepsilon^{\zeta,x} \tag{B.85}$$

exists for almost every $x \in \mathbb{R}$. Next, we want to prove that $\sup_{\varepsilon>0} |F_\varepsilon^{\zeta,x}| \in L^1_\zeta(\gamma_+^1)$ for almost every $x \in \mathbb{R}$. To see the latter we write

$$\int_{\mathbb{R}}\left|\int_{\gamma_+^1} \sup_{\varepsilon>0} |F_\varepsilon^{\zeta,x}|\,d\zeta\right|^2 dx \le \int_{\gamma_+^1}\int_{\mathbb{R}} \bigl(\sup_{\varepsilon>0} |F_\varepsilon^{\zeta,x}|\bigr)^2\,dx\,d\zeta \le C\|f\|_{L^2(\mathbb{R})}. \tag{B.86}$$

The first inequality in (B.86) is standard, while for the second one we have used (a suitable version of) the maximal inequality (B.80). The above analysis provides all the ingredients necessary for invoking

Lebesgue's dominated convergence theorem, which, in turn, allows us to conclude that

$$\lim_{\varepsilon\to 0^+} I_+ = \lim_{\varepsilon\to 0^+}\left\{-\frac{1}{2\pi}\int_{\gamma_+^1} F_\varepsilon^{\zeta,x}\,d\zeta\right\} \quad \text{exists at almost every point } x\in\mathbb{R}. \tag{B.87}$$

Similarly, one shows that $\lim_{\varepsilon\to 0^+} I_-$, $\lim_{\varepsilon\to 0^+} II_\pm$ exist for almost every $x\in\mathbb{R}$, and thus the earlier claim is proved.

**Step 12.** *Pointwise convergence for the operator* (B.58).

The fact that for $f\in L^p(\mathbb{R})$, the limit $\lim_{\varepsilon\to 0^+} K_{A,F,\varepsilon} f(x)$ exists for almost every $x\in\mathbb{R}$ follows by a reasoning similar to the one in Step 11. This time the identity (B.59) replaces the expressions in (B.55) and the decay properties of the Fourier transform $\widehat{F}(\xi)$ in are used when applying Lebesgue's dominated convergence theorem.

**Step 13.** *For each given* $f\in L^p(\mathbb{R}^n)$, *the limit* (B.29) *exists for a.e.* $x\in\mathbb{R}^n$.

Indeed, the case $n=1$ has been treated in Step 12. Finally, in the case $n>1$, the existence of the limit in question for $f\in\mathscr{C}_0^\infty(\mathbb{R}^n)$ follows via the rotation method from the one-dimensional result (and Lebesgue's dominated convergence theorem). Granted this, we may invoke Lemma B.4 and the maximal inequality (B.80) in order to finish, keeping in mind that $\mathscr{C}_0^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$.

In summary, at this point we know that

$$\text{for each } f\in L^p(\mathbb{R}^n), \text{ the limit } \lim_{\varepsilon\to 0^+} T_{I,\varepsilon} f(x) \text{ exists for a.e. } x\in\mathbb{R}^n. \tag{B.88}$$

In turn, this readily yields that

$$\lim_{\varepsilon\searrow 0}\int_{\{y\in\mathbb{R}^n:1>|x-y|>\varepsilon\}} \frac{1}{|x-y|^n} F\left(\frac{A(x)-A(y)}{|x-y|}\right) dy \quad \text{exists for a.e. } x\in\mathbb{R}^n. \tag{B.89}$$

With this in hand and relying on Lemma B.1, we deduce that, for each bi-Lipschitz function $B$,

$$\lim_{\varepsilon\searrow 0}\int_{\{y\in\mathbb{R}^n:|B(x)-B(y)|>\varepsilon,\,|x-y|<1\}} \frac{1}{|x-y|^n} F\left(\frac{A(x)-A(y)}{|x-y|}\right) dy \quad \text{exists for a.e. } x\in\mathbb{R}^n \tag{B.90}$$

and the limits in (B.89) and (B.90) are equal. Having proved this, it follows that

$$\text{for each } f\in\mathscr{C}_0^\infty(\mathbb{R}^n),\ \lim_{\varepsilon\to 0^+} T_{B,\varepsilon} f(x) \text{ exists for a.e. } x\in\mathbb{R}^n \text{ and is equal to } \lim_{\varepsilon\to 0^+} T_{I,\varepsilon} f(x). \tag{B.91}$$

Let us also note that, thanks to (B.80) and Lemma B.3,

$$\|T_{B,*} f\|_{L^p(\mathbb{R}^n)} \le C\|f\|_{L^p(\mathbb{R}^n)} \quad \text{for all } f\in L^p(\mathbb{R}^n). \tag{B.92}$$

From (B.91), (B.92) and Lemma B.4 we may finally conclude that for each fixed $f\in L^p(\mathbb{R}^n)$ the limit (B.28) exists at a.e. point $x\in\mathbb{R}^n$ and is equal to (B.29). This finishes the proof of Proposition B.2. □

**Appendix C.** ***Background on*** $\mathbf{OP}(L^\infty \cap \mathbf{vmo})S^{\mathbf{0}}_{\mathbf{cl}}$**.** If $X$ is a Banach space of functions on $\mathbb{R}^n$, we say a function $p$ on points $(x,\xi) \in \mathbb{R}^n \times \mathbb{R}^n$ belongs to the symbol class $XS^m_{1,0}$,

$$p \in XS^m_{1,0}, \tag{C.1}$$

provided $p(\cdot,\xi) \in X$ for each $\xi \in \mathbb{R}^n$ and

$$\|\partial_\xi^\alpha p(\cdot,\xi)\|_X \le C_\alpha \langle\xi\rangle^{m-|\alpha|} \quad \text{for all } \alpha \in \mathbb{N}_0^n, \tag{C.2}$$

where $\langle\xi\rangle := (1+|\xi|^2)^{1/2}$ and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. If, in addition,

$$p(x,\xi) \sim \sum_{j\ge 0} p_j(x,\xi), \quad p_j(x,r\xi) = r^{m-j} p_j(x,\xi) \quad \text{for } r,\ |\xi| \ge 1, \tag{C.3}$$

in the sense that for every $k \in \mathbb{N}$ the difference $p - \sum_{j=0}^{k-1} p_j$ belongs to $XS^{m-k}_{1,0}$, we say

$$p \in XS^m_{\mathrm{cl}}. \tag{C.4}$$

The associated operator $p(x,D)$ is given by

$$p(x,D)u = (2\pi)^{-n/2} \int p(x,\xi)\hat{u}(\xi)e^{ix\cdot\xi}\,d\xi. \tag{C.5}$$

If (C.1) holds, we say $p(x,D) \in \mathrm{OPXS}^m_{1,0}$, and if (C.4) holds, we say $p(x,D) \in \mathrm{OPXS}^m_{\mathrm{cl}}$.

Here we single out the spaces

$$L^\infty(\mathbb{R}^n), \quad \mathrm{bmo}(\mathbb{R}^n), \quad \mathrm{vmo}(\mathbb{R}^n), \quad L^\infty(\mathbb{R}^n) \cap \mathrm{vmo}(\mathbb{R}^n) \tag{C.6}$$

to play the role of $X$. Here bmo is the localized variant of BMO, and vmo that of VMO. We summarize some results about the associated pseudodifferential operators. Details can be found in [Taylor 2000, Chapter 1, §11], which builds on work in [Chiarenza et al. 1991; Taylor 1997, §6]. A key ingredient in the proofs of these results is the classical commutator estimate of [Coifman et al. 1976],

$$\|[M_g, B]u\|_{L^p} \le C_p \|g\|_{\mathrm{bmo}} \|u\|_{L^p} \tag{C.7}$$

given $B \in \mathrm{OPS}^0_{1,0}$. Here $M_g u := gu$ is the operator of multiplication by $g$.

The following extension appears in [Taylor 2000, Proposition 11.1]:

**Proposition C.1.** *If* $p(x,D) \in \mathrm{OP(bmo)}S^0_{\mathrm{cl}}$ *and* $B = b(x,D) \in \mathrm{OPS}^0_{1,\delta}$, $\delta < 1$, *with* $B$ *scalar*, *then*

$$[p(x,D), B] : L^p(\mathbb{R}^n) \longrightarrow L^p(\mathbb{R}^n), \quad 1 < p < \infty. \tag{C.8}$$

*If* $p \in \mathrm{vmo}\, S^0_{\mathrm{cl}}$ *and* $b \in S^0_{1,\delta}$ *have compact* $x$*-support*, *this commutator is compact.*

This result in turn helps prove the following, which may be found in [Taylor 2000, Proposition 11.3].

**Proposition C.2.** *Assume that*

$$p \in L^\infty S^0_{\mathrm{cl}}, \quad q \in (L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}, \tag{C.9}$$

*with compact* $x$*-support. Then*

$$p(x,D)q(x,D) = a(x,D) + K, \quad a(x,\xi) = p(x,\xi)q(x,\xi), \tag{C.10}$$

*with $K$ compact on $L^p(\mathbb{R}^n)$ for $1 < p < \infty$.*

The following result has a proof parallel to that of Proposition C.2:

**Proposition C.3.** *Assume $q \in (L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}$, with compact $x$-support, and set*

$$q^*(x, \xi) = q(x, \xi)^*. \tag{C.11}$$

*Then*

$$q(x, D)^* = q^*(x, D) + K, \tag{C.12}$$

*with $K$ compact on $L^p(\mathbb{R}^n)$ for $1 < p < \infty$.*

To proceed, we have the following useful result, which appears in [Taylor 2000, Proposition 11.4].

**Proposition C.4.** *The space $L^\infty \cap \mathrm{vmo}$ is a closed subalgebra of $L^\infty(\mathbb{R}^n)$.*

Putting Propositions C.2 and C.4 together yields the following:

**Corollary C.5.** *Assume that*

$$p,\ q \in (L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}, \tag{C.13}$$

*with compact $x$-support. Then*

$$p(x, D)q(x, D) = a(x, D) + K, \tag{C.14}$$

*with $K$ compact on $L^p(\mathbb{R}^n)$ for $1 < p < \infty$ and*

$$a = pq \in (L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}. \tag{C.15}$$

Generally, if $\mathscr{A}$ is a $C^*$-algebra and $\mathscr{B}$ a closed *-subalgebra of $\mathscr{A}$ containing the identity element, and if $f \in \mathscr{B}$, then $f$ is invertible in $\mathscr{B}$ if and only if it is invertible in $\mathscr{A}$. To see this, consider $h = f^* f$ and expand $H(z) = (h + 1 - z)^{-1}$ in a power series about $z = 0$. The radius of convergence is greater than 1 if $f$ is invertible in $\mathscr{A}$. Clearly, $H(z) \in \mathscr{B}$ for $|z| < 1$ if $f \in \mathscr{B}$, so $H(1) \in \mathscr{B}$.

Consequently, we have

$$a \in L^\infty \cap \mathrm{vmo}, \quad a^{-1} \in L^\infty \quad \Longrightarrow \quad a^{-1} \in L^\infty \cap \mathrm{vmo}. \tag{C.16}$$

This holds for matrix-valued $a(x)$. Similarly, if

$$p \in (L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}} \quad \text{is elliptic,} \tag{C.17}$$

so that there exist $C_j < \infty$ such that

$$|p(x, \xi)^{-1}| \le C_1 \quad \text{for } |\xi| \ge C_2, \tag{C.18}$$

then

$$(1 - \varphi(\xi))p(x, \xi)^{-1} \in (L^\infty \cap \mathrm{vmo})S^0_{\mathrm{cl}}, \tag{C.19}$$

where $\varphi \in C_0^\infty(\mathbb{R}^n)$ is equal to 1 for $|\xi| \le C_2$. This allows the construction of Fredholm inverses of elliptic operators with coefficients in $L^\infty \cap \mathrm{vmo}$.

**Appendix D.** ***Analysis on spaces of homogeneous type.*** We begin by discussing a few results of a general nature, valid in the context of spaces of homogeneous type. Recall that $(\Sigma, \rho)$ is a quasimetric space if $\Sigma$ is a set (of cardinality at least two) and the mapping $\rho : \Sigma \times \Sigma \to [0, \infty)$ is a quasidistance; that is, there exists $C \in [1, \infty)$ such that, for every $x, y, z \in \Sigma$, $\rho$ satisfies

$$\rho(x, y) = 0 \Longleftrightarrow x = y, \quad \rho(y, x) = \rho(x, y), \quad \rho(x, y) \le C(\rho(x, z) + \rho(z, y)). \tag{D.1}$$

A space of homogeneous type in the sense of Coifman and Weiss [1977] is a triplet $(\Sigma, \rho, \mu)$ such that $(\Sigma, \rho)$ is a quasimetric space and $\mu$ is a Borel measure on $\Sigma$ (equipped with the topology canonically induced by $\rho$) that is doubling. That is, there exists $C \in (0, \infty)$ such that

$$0 < \mu(B_\rho(x, 2r)) \le C\mu(B_\rho(x, r)) \quad \text{for all } x \in \Sigma \text{ and } r > 0, \tag{D.2}$$

where $B_\rho(x, r)$ is the $\rho$-ball of center $x$ and radius $r$ given by $\{y \in \Sigma : \rho(x, y) < r\}$.

Then the John–Nirenberg space of functions of bounded mean oscillations, $\mathrm{BMO}(\Sigma, \mu)$, consists of functions $f \in L^1_{\mathrm{loc}}(\Sigma, \mu)$ for which $\|f\|_{\mathrm{BMO}(\Sigma,\mu)} < +\infty$. As usual, we have set

$$\|f\|_{\mathrm{BMO}(\Sigma,\mu)} := \begin{cases} \sup_{R>0} M_1(f; R) & \text{if } \mu(\Sigma) = +\infty, \\ \left|\int_\Sigma f\, d\mu\right| + \sup_{R>0} M_1(f; R) & \text{if } \mu(\Sigma) < +\infty, \end{cases} \tag{D.3}$$

where, for $p \in [1, \infty)$, we have set

$$\begin{aligned} M_p(f; R) &:= \sup_{x\in\Sigma} \sup_{r\in(0,R]} \left( ⨍_{B_\rho(x,r)} \left| f - ⨍_{B_\rho(x,r)} f\, d\mu \right|^p d\mu \right)^{\frac{1}{p}}, \\ \text{and} \quad ⨍_{B_\rho(x,r)} f\, d\mu &:= \frac{1}{\mu(B_\rho(x, r))} \int_{B_\rho(x,r)} f\, d\mu. \end{aligned} \tag{D.4}$$

Following [Sarason 1975], if $\mathrm{UC}(\Sigma, \mu)$ stands for the space of uniformly continuous functions on $X$, we introduce $\mathrm{VMO}(\Sigma, \mu)$, the space of functions of vanishing mean oscillations on $\Sigma$, where

$$\mathrm{VMO}(\Sigma, \mu) \text{ is the closure of } \mathrm{UC}(\Sigma, \mu) \cap \mathrm{BMO}(\Sigma, \mu) \text{ in } \mathrm{BMO}(\Sigma, \mu). \tag{D.5}$$

We have the following useful equivalent characterization of VMO on compact spaces of homogeneous type. To state it, we denote by $\mathscr{C}^\alpha(\Sigma, \rho)$ the space of real-valued Hölder functions of order $\alpha > 0$ on the quasimetric space $(\Sigma, \rho)$. That is, $\mathscr{C}^\alpha(\Sigma, \rho)$ is the collection of all real-valued functions $f$ on $\Sigma$ with the property that

$$\|f\|_{\mathscr{C}^\alpha(\Sigma,\rho)} := \sup_{x\in\Sigma} |f(x)| + \sup_{x,y\in\Sigma,\, x\neq y} \frac{|f(x) - f(y)|}{\rho(x, y)^\alpha} < +\infty. \tag{D.6}$$

For further reference, let us also set

$$\mathscr{C}^\alpha_0(X, \rho) := \{f \in \mathscr{C}^\alpha(\Sigma, \rho) : \operatorname{supp} f \text{ bounded}\}. \tag{D.7}$$

The following two propositions contain results proved in [Hofmann et al. 2010; Mitrea et al. 2013].

**Proposition D.1.** *Assume that* $(\Sigma, \rho, \mu)$ *is a compact space of homogeneous type. Then*

$$\mathrm{VMO}(\Sigma, \mu) \text{ is the closure of } \mathscr{C}^{\alpha}(\Sigma, \rho) \cap \mathrm{BMO}(\Sigma, \mu) \text{ in } \mathrm{BMO}(\Sigma, \mu) \tag{D.8}$$

*for every* $\alpha \in \mathbb{R}$ *such that*

$$0 < \alpha \leq \left[\log_2\left(\sup_{\substack{x,y,z\in\Sigma \\ \text{not all equal}}} \frac{\rho(x, y)}{\max\{\rho(x, z), \rho(z, y)\}}\right)\right]^{-1}. \tag{D.9}$$

**Proposition D.2.** *Let* $(\Sigma, \rho, \mu)$ *be a space of homogeneous type. Then, for each* $p \in [1, \infty)$,

$$\begin{aligned} \mathrm{dist}_{\mathrm{BMO}}(f, \mathrm{VMO}(\Sigma, \mu)) &\approx \limsup_{r\to 0^+}\left\{\sup_{x\in\Sigma} ⨍_{B_\rho(x,r)} ⨍_{B_\rho(x,r)} |f(y) - f(z)|^p \, d\mu(y) \, d\mu(z)\right\}^{\frac{1}{p}} \\ &\approx \limsup_{r\to 0^+}\left\{\sup_{x\in\Sigma} ⨍_{B_\rho(x,r)} \left|f - ⨍_{B_\rho(x,r)} f \, d\mu\right|^p d\mu\right\}^{\frac{1}{p}} \end{aligned} \tag{D.10}$$

*uniformly for* $f \in \mathrm{BMO}(\Sigma, \mu)$ (*i.e.*, *the constants do not depend on* $f$), *where the distance is measured in the* BMO *norm. In particular, for each* $p \in [1, \infty)$,

$$\mathrm{dist}_{\mathrm{BMO}}(f, \mathrm{VMO}(\Sigma, \mu)) \approx \lim_{R\to 0^+} M_p(f; R) \quad \text{uniformly for } f \in \mathrm{BMO}(\Sigma, \mu), \tag{D.11}$$

*where* $M_p(f; R)$ *is defined as in* (D.4). *Moreover, for each function* $f \in \mathrm{BMO}(\Sigma, \mu)$ *and each* $p \in [1, \infty)$,

$$f \in \mathrm{VMO}(\Sigma, \mu) \iff \lim_{r\to 0^+}\left\{\sup_{x\in\Sigma} ⨍_{B_\rho(x,r)} \left|f - ⨍_{B_\rho(x,r)} f \, d\mu\right|^p d\mu\right\}^{\frac{1}{p}} = 0. \tag{D.12}$$

For future purposes, we find it convenient to restate (D.11) in a slightly different form. More specifically, in the context of Proposition D.2, given $f \in L^2_{\mathrm{loc}}(\Sigma, \mu)$, $x \in \Sigma$ and $R > 0$, we set

$$\|f\|_*(B_\rho(x, R)) := \sup_{B\subseteq B_\rho(x,R)} \left(⨍_B |f - f_B|^2 \, d\mu\right)^{\frac{1}{2}}, \tag{D.13}$$

where the supremum is taken over all $\rho$-balls $B$ included in $B_\rho(x, R)$ and $f_B := \mu(B)^{-1} \int_B f \, d\mu$. It is then clear from the definitions that

$$\sup_{x\in\Sigma} \|f\|_*(B_\rho(x, R)) \approx M_2(f; R). \tag{D.14}$$

Consequently, (D.11) yields:

**Corollary D.3.** *With the above notation and conventions*,

$$\lim_{R\to 0^+}\left[\sup_{x\in\Sigma} \|f\|_*(B_\rho(x, R))\right] \approx \mathrm{dist}_{\mathrm{BMO}}(f, \mathrm{VMO}(\Sigma, \mu)) \tag{D.15}$$

*uniformly for* $f \in \mathrm{BMO}(\Sigma, \mu)$.

We continue by translating Proposition C.4 (which was formulated in the Euclidean context) to spaces of homogeneous type.

**Proposition D.4.** *Assume that* $(\Sigma, \rho, \mu)$ *is a space of homogeneous type. Then there exists a constant* $C \in (0, \infty)$ *such that*

$$\begin{aligned} &\mathrm{dist}_{\mathrm{BMO}}(fg, \mathrm{VMO}(\Sigma, \mu)) \\ &\qquad \leq C\|f\|_{L^\infty(\Sigma,\mu)}\,\mathrm{dist}_{\mathrm{BMO}}(g, \mathrm{VMO}(\Sigma, \mu)) + C\|g\|_{L^\infty(\Sigma,\mu)}\,\mathrm{dist}_{\mathrm{BMO}}(f, \mathrm{VMO}(\Sigma, \mu)), \end{aligned} \tag{D.16}$$

*for any* $f, g \in L^\infty(\Sigma, \mu)$, *where all distances are considered in the space* $\mathrm{BMO}(\Sigma, \mu)$.

*Moreover*,

$$\mathrm{VMO}(\Sigma, \mu) \cap L^\infty(\Sigma, \mu) \quad \text{is a closed } C^* \text{ subalgebra of } L^\infty(\Sigma, \mu), \tag{D.17}$$

*and*

$$f \in \mathrm{VMO}(\Sigma, \mu) \cap L^\infty(\Sigma, \mu) \text{ and } \frac{1}{f} \in L^\infty(\Sigma, \mu) \quad \Longrightarrow \quad \frac{1}{f} \in \mathrm{VMO}\,(\Sigma, \mu) \cap L^\infty(\Sigma, \mu). \tag{D.18}$$

*Proof.* Note that (D.16) implies (D.17) and also (D.18), via the same type of argument used to establish (C.16). As such, it suffices to prove (D.16). To this end, if $f, g \in L^\infty(\Sigma, \mu)$ then, for any $x \in \Sigma$, $r > 0$ and $y, z \in B_\rho(x, r)$, we have

$$\begin{aligned} |f(y)g(y) - f(z)g(z)| &\leq |f(y)||g(y) - g(z)| + |g(z)||f(y) - f(z)| \\ &\leq \|f\|_{L^\infty(X,\mu)}|g(y) - g(z)| + \|g\|_{L^\infty(X,\mu)}|f(y) - f(z)|. \end{aligned} \tag{D.19}$$

With this in hand, (D.16) follows with the help of the first equivalence in (D.10). ☐

Another useful result pertains to the manner in which one can control the distance to VMO under composition by a Lipschitz function.

**Proposition D.5.** *Assume that* $(\Sigma, \rho, \mu)$ *is a space of homogeneous type. Let* $F : \mathbb{R}^m \to \mathbb{R}$ *be a Lipschitz function. Then there exists a constant* $C \in (0, \infty)$ *such that, for every* $f : \Sigma \to \mathbb{R}^m$ *with components in* $\mathrm{BMO}(\Sigma, \mu)$,

$$\mathrm{dist}_{\mathrm{BMO}}(F \circ f, \mathrm{VMO}(\Sigma, \mu)) \leq C\|\nabla F\|_{L^\infty(\mathbb{R}^m)}\,\mathrm{dist}_{\mathrm{BMO}}(f, \mathrm{VMO}(\Sigma, \mu)). \tag{D.20}$$

*where the distances are considered in the space* $\mathrm{BMO}(\Sigma, \mu)$. *In particular*,

$$f \in \mathrm{VMO}(\Sigma, \mu) \quad \Longrightarrow \quad F \circ f \in \mathrm{VMO}(\Sigma, \mu). \tag{D.21}$$

*Proof.* Fix $x \in \Sigma$ and $r > 0$, arbitrary. Using the fact that $F$ is Lipschitz we may then estimate, for every $y, z \in B_\rho(x, r)$,

$$|F(f(y)) - F(f(z))| \leq \|\nabla F\|_{L^\infty(\mathbb{R}^m)}|f(y) - f(x)|. \tag{D.22}$$

Then the desired conclusion readily follows from this and the first equivalence in (D.10). ☐

**Appendix E.** ***On the class of* Lip ∩ vmo₁ *domains.*** The starting point in this appendix is the following result.

**Lemma E.1.** *Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function, with graph*

$$\Sigma := \{(x, \varphi(x)) : x \in \mathbb{R}^n\} \subset \mathbb{R}^{n+1}. \tag{E.1}$$

*Set $\mu := \mathcal{H}^n \lfloor \Sigma$, where $\mathcal{H}^n$ is the $n$-dimensional Hausdorff measure in $\mathbb{R}^{n+1}$. Then*

$$f \in \mathrm{VMO}(\Sigma, \mu) \quad \Longleftrightarrow \quad f(\cdot\,, \varphi(\cdot)) \in \mathrm{VMO}(\mathbb{R}^n). \tag{E.2}$$

*Proof.* For each given point $X = (x, \varphi(x)) \in \Sigma$ with $x \in \mathbb{R}^n$ and each given radius $r > 0$, set $\Delta(X, r) := \{Y \in \Sigma : |Y - X| < r\}$. Then fix $X_0 = (x_0, \varphi(x_0)) \in \Sigma$ for some $x_0 \in \mathbb{R}^n$ and pick some $r > 0$. Consider $c := ⨍_{B(x_0,r)} f(x, \varphi(x))\, dx$. Then

$$\begin{aligned}
&⨍_{\Delta(X_0,r)} \Big| f - ⨍_{\Delta(X_0,r)} f \, d\mu \Big| d\mu \\
&\qquad = ⨍_{\Delta(X_0,r)} \Big| (f - c) - ⨍_{\Delta(X_0,r)} (f - c)\, d\mu \Big| d\mu \\
&\qquad \le 2 ⨍_{\Delta(X_0,r)} |f - c|\, d\mu \\
&\qquad = 2 ⨍_{\{x \in \mathbb{R}^n : |x - x_0|^2 + (\varphi(x) - \varphi(x_0))^2 < r^2\}} |f(x, \varphi(x)) - c| \sqrt{1 + |\nabla\varphi(x)|^2}\, dx \\
&\qquad \le C ⨍_{\{x \in \mathbb{R}^n : |x - x_0| < r\}} |f(x, \varphi(x)) - c|\, dx.
\end{aligned} \tag{E.3}$$

Bearing in mind the significance of $c$, the left-pointing implication in (E.2) follows from (D.12) (with $p = 1$). For the opposite implication, pick $c' := ⨍_{\Delta(X_0,r)} f\, d\mu$. Then, for some sufficiently large $M > 0$ depending on the Lipschitz constant of $\varphi$, we have

$$\begin{aligned}
&⨍_{B(x_0,r)} \Big| f(x, \varphi(x)) - ⨍_{B(x_0,r)} f(y, \varphi(y))\, dy \Big|\, dx \\
&\qquad \le 2 ⨍_{\{x \in \mathbb{R}^n : |x - x_0| < r\}} |f(x, \varphi(x)) - c'|\, dx \\
&\qquad \le C ⨍_{\{x \in \mathbb{R}^n : |x - x_0|^2 + (\varphi(x) - \varphi(x_0))^2 < (Mr)^2\}} |f(x, \varphi(x)) - c'| \sqrt{1 + |\nabla\varphi(x)|^2}\, dx \\
&\qquad \le C ⨍_{\Delta(X_0,r)} \Big| f - ⨍_{\Delta(X_0,r)} f\, d\mu \Big| d\mu.
\end{aligned} \tag{E.4}$$

Based on this and (D.12), the right-pointing implication in (E.2) now follows. □

In turn, Lemma E.1 is an important ingredient in the proof of the following result:

**Lemma E.2.** *Assume that $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a Lipschitz function, and let $\Sigma$ as in* (E.1) *denote its graph. Set $\mu := \mathcal{H}^n \lfloor \Sigma$, where $\mathcal{H}^n$ is the $n$-dimensional Hausdorff measure in $\mathbb{R}^{n+1}$, and let $\nu = (\nu_1, \dots, \nu_{n+1})$*

*stand for the unit normal to* $\Sigma$ *(defined* $\nu$*-a.e.). Then*

$$\nu_j \in \mathrm{VMO}(\Sigma,\mu) \quad \text{for } 1 \le j \le n+1 \quad \Longleftrightarrow \quad \partial_j\varphi \in \mathrm{VMO}(\mathbb{R}^n) \quad \text{for } 1 \le j \le n. \tag{E.5}$$

*Proof.* Recall that the components $\nu_j : \Sigma \to \mathbb{R}$ of the unit normal to the Lipschitz surface $\Sigma$ satisfy

$$\nu_j(x,\varphi(x)) = \begin{cases} \partial_j\varphi(x)/\sqrt{1+|\nabla\varphi(x)|^2} & \text{if } 1 \le j \le n, \\ -1/\sqrt{1+|\nabla\varphi(x)|^2} & \text{if } j = n+1 \end{cases} \tag{E.6}$$

for a.e. $x \in \mathbb{R}^n$. As regards (E.5), assume first that

$$\partial_j\varphi \in \mathrm{VMO}(\mathbb{R}^n) \quad \text{for each } j \in \{1,\dots,n\} \tag{E.7}$$

and consider the functions $F_j : \mathbb{R}^n \to \mathbb{R}$, $1 \le j \le n+1$, given by

$$F_j(x) := \begin{cases} x_j/\sqrt{1+|x|^2} & \text{if } 1 \le j \le n, \\ -1/\sqrt{1+|x|^2} & \text{if } j = n+1 \end{cases} \tag{E.8}$$

for each $x = (x_1,\dots,x_n) \in \mathbb{R}^n$. A straightforward computation gives that there exists a dimensional constant such that, for every $x \in \mathbb{R}^n$,

$$|\nabla F_j(x)| \le \begin{cases} C_n/\sqrt{1+|x|^2} & \text{if } 1 \le j \le n, \\ C_n/(1+|x|^2) & \text{if } j = n+1. \end{cases} \tag{E.9}$$

In particular, each function $F_j : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz. Upon noting from (E.6) and (E.8) that $\nu_j(x,\varphi(x)) = F_j(\nabla\varphi(x))$ for a.e. $x \in \mathbb{R}^n$, this implies, in concert with (E.7) and (D.21), that $\nu_j(\cdot,\varphi(\cdot)) \in \mathrm{VMO}(\mathbb{R}^n)$ for each $j \in \{1,\dots,n+1\}$. Having established this, we may then conclude that $\nu_j \in \mathrm{VMO}(\Sigma,\mu)$ for $1 \le j \le n+1$ by invoking Lemma E.1. This proves the left-pointing implication in (E.5).

In the opposite direction, assume

$$\nu_j \in \mathrm{VMO}(\Sigma,\mu) \quad \text{for each } j \in \{1,\dots,n+1\}. \tag{E.10}$$

Then Lemma E.1 gives

$$\nu_j(\cdot,\varphi(\cdot)) \in \mathrm{VMO}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n) \quad \text{for each } j \in \{1,\dots,n+1\}. \tag{E.11}$$

Since, from (E.6) and the fact that $\varphi$ is Lipschitz, we have

$$1/\nu_{n+1}(\cdot,\varphi(\cdot)) \in L^\infty(\mathbb{R}^n), \tag{E.12}$$

we deduce from (D.18), (E.11) with $j = n+1$, and (E.12) that

$$1/\nu_{n+1}(\cdot,\varphi(\cdot)) \in \mathrm{VMO}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n). \tag{E.13}$$

Given that $\mathrm{VMO}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ is an algebra (see (D.17) in Proposition D.4), it follows from (E.11) and (E.13) that

$$\nu_j(\cdot,\varphi(\cdot))/\nu_{n+1}(\cdot,\varphi(\cdot)) \in \mathrm{VMO}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n) \quad \text{for each } j \in \{1,\dots,n\}. \tag{E.14}$$

In light of (E.6) this ultimately entails $\partial_j\varphi \in \mathrm{VMO}(\mathbb{R}^n)$ for $1 \le j \le n$, as wanted. □

We are now in a position to define the class of $\mathrm{Lip} \cap \mathrm{vmo}_1$ domains.

**Definition E.3.** Assume that $C \in (0, \infty)$ and let $\Omega$ be a nonempty, open subset of $\mathbb{R}^n$, with diameter at most $C$. One calls $\Omega$ a bounded Lipschitz domain, with Lipschitz character controlled by $C$, if there exists $r \in (0, C)$ with the property that for every $x_0 \in \partial\Omega$ one can find a rigid transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ and a Lipschitz function $\varphi : \mathbb{R}^{n-1} \to \mathbb{R}$ with $\|\nabla\varphi\|_{L^\infty(\mathbb{R}^{n-1})} \le C$ such that

$$T(\Omega \cap B(x_0, r)) = T(B(x_0, r)) \cap \{(x', x_n) \in \mathbb{R}^{n-1} \times \mathbb{R} : x_n > \varphi(x')\}. \tag{E.15}$$

Whenever this is the case, call $\phi(x') := (x', \varphi(x'))$ a coordinate chart for $\partial\Omega$.

If, in addition, $\partial_j \varphi \in \mathrm{vmo}(\mathbb{R}^{n-1})$ for each $j \in \{1, \dots, n-1\}$, then we shall say that $\Omega$ is a bounded $\mathrm{Lip} \cap \mathrm{vmo}_1$ domain.

Both the class of Lipschitz domains and the class of $\mathrm{Lip} \cap \mathrm{vmo}_1$ domains may be naturally defined in the manifold setting by working in local coordinates, in a similar fashion as above (see also the discussion in [Hofmann et al. 2007]).

We conclude this appendix by proving the following characterization of the class of $\mathrm{Lip} \cap \mathrm{vmo}_1$ domains:

**Proposition E.4.** *Let $\Omega$ be a Lipschitz domain with outward unit normal $\nu$. Then*

$$\nu \in \mathrm{vmo}(\partial\Omega) \quad \Longleftrightarrow \quad \Omega \text{ is a } \mathrm{Lip} \cap \mathrm{vmo}_1 \text{ domain.} \tag{E.16}$$

*Proof.* This is a consequence of Lemma E.2 and definitions. □

## References


[Bourdaud et al. 2002] G. Bourdaud, M. Lanza de Cristoforis, and W. Sickel, "Functional calculus on BMO and related spaces", *J. Funct. Anal.* **189**:2 (2002), 515–538. MR 2003c:47059 Zbl 1007.47028

[Calderón 1977] A.-P. Calderón, "Cauchy integrals on Lipschitz curves and related operators", *Proc. Nat. Acad. Sci. U.S.A.* **74**:4 (1977), 1324–1327. MR 57 #6445 Zbl 0373.44003

[Calderón 1985] A. P. Calderón, "Boundary value problems for the Laplace equation in Lipschitzian domains", pp. 33–48 in *Recent progress in Fourier analysis* (El Escorial, 1983), edited by I. Peral and J. L. Rubio de Francia, North-Holland Math. Stud. **111**, North-Holland, Amsterdam, 1985. MR 87k:35062 Zbl 0608.31001

[Chiarenza et al. 1991] F. Chiarenza, M. Frasca, and P. Longo, "Interior $W^{2,p}$ estimates for nondivergence elliptic equations with discontinuous coefficients", *Ricerche Mat.* **40**:1 (1991), 149–168. MR 93k:35051 Zbl 0772.35017

[Coifman and Weiss 1977] R. R. Coifman and G. Weiss, "Extensions of Hardy spaces and their use in analysis", *Bull. Amer. Math. Soc.* **83**:4 (1977), 569–645. MR 56 #6264 Zbl 0358.30023

[Coifman et al. 1976] R. R. Coifman, R. Rochberg, and G. Weiss, "Factorization theorems for Hardy spaces in several variables", *Ann. of Math.* (2) **103**:3 (1976), 611–635. MR 54 #843 Zbl 0326.32011

[Coifman et al. 1982] R. R. Coifman, A. McIntosh, and Y. Meyer, "L'intégrale de Cauchy définit un opérateur borné sur $L^2$ pour les courbes lipschitziennes", *Ann. of Math.* (2) **116**:2 (1982), 361–387. MR 84m:42027 Zbl 0497.42012

[Dautray and Lions 1990] R. Dautray and J.-L. Lions, *Mathematical analysis and numerical methods for science and technology, III: Spectral theory and applications*, Springer, Berlin, 1990. MR 91h:00004a Zbl 0944.47002

[Fabes et al. 1978] E. B. Fabes, M. Jodeit, Jr., and N. M. Rivière, "Potential techniques for boundary value problems on $C^1$-domains", *Acta Math.* **141**:3-4 (1978), 165–186. MR 80b:31006 Zbl 0402.31009

[Fabes et al. 1998] E. Fabes, O. Mendez, and M. Mitrea, "Boundary layers on Sobolev–Besov spaces and Poisson's equation for the Laplacian in Lipschitz domains", *J. Funct. Anal.* **159**:2 (1998), 323–368. MR 99j:35036 Zbl 0930.35045

[Grüter and Widman 1982] M. Grüter and K.-O. Widman, "The Green function for uniformly elliptic equations", *Manuscripta Math.* **37**:3 (1982), 303–342. MR 83h:35033 Zbl 0485.35031

[Hofmann 1994] S. Hofmann, "On singular integrals of Calderón-type in $\mathbf{R}^n$, and BMO", *Rev. Mat. Iberoamericana* **10**:3 (1994), 467–505. MR 96d:42019 Zbl 0874.42011

[Hofmann et al. 2007] S. Hofmann, M. Mitrea, and M. Taylor, "Geometric and transformational properties of Lipschitz domains, Semmes–Kenig–Toro domains, and other classes of finite perimeter domains", *J. Geom. Anal.* **17**:4 (2007), 593–647. MR 2009b:49098 Zbl 1142.49021

[Hofmann et al. 2010] S. Hofmann, M. Mitrea, and M. Taylor, "Singular integrals and elliptic boundary problems on regular Semmes–Kenig–Toro domains", *Int. Math. Res. Not.* **2010**:14 (2010), 2567–2865. MR 2011h:42021 Zbl 1221.31010

[Kalton and Mitrea 1998] N. Kalton and M. Mitrea, "Stability results on interpolation scales of quasi-Banach spaces and applications", *Trans. Amer. Math. Soc.* **350**:10 (1998), 3903–3922. MR 98m:46094 Zbl 0902.46002

[Kenig and Pipher 1988] C. E. Kenig and J. Pipher, "The oblique derivative problem on Lipschitz domains with $L^p$ data", *Amer. J. Math.* **110**:4 (1988), 715–737. MR 89i:35047 Zbl 0676.35019

[Kenig and Toro 1997] C. E. Kenig and T. Toro, "Harmonic measure on locally flat domains", *Duke Math. J.* **87**:3 (1997), 509–551. MR 98k:31010 Zbl 0878.31002

[Lewis et al. 1993] J. E. Lewis, R. Selvaggi, and I. Sisto, "Singular integral operators on $C^1$ manifolds", *Trans. Amer. Math. Soc.* **340**:1 (1993), 293–308. MR 94a:58194 Zbl 0786.35156

[Melnikov and Verdera 1995] M. S. Melnikov and J. Verdera, "A geometric proof of the $L^2$ boundedness of the Cauchy integral on Lipschitz graphs", *Internat. Math. Res. Notices* **1995**:7 (1995), 325–331. MR 96f:45011 Zbl 0923.42006

[Mitrea 2001] M. Mitrea, "Generalized Dirac operators on nonsmooth manifolds and Maxwell's equations", *J. Fourier Anal. Appl.* **7**:3 (2001), 207–256. MR 2002k:58046 Zbl 0979.31006

[Mitrea and Taylor 1999] M. Mitrea and M. Taylor, "Boundary layer methods for Lipschitz domains in Riemannian manifolds", *J. Funct. Anal.* **163**:2 (1999), 181–251. MR 2000b:35050 Zbl 0930.58014

[Mitrea and Taylor 2000] M. Mitrea and M. Taylor, "Potential theory on Lipschitz domains in Riemannian manifolds: Sobolev–Besov space results and the Poisson problem", *J. Funct. Anal.* **176**:1 (2000), 1–79. MR 2002e:58044 Zbl 0968.58023

[Mitrea et al. 2001] D. Mitrea, M. Mitrea, and M. Taylor, "Layer potentials, the Hodge Laplacian, and global boundary problems in nonsmooth Riemannian manifolds", *Mem. Amer. Math. Soc.* **150**:713 (2001), x+120. MR 2002g:58026 Zbl 1003.35001

[Mitrea et al. 2013] D. Mitrea, I. Mitrea, M. Mitrea, and S. Monniaux, *Groupoid metrization theory, with applications to analysis on quasimetric spaces and functional analysis*, Birkhäuser/Springer, New York, 2013. MR 2987059 Zbl 1269.46002

[Mitrea et al. 2015] I. Mitrea, M. Mitrea, and M. Taylor, "Cauchy integrals, Calderón projectors, and Toeplitz operators on uniformly rectifiable domains", *Adv. Math.* **268** (2015), 666–757. MR 3276607 Zbl 06370015

[Mitrea et al. $\geq$ 2015] I. Mitrea, M. Mitrea, and M. Taylor, "The Riemann–Hilbert problem, Cauchy integrals, and Toeplitz operators on uniformly rectifiable domains", in preparation.

[Pipher 1987] J. Pipher, "Oblique derivative problems for the Laplacian in Lipschitz domains", *Rev. Mat. Iberoamericana* **3**:3-4 (1987), 455–472. MR 90g:35040 Zbl 0686.35028

[Rademacher 1919] H. Rademacher, "Über partielle und totale differenzierbarkeit von Funktionen mehrerer Variabeln und über die Transformation der Doppelintegrale", *Math. Ann.* **79**:4 (1919), 340–359. MR 1511935 Zbl 47.0243.01

[Sarason 1975] D. Sarason, "Functions of vanishing mean oscillation", *Trans. Amer. Math. Soc.* **207** (1975), 391–405. MR 51 #13690 Zbl 0319.42006

[Semmes 1991] S. Semmes, "Chord-arc surfaces with small constant, I", *Adv. Math.* **85**:2 (1991), 198–223. MR 93d:42019a Zbl 0733.42015

[Šneĭberg 1974] I. J. Šneĭberg, "Spectral properties of linear operators in interpolation families of Banach spaces", *Mat. Issled.* **9**:2(32) (1974), 214–229, 254–255. In Russian. MR 58 #30362 Zbl 0314.46033

[Stein 1970] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series **30**, Princeton University Press, Princeton, N.J., 1970. MR 44 #7280 Zbl 0207.13501

[Taylor 1996] M. E. Taylor, *Partial differential equations, II: Qualitative studies of linear equations*, Applied Mathematical Sciences **116**, Springer, 1996. MR 98b:35003 Zbl 1206.35003

[Taylor 1997] M. E. Taylor, "Microlocal analysis on Morrey spaces", pp. 97–135 in *Singularities and oscillations* (Minneapolis, MN, 1994/1995), edited by J. Rauch and M. Taylor, IMA Vol. Math. Appl. **91**, Springer, New York, 1997. MR 99a:35010 Zbl 0904.58012

[Taylor 2000] M. E. Taylor, *Tools for PDE: pseudodifferential operators, paradifferential operators, and layer potentials*, Mathematical Surveys and Monographs **81**, American Mathematical Society, Providence, RI, 2000. MR 2001g:35004 Zbl 0963.35211

[Taylor 2009] M. Taylor, "Hardy spaces and BMO on manifolds with bounded geometry", *J. Geom. Anal.* **19**:1 (2009), 137–190. MR 2009k:58053 Zbl 1189.46030

[Verchota 1984] G. Verchota, "Layer potentials and regularity for the Dirichlet problem for Laplace's equation in Lipschitz domains", *J. Funct. Anal.* **59**:3 (1984), 572–611. MR 86e:35038 Zbl 0589.31005

Received 12 Mar 2014. Accepted 5 Jan 2015.

STEVE HOFMANN: hofmanns@missouri.edu
*Department of Mathematics, University of Missouri at Columbia, Math. Building, University of Missouri, Columbia, MO 65211, United States*

MARIUS MITREA: mitream@missouri.edu
*Department of Mathematics, University of Missouri at Columbia, Math. Building, University of Missouri, Columbia, MO 65211, United States*

MICHAEL E. TAYLOR: met@math.unc.edu
*Department of Mathematics, University of North Carolina, Phillips Hall, UNC, Chapel Hill, NC 27599-3250, United States*

ANALYSIS AND PDE
Vol. 8, No. 1, 2015
dx.doi.org/10.2140/apde.2015.8.183

# CRITERIA FOR HANKEL OPERATORS TO BE SIGN-DEFINITE

DIMITRI R. YAFAEV

We show that the total multiplicities of negative and positive spectra of a self-adjoint Hankel operator $H$ in $L^2(\mathbb{R}_+)$ with integral kernel $h(t)$ and of the operator of multiplication by the inverse Laplace transform of $h(t)$, the distribution $\sigma(\lambda)$, coincide. In particular, $\pm H \geq 0$ if and only if $\pm\sigma(\lambda) \geq 0$. To construct $\sigma(\lambda)$, we suggest a new method of inversion of the Laplace transform in appropriate classes of distributions. Our approach directly applies to various classes of Hankel operators. For example, for Hankel operators of finite rank, we find an explicit formula for the total numbers of their negative and positive eigenvalues.

## 1. Introduction

**1.1.** Hankel operators can be defined as integral operators

$$(Hf)(t) = \int_0^\infty h(t+s) f(s)\, ds \tag{1-1}$$

in the space $L^2(\mathbb{R}_+)$ with kernels $h$ that depend on the sum of variables only. Of course $H$ is symmetric if $h(t) = \overline{h(t)}$. In the fundamental paper [Megretskiĭ et al. 1995], A. V. Megretskiĭ, V. V. Peller, and S. R. Treil characterized the spectra of all bounded self-adjoint Hankel operators by a certain balance between the positive and negative parts of their spectra. The result of [Megretskiĭ et al. 1995] applies to *all* Hankel operators, and so it does not allow one to distinguish spectral properties of particular operators.

The cases where Hankel operators can be explicitly diagonalized are very scarce. We mention here the kernels $h(t) = t^{-1}$ [Carleman 1923], $h(t) = (t+1)^{-1}$ [Mehler 1881], and $h(t) = t^{-1}e^{-t}$ [Magnus 1950; Rosenblum 1958a; 1958b]. These kernels are treated in a unified way in [Yafaev 2010], where some new examples are also considered.

Our goal here is to find explicit expressions for the total numbers $N_+(H)$ and $N_-(H)$ of (strictly) positive and negative eigenvalues of self-adjoint Hankel operators $H$. Actually, we show that $N_\pm(H) = N_\pm(\Sigma)$, where $\Sigma$ is the operator[1] of multiplication by the function (distribution) $\sigma(\lambda)$ obtained through the inversion of the Laplace transform

$$h(t) = \int_0^\infty e^{-t\lambda}\sigma(\lambda)\, d\lambda. \tag{1-2}$$

We call $\sigma(\lambda)$ the *sigma function* of a Hankel operator $H$ or of its kernel $h(t)$.

*MSC2010:* primary 47A40; secondary 47B25.
*Keywords:* Hankel operators, convolutions, necessary and sufficient conditions for positivity, sign function, operators of finite rank, the Carleman operator and its perturbations.

[1] To be more precise, we consider the quadratic forms $(Hf, f)$ and $(\Sigma\varphi, \varphi)$ instead of the operators $H$ and $\Sigma$.

In particular, we obtain necessary and sufficient conditions for the sign-definiteness of Hankel operators. Indeed, it formally follows from (1-2) that

$$(Hf, f) = \int_0^\infty d\lambda\, \sigma(\lambda) \left| \int_0^\infty f(t) e^{-t\lambda}\, dt \right|^2,$$

and hence $\pm H \geq 0$ if and only if $\pm\sigma \geq 0$. We usually discuss conditions for $H \geq 0$, but of course replacing $H$ by $-H$ we obtain conditions for $H \leq 0$. Note that positive[2] distributions are always given by some measures, so that for positive Hankel operators $H$, representation (1-2) reduces to

$$h(t) = \int_0^\infty e^{-t\lambda}\, dm(\lambda), \tag{1-3}$$

where $dm(\lambda)$ is a (positive) measure on $[0, \infty)$.

**1.2.** If a function $\sigma(\lambda)$ is sufficiently regular (for example, bounded), then its Laplace transform (1-2) is analytic in the right half-plane and satisfies certain decay conditions for $|t| \to \infty$. For example, such simple functions as the characteristic functions of intervals or $h(t) = e^{-t^2}$ do not satisfy these conditions. A regular function $\sigma(\lambda)$ can be recovered from its Laplace transform $h(t)$ by the integral of $h(a + i\tau)$, $a > 0$, over $\tau \in \mathbb{R}$; alternatively, it can also be recovered (see, for example, [Paley and Wiener 1934, Section 13]) from the values of $h(t)$ for $t > 0$. These methods are not sufficient for our purposes since, for example, for $h(t) = t^k e^{-\alpha t}$, $k \in \mathbb{Z}_+$, $\operatorname{Re}\alpha > 0$ (such Hankel operators have rank $k$), the corresponding function

$$\sigma(\lambda) = \delta^{(k)}(\lambda - \alpha) \tag{1-4}$$

($\delta(\cdot)$ is the Dirac function) is a highly singular distribution, especially if $\operatorname{Im}\alpha \neq 0$.

Thus we are led to a solution of (1-2) for $\sigma(\lambda)$ in a class of distributions. Put

$$b(\xi) = \frac{1}{2\pi} \frac{\int_0^\infty h(t) t^{-i\xi}\, dt}{\int_0^\infty e^{-t} t^{-i\xi}\, dt}, \tag{1-5}$$

and let $s(x) = \sqrt{2\pi}\,(\Phi^* b)(x)$, where $\Phi$ is the Fourier transform. We show that the function

$$\sigma(\lambda) := s(-\ln\lambda) \tag{1-6}$$

satisfies (1-2). We call $b(\xi)$ the *b-function* and $s(x)$ the *sign function* (or *s-function*) of the Hankel operator $H$ (or of its integral kernel $h(t)$). The sigma function $\sigma(\lambda)$ differs from $s(x)$ by a change of variables only. In specific examples we consider, functions $b(\xi)$ and $s(x)$ may be of a quite different nature. For instance, $s(x)$ may be a polynomial or, on the contrary, it may be a highly singular distribution such as a combination of delta functions and their derivatives. We emphasize that all our formulas are understood in the sense of distributions and of course no analyticity of $h(t)$ is required. From a purely formal point of view, our method of inversion of the Laplace transform is not too far from one of the methods described in [Paley and Wiener 1934], but the classes of functions (distributions) are quite different.

[2] We use the term "positive" instead of the more precise but lengthy term "nonnegative".

The precise meaning of formula (1-5) requires some discussion. Observe that the denominator in (1-5) coincides with the numerator for the special case $h(t)=e^{-t}$. It equals $\Gamma(1-i\xi)$, and hence exponentially tends to zero as $|\xi|\to\infty$. Therefore $b(\xi)$ is a "nice" function of $\xi$ only under very restrictive assumptions on the kernel $h(t)$. Thus, to cover natural examples, we are obliged to work with distributions $b(\xi)$ and $s(x)$. The choice of appropriate spaces of distributions is also very important. The Schwartz space $\mathcal{S}(\mathbb{R})'$ is too restrictive for our purposes, which is seen already in the example of finite-rank Hankel operators. In order to be able to divide in (1-5) by an exponentially decaying function, we assume that the numerator belongs to the class of distributions $C_0^\infty(\mathbb{R})'$. This means that the Fourier transform of the function $\theta(x)=e^x h(e^x)$ belongs to $C_0^\infty(\mathbb{R})'$, that is, $\theta$ belongs to the space $\mathcal{Z}'$ dual to the space $\mathcal{Z}=\mathcal{Z}(\mathbb{R})$ of analytic test functions. The class of distributions $h(t)$ such that the corresponding function $\theta$ is in $\mathcal{Z}'$ will be denoted by $\mathcal{Z}'_+$. It follows from (1-5) that $b\in C_0^\infty(\mathbb{R})'$ and $s\in\mathcal{Z}'$ if $h\in\mathcal{Z}'_+$.

A remarkable circumstance is that, in these classes, there is a one-to-one correspondence between kernels of Hankel operators and their sigma functions. To be precise, let us put $h^\natural(\lambda)=\lambda^{-1}\sigma(\lambda)$. We show that $h\in\mathcal{Z}'_+$ if and only if $h^\natural\in\mathcal{Z}'_+$, and the correspondence $h\mapsto h^\natural$ is a continuous one-to-one mapping of $\mathcal{Z}'_+$ onto itself. As an example, note that although the functions $h(t)=t^k e^{-\alpha t}$ and $h^\natural(\lambda)=\lambda^{-1}\delta^{(k)}(\lambda-\alpha)$ are of a completely different nature, both of them belong to the class $\mathcal{Z}'_+$.

In the case $h\in L^1_{\rm loc}(\mathbb{R}_+)$, the condition $h\in\mathcal{Z}'_+$ means that

$$\int_0^\infty |h(t)|(1+|\ln t|)^{-\kappa}\,dt<\infty \tag{1-7}$$

for some $\kappa$. Condition (1-7) is also quite general and does not require that the corresponding Hankel operator be bounded. For example, it admits kernels

$$h(t)=P(\ln t)t^{-1}, \tag{1-8}$$

where $P(x)$ is an arbitrary polynomial. Note that Hankel operators with such kernels are bounded for $P(x)=\text{const}$ only.

**1.3.** Our study of spectral properties of Hankel operators $H$ relies on their reduction to the operators $S$ of multiplication by the corresponding sign functions. This reduction is given by a transformation which is, in a suitable sense, invertible but not unitary. Let $B$,

$$(Bg)(\xi)=\int_{-\infty}^{\infty} b(\xi-\eta)g(\eta)\,d\eta, \tag{1-9}$$

be the operator of convolution with the function (1-5), and let $S$ be the operator of multiplication by $s(x)$ so that $S=\Phi^*B\Phi$. If $h(t)=\overline{h(t)}$, then $b(-\xi)=\overline{b(\xi)}$ and $s(x)=\overline{s(x)}$ so that the operators $B$ and $S$ are formally symmetric.

We establish the identity

$$(Hf,f)=(Bg,g)=(Su,u), \tag{1-10}$$

where

$$g(\xi)=\Gamma\big(\tfrac12+i\xi\big)\tilde f(\xi)=:(\Xi f)(\xi),\quad u(x)=(\Phi^*g)(x), \tag{1-11}$$

$\tilde{f}(\xi)$ is the Mellin transform of $f(t)$, and $\Gamma(\,\cdot\,)$ is the gamma function. We often write the identity (1-10) in short form as

$$H = \Xi^* B \Xi = \widehat{\Xi}^* S \widehat{\Xi}, \tag{1-12}$$

where $\widehat{\Xi} = \Phi^* \Xi$.

It follows from (1-12) that the total multiplicities of the strictly positive (negative) spectra of the operators $H$ and $B$, or $S$, coincide:

$$N_\pm(H) = N_\pm(B) = N_\pm(S). \tag{1-13}$$

This result can be compared with Sylvester's inertia theorem, which states the same for Hermitian matrices $H$ and $B$, or $S$, related by (1-12) provided the matrix $\Xi$, or $\widehat{\Xi}$, is invertible. In contrast to linear algebra, in our case the operators $H$ and $B$, or $S$, are of a completely different nature and $B$ and $S$ (but not $H$) admit explicit spectral analysis.

Thus our calculation of the numbers $N_\pm(H)$ consists of two parts. The first is the construction of the sign function (distribution) $s(x)$. The second is the study of the operator $S$ of multiplication by $s(x)$. Observe that since $s(x)$ is a distribution, the numbers $N_\pm(S)$ are not necessarily zero or infinity. We also note that $N_\pm(S) = N_\pm(\Sigma)$ because the functions $s(x)$ and $\sigma(\lambda)$ differ by the change of variables (1-6) only.

In particular, we see that the Hankel operator $H$ is positive if and only if $B \geq 0$ or, equivalently, $S \geq 0$. This means that a Hankel operator $H$ is positive if and only if its sign function $s(x)$ is positive. In some cases the calculation of the sign function is not necessary. Actually, we show that if $|b(\xi)| \to \infty$ as $|\xi| \to \infty$, then $H$ is not sign-definite.

Under the assumption $h \in \mathscr{Z}'_+$, we prove the identity (1-10) for test functions $f(t)$ whose Mellin transforms $\tilde{f}$ are in $C_0^\infty(\mathbb{R})$. Then functions (1-11) belong to $C_0^\infty(\mathbb{R})$ and both sides of (1-10) are well defined. The condition $h \in \mathscr{Z}'_+$ is very general. It is satisfied for *all* bounded, but also for a wide class of unbounded, Hankel operators $H$. More than that, it is not even required that $H$ be defined by formula (1-1) on some dense set. Therefore we work with quadratic forms $(Hf, f)$, which is more convenient and yields more general results. This context allows us to accommodate distributions $h(t)$ as kernels of Hankel operators and makes the theory self-consistent. Note that for bounded operators $H$, the identity (1-10) extends to all elements $f \in L^2(\mathbb{R}_+)$.

**1.4.** Representation (1-2) does not require the positivity of $H$. If, however, $H \geq 0$, then combining our results with the Bochner–Schwartz theorem, we obtain that $\sigma(\lambda)\,d\lambda = dm(\lambda)$, where $dm(\lambda)$ is a positive measure on $\mathbb{R}_+$ ($m(\{0\}) = 0$). In this case, representation (1-2) reduces to (1-3), with the measure $dm(\lambda)$ satisfying for some $\varkappa$ the condition

$$\int_0^\infty (1 + |\ln\lambda|)^{-\varkappa} \lambda^{-1}\, dm(\lambda) < \infty, \tag{1-14}$$

which follows from the assumption $h \in \mathscr{Z}'_+$.

Recall that according to the Bernstein theorem (see the original paper [1929] or [Akhiezer 1965; Widder 1941]), the representation (1-3) is true if and only if the function $h(t)$ is completely monotonic. In contrast to this classical result, we link the representation (1-3) to the positivity of the Hankel operator $H$ with kernel $h(t)$. This fact is not very surprising in view of the analogy with the discrete case when Hankel operators are given in the space $\ell^2(\mathbb{Z}_+)$ by infinite matrices with elements $h_{n+m}$ where $n, m \in \mathbb{Z}_+$. Indeed, according to the classical Hamburger theorem (see, e.g., [Akhiezer 1965]), the positivity of a discrete Hankel operator is equivalent to the existence of a solution of the moment problem with moments $h_n$. In the continuous case, the role of the moment problem is played by the exponential representation (1-3).

We mention that Hankel operators $H$ with kernels $h(t)$ admitting representation (1-3) were considered by H. Widom [1966] and J. S. Howland [1971]. Such kernels $h(t)$ and operators $H$ are necessarily positive. Widom proved that $H$ is bounded if and only if $m([0,\lambda)) = O(\lambda)$ as $\lambda \to 0$ and as $\lambda \to \infty$. In this case, $h(t) \le Ct^{-1}$ for some $C > 0$. Howland showed that $H$ belongs to the trace class if and only if condition (1-14) is satisfied for $\varkappa = 0$.

**1.5.** A large part of this paper is devoted to applying the general theory to various classes of Hankel operators $H$, although we do not try to cover all possible cases. In some examples, the sign-definiteness of $H$ can also be verified or refuted with the help of Bernstein's theorems. Note, however, that our approach yields additionally an explicit formula for the total numbers $N_\pm(H)$ of positive and negative eigenvalues of $H$.

In Section 5, we calculate $N_\pm(H)$ for Hankel operators $H$ of finite rank. Then we consider two specific examples. The first one is given by the formula

$$h(t) = t^k e^{-\alpha t}, \quad \alpha > 0,\ k \ge -1. \tag{1-15}$$

Note that the Hankel operator $H$ with such kernel has finite rank for $k \in \mathbb{Z}_+$ only. We show that $H$ is positive if and only if $k \le 0$. The second class of kernels is defined by the formula

$$h(t) = e^{-t^r}, \quad r > 0. \tag{1-16}$$

It turns out that the corresponding Hankel operator is positive if and only if $r \le 1$.

Section 6 is devoted to a study of Hankel operators $H$ with kernels $h(t)$ having a singularity at a single point $t_0 > 0$. In this case the operators $H$ are compact, but both numbers $N_\pm(H)$ are infinite. We find the asymptotics of positive ($\lambda_n^{(+)}$) and negative ($\lambda_n^{(-)}$) eigenvalues of $H$ as $n \to \infty$ for singularities of different strengths.

Finally, in Section 7, we consider perturbations of the Carleman operator $\boldsymbol{C}$, that is, of the Hankel operator with kernel $t^{-1}$, by various classes of compact Hankel operators $V$. The operator $\boldsymbol{C}$ can be explicitly diagonalized by the Mellin transform. We recall that it has the absolutely continuous spectrum $[0, \pi]$ of multiplicity 2. The Carleman operator plays a distinguished role in the theory of Hankel operators. In particular, it is important for us that its sign function $s(x)$ equals 1. As was pointed out by Howland [1992], Hankel operators are to a certain extent similar to differential operators. In this analogy, the Carleman operator $\boldsymbol{C}$ plays the role of the "free" Schrödinger operator $D^2$, $D = -id/dx$, in the space $L^2(\mathbb{R})$. Furthermore, Hankel operators $H = \boldsymbol{C} + V$ with "perturbed" kernels $h(t) = t^{-1} + v(t)$ can be compared to Schrödinger operators $D^2 + \mathsf{V}(x)$. The assumption that $v(t)$ decays sufficiently rapidly as

$t \to \infty$ and is not too singular as $t \to 0$ corresponds to a sufficiently rapid decay of the potential $\mathsf{V}(x)$ as $|x| \to \infty$.

As shown in [Yafaev 2013], the results on the discrete spectrum of the operator $H$ lying *above* its essential spectrum $[0, \pi]$ are close in spirit to the results on the discrete spectrum of the Schrödinger operator $D^2 + \mathsf{V}(x)$. On the other hand, the results on the negative spectrum of the Hankel operator $H$ are drastically different. In particular, contrary to the case of differential operators with decaying coefficients, the finiteness of the negative spectrum of the Hankel operator $H$ is not determined by the behavior of $v(t)$ at singular points $t = 0$ and $t = \infty$. As an example, consider the Hankel operator with kernel

$$h(t) = t^{-1} - \gamma e^{-t^r}, \quad r \in (0, 1).$$

Now the kernel of the perturbation is a function that decays faster than any power of $t^{-1}$ as $t \to \infty$, and it has a finite limit as $t \to 0$. Nevertheless, we show that the negative spectrum of $H$ is infinite if $\gamma > \gamma_0$ (here $\gamma_0 = \gamma_0(r)$ is an explicit constant), while $H$ is positive if $\gamma \le \gamma_0$. Such a phenomenon has no analogy for Schrödinger operators with decaying potentials, although a somewhat similar effect (known as the Efimov effect) occurs for three-particle Schrödinger operators. Note, however, that for $\gamma > \gamma_0$, a new band of the continuous spectrum appears for three-particle systems, while in our case, the continuous spectrum of $H$ is $[0, \pi]$ for all values of $\gamma$.

We also study perturbations of the Carleman operator $\boldsymbol{C}$ by Hankel operators $V$ of finite rank. Here we obtain a striking result: the total numbers of negative eigenvalues of the operators $H = \boldsymbol{C} + V$ and $V$ coincide.

As examples, we consider only bounded Hankel operators in this paper. However, our general results directly apply to a wide class of unbounded operators, such as Hankel operators with kernels (1-8); see [Yafaev 2014a].

**1.6.** Let us briefly describe the structure of the paper. In Section 2, we define the basic objects, establish the inversion formula (1-2), and obtain the main identity (1-10). Necessary information on bounded Hankel operators (including a continuous version of the Nehari theorem) is collected in Section 3. In Sections 2 and 3, we do not assume that the function $h$ is real, i.e., the corresponding Hankel operator $H$ is not necessarily symmetric. Spectral consequences of the formula (1-10) and, in particular, criteria for the sign-definiteness of Hankel operators are stated in Section 4. In Sections 5, 6, and 7, we apply the general theory to particular classes of Hankel operators.

Let us introduce some standard notation. We denote by $\Phi$,

$$(\Phi u)(\xi) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} u(x) e^{-ix\xi}\, dx,$$

the Fourier transform. The space $\mathcal{Z} = \mathcal{Z}(\mathbb{R})$ of test functions is defined as the subset of the Schwartz space $\mathcal{S} = \mathcal{S}(\mathbb{R})$ which consists of functions $\varphi$ admitting the analytic continuation to entire functions in the complex plane $\mathbb{C}$ and satisfying bounds

$$|\varphi(z)| \le C_n (1 + |z|)^{-n} e^{r|\operatorname{Im} z|}, \quad \text{for all } z \in \mathbb{C},$$

for some $r = r(\varphi) > 0$ and all $n$. Note that $\mathcal{Z}$ is invariant with respect to the complex conjugation $\varphi(z) \mapsto \varphi^*(z) = \overline{\varphi(\bar{z})}$. We recall that the Fourier transform $\Phi$ maps $\mathcal{Z} \to C_0^\infty(\mathbb{R})$ and that $\Phi^* : C_0^\infty(\mathbb{R}) \to \mathcal{Z}$. The dual classes of distributions (continuous antilinear functionals) are denoted by $\mathcal{S}'$, $C_0^\infty(\mathbb{R})'$, and $\mathcal{Z}'$, respectively. In general, for a linear topological space $\mathcal{L}$, we use the notation $\mathcal{L}'$ for its dual space.

We use the notation $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ and $\langle \cdot, \cdot \rangle$ for the duality symbols in $L^2(\mathbb{R}_+)$ and $L^2(\mathbb{R})$, respectively. They are always linear in the first argument and antilinear in the second argument. The letter $C$ (sometimes with indices) denotes various positive constants whose precise values are inessential.

## 2. The main identity

**2.1.** Let us consider the Hankel operator $H$ defined by equality (1-1) in the space $L^2(\mathbb{R}_+)$. Actually, it is more convenient to work with sesquilinear forms $(Hf_1, f_2)$ instead of operators.

Before giving precise definitions, let us explain our construction at a formal level. It follows from (1-1) that

$$(Hf_1, f_2) = \int_0^\infty \int_0^\infty h(t+s) f_1(s) \overline{f_2(t)}\, dt\, ds = \int_0^\infty h(t) \overline{F(t)}\, dt =: \langle\!\langle h, F \rangle\!\rangle, \tag{2-1}$$

where

$$F(t) = \int_0^t \overline{f_1(s)} f_2(t-s)\, ds =: (\bar{f}_1 \star f_2)(t) \tag{2-2}$$

is the Laplace convolution of the functions $\bar{f}_1$ and $f_2$. Formula (2-1) allows us to consider $h$ as a distribution with the test function $F$ defined by (2-2). Thus the Hankel quadratic form will be defined by the relation

$$h[f_1, f_2] := \langle\!\langle h, \bar{f}_1 \star f_2 \rangle\!\rangle. \tag{2-3}$$

Let us introduce the test function

$$\Omega(x) = F(e^x) =: (\mathcal{R}F)(x) \tag{2-4}$$

and the distribution

$$\theta(x) = e^x h(e^x) \tag{2-5}$$

defined for $x \in \mathbb{R}$. Setting $t = e^x$ in (2-1), we see that

$$\langle\!\langle h, F \rangle\!\rangle = \int_{-\infty}^\infty \theta(x) \overline{\Omega(x)}\, dx =: \langle \theta, \Omega \rangle. \tag{2-6}$$

We are going to consider the form (2-6) on pairs $F, h$ such that the corresponding test function $\Omega$ defined by (2-4) is an element of the space $\mathcal{Z}$ of analytic functions and the corresponding distribution $\theta$ defined by (2-5) is an element of the dual space $\mathcal{Z}'$. The set of all such $F$ and $h$ will be denoted by $\mathcal{Z}_+$ and $\mathcal{Z}'_+$, respectively; that is,

$$F \in \mathcal{Z}_+ \iff \Omega \in \mathcal{Z} \quad \text{and} \quad h \in \mathcal{Z}'_+ \iff \theta \in \mathcal{Z}'. \tag{2-7}$$

Of course, the topology in $\mathcal{Z}_+$ is induced by that in $\mathcal{Z}$ and $\mathcal{Z}'_+$ is dual to $\mathcal{Z}_+$. Note that $h \in \mathcal{Z}'_+$ if $h \in L^1_{\mathrm{loc}}(\mathbb{R}_+)$ and integral (1-7) is convergent for some $\kappa$. In this case, the corresponding function (2-5) satisfies the condition

$$\int_{-\infty}^{\infty} |\theta(x)|(1+|x|)^{-\kappa}\, dx < \infty,$$

and hence $\theta \in \mathcal{S}' \subset \mathcal{Z}'$.

Define the unitary operator $U : L^2(\mathbb{R}_+) \to L^2(\mathbb{R})$ by the equality

$$(Uf)(x) = e^{x/2} f(e^x). \tag{2-8}$$

Let the set $\mathcal{D}$ consist of functions $f(t)$ such that $Uf \in \mathcal{Z}$. Since

$$f(t) = t^{-1/2}(Uf)(\ln t)$$

and $\mathcal{Z} \subset \mathcal{S}$, we see that functions $f \in \mathcal{D}$ and their derivatives satisfy the estimates

$$|f^{(m)}(t)| = C_{n,m} t^{-1/2-m}(1+|\ln t|)^{-n}$$

for all $n$ and $m$. Obviously, $f \in \mathcal{D}$ if and only if $\varphi(t) = t^{1/2} f(t)$ belongs to the class $\mathcal{Z}_+$.

Let us show that form (2-3) is correctly defined on functions $f_1, f_2 \in \mathcal{D}$. To that end, we have to verify that function (2-2) belongs to the space $\mathcal{Z}_+$ or, equivalently, that function (2-4) belongs to the space $\mathcal{Z}$. This requires some preliminary study, which will also allow us to derive a convenient representation for form (2-3).

Recall that the Mellin transform $M : L^2(\mathbb{R}_+) \to L^2(\mathbb{R})$ is defined by the formula

$$(Mf)(\xi) = (2\pi)^{-1/2} \int_0^{\infty} f(t) t^{-1/2-i\xi}\, dt. \tag{2-9}$$

Of course, $M = \Phi U$, where $\Phi$ is the Fourier transform and $U$ is operator (2-8). Since both $\Phi$ and $U$ are unitary, the operator $M$ is also unitary. The inversion of the formula (2-9) is given by the relation

$$f(t) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} \tilde{f}(\xi) t^{-1/2+i\xi}\, d\xi, \quad \tilde{f} = Mf. \tag{2-10}$$

Let $\Gamma(z)$ be the gamma function. Recall that $\Gamma(z)$ is a holomorphic function in the right half-plane and $\Gamma(z) \neq 0$ for all $z \in \mathbb{C}$. According to the Stirling formula, the function $\Gamma(z)$ tends to zero exponentially as $|z| \to \infty$ parallel with the imaginary axis. To be more precise, we have

$$\Gamma(\alpha + i\lambda) = e^{\pi i(2\alpha-1)/4} \left(\frac{2\pi}{e}\right)^{1/2} \lambda^{\alpha-1/2} e^{i\lambda(\ln\lambda - 1)} e^{-\pi\lambda/2}(1 + O(\lambda^{-1})) \tag{2-11}$$

for a fixed $\alpha > 0$ and $\lambda \to +\infty$. Since $\Gamma(\alpha - i\lambda) = \overline{\Gamma(\alpha + i\lambda)}$, this yields also the asymptotics of $\Gamma(\alpha + i\lambda)$ as $\lambda \to -\infty$.

If $f_j \in \mathcal{D}$, $j = 1, 2$, then $\tilde{f}_j = Mf_j = \Phi U f_j \in C_0^\infty(\mathbb{R})$, and hence the functions $g_j(\xi)$ defined by formula (1-11) also belong to the class $C_0^\infty(\mathbb{R})$. Let us introduce the convolution of the functions $g_1$ and $g_2$,

$$(g_1 * g_2)(\xi) = \int_{-\infty}^{\infty} g_1(\xi - \eta) g_2(\eta)\, d\eta,$$

and set

$$(\mathcal{J} g)(\xi) = g(-\xi).$$

We have the following result.

**Lemma 2.1.** *Suppose that* $f_j \in \mathcal{D}$, $j = 1, 2$, *and define functions* $g_j(\xi)$ *by equality* (1-11). *Let the function* $\Omega(x)$ *be defined by formulas* (2-2) *and* (2-4). *Then*

$$(\Phi\Omega)(\xi) = (2\pi)^{-1/2}\Gamma(1+i\xi)^{-1}((\mathcal{J}\bar{g}_1) * g_2)(\xi). \tag{2-12}$$

*Proof.* Substituting (2-10) into (2-2), we see that

$$F(t) = (2\pi)^{-1}\int_0^t ds \int_{-\infty}^{\infty} \overline{\tilde{f}_1(\tau)}(t-s)^{-1/2-i\tau}\, d\tau \int_{-\infty}^{\infty} \tilde{f}_2(\sigma)s^{-1/2+i\sigma}\, d\sigma.$$

Observe that

$$\int_0^t (t-s)^{-1/2-i\tau}s^{-1/2+i\sigma}\, ds = t^{i(\sigma-\tau)}\frac{\Gamma\left(\frac{1}{2}-i\tau\right)\Gamma\left(\frac{1}{2}+i\sigma\right)}{\Gamma(1+i(\sigma-\tau))}.$$

Then using (1-11), we obtain the representation

$$\begin{aligned} F(t) &= (2\pi)^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} t^{i(\sigma-\tau)}\Gamma(1+i(\sigma-\tau))^{-1}\overline{g_1(\tau)}g_2(\sigma)\, d\tau\, d\sigma \\ &= (2\pi)^{-1}\int_{-\infty}^{\infty} t^{i\xi}\Gamma(1+i\xi)^{-1}((\mathcal{J}\bar{g}_1) * g_2)(\xi)\, d\xi, \end{aligned}$$

whence

$$\Omega(x) = (2\pi)^{-1}\int_{-\infty}^{\infty} e^{i\xi x}\Gamma(1+i\xi)^{-1}((\mathcal{J}\bar{g}_1) * g_2)(\xi)\, d\xi.$$

This is equivalent to formula (2-12). □

Observe that the function $\Gamma(1+i\xi)^{-1}$ on the right-hand side of (2-12) tends to infinity exponentially as $|\xi| \to \infty$. Nevertheless, $\Phi\Omega \in C_0^\infty(\mathbb{R})$ because $(\mathcal{J}\bar{g}_1) * g_2 \in C_0^\infty(\mathbb{R})$ for $g_1, g_2 \in C_0^\infty(\mathbb{R})$. Thus we have:

**Corollary 2.2.** *Let* $f_j \in \mathcal{D}$, $j = 1, 2$, *and let the function* $\Omega(x)$ *be defined by formulas* (2-2) *and* (2-4). *Then* $\Omega \in \mathcal{Z}$ *or, equivalently,* $F \in \mathcal{Z}_+$.

Now we are in a position to give the precise definition.

**Definition 2.3.** Let $h \in \mathcal{Z}'_+$ and $f_j \in \mathcal{D}$, $j = 1, 2$. Then $\bar{f}_1 \star f_2 \in \mathcal{Z}_+$ and the Hankel sesquilinear form is defined by the relation (2-3).

We shall see in Section 2.4 that $h \in \mathcal{Z}'_+$ is determined uniquely by the values $\langle h, \bar{f}_1 \star f_2\rangle$ on $f_1, f_2 \in \mathcal{D}$, that is, $h = 0$ if $\langle h, \bar{f}_1 \star f_2\rangle = 0$ for all $f_1, f_2 \in \mathcal{D}$.

Of course (2-3) can be rewritten as

$$h[f_1, f_2] = \langle \theta, \Omega\rangle, \tag{2-13}$$

where $\theta$ is distribution (2-5) and

$$\Omega(x) = (\bar{f}_1 \star f_2)(e^x).$$

We sometimes write $h[f_1, f_2]$ as integral (2-1), keeping in mind that its precise meaning is given by Definition 2.3.

**2.2.** Our next goal is to show that (2-13) is the sesquilinear form of the convolution operator $B$, that is, it equals the right-hand side of (1-10). Here the representation of Lemma 2.1 for the function

$$G(\xi) := \sqrt{2\pi}\,\Gamma(1+i\xi)(\Phi\Omega)(\xi) \tag{2-14}$$

plays a crucial role.

Since $\theta$ is in $\mathfrak{X}'$, its Fourier transform $a = \Phi\theta$ is correctly defined as an element of $C_0^\infty(\mathbb{R})'$. Formally,

$$a(\xi) = (\Phi\theta)(\xi) = (2\pi)^{-1/2} \int_0^\infty h(t) t^{-i\xi}\, dt, \tag{2-15}$$

that is, $a(\xi)$ is the Mellin transform of the function $h(t)t^{1/2}$. Let $\Omega \in \mathfrak{X}$. Passing to the Fourier transforms and using notation (2-14), we see that

$$\langle \theta, \Omega \rangle = \langle a, \Phi\Omega \rangle = \langle b, G \rangle, \tag{2-16}$$

where $G \in C_0^\infty(\mathbb{R})$ and the distribution $b \in C_0^\infty(\mathbb{R})'$ is given by the relation

$$b(\xi) = (2\pi)^{-1/2} a(\xi) \Gamma(1 - i\xi)^{-1}, \tag{2-17}$$

which is of course the same as (1-5). Thus we are led to the following.

**Definition 2.4.** Let $h \in \mathfrak{X}'_+$. The distribution $b \in C_0^\infty(\mathbb{R})'$ defined by formulas (2-5), (2-15), and (2-17) is called the $b$-function of the kernel $h(t)$ (or of the Hankel operator $H$). Its Fourier transform $s = \sqrt{2\pi}\,\Phi^* b \in \mathfrak{X}'$ is called the $s$-function or the sign function.

Recall that the distribution $\sigma$ was defined by relation (1-6). It is convenient to also introduce

$$h^\natural(\lambda) = \lambda^{-1}\sigma(\lambda) = \lambda^{-1} s(-\ln \lambda). \tag{2-18}$$

The following assertion is an immediate consequence of formulas (2-5), (2-15), and (2-17).

**Proposition 2.5.** *The mappings*

$$h \mapsto \theta \mapsto a \mapsto b \mapsto s \mapsto h^\natural$$

*yield one-to-one correspondences* (*bijections*)

$$\mathfrak{X}'_+ \to \mathfrak{X}' \to C_0^\infty(\mathbb{R}) \to C_0^\infty(\mathbb{R}) \to \mathfrak{X}' \to \mathfrak{X}'_+.$$

*All of them*, *as well as their inverse mappings*, *are continuous.*

Putting together equalities (2-6) and (2-16), we see that

$$\langle\!\langle h, F \rangle\!\rangle = \langle b, G \rangle. \tag{2-19}$$

Combining this relation with Lemma 2.1 and Definitions 2.3, 2.4 and using notation (1-11), we obtain the main identity (1-10). To be more precise, we have the following result.

**Theorem 2.6.** *Suppose that $h \in \mathcal{Z}'_+$, and let $b \in C_0^\infty(\mathbb{R})'$ be the corresponding b-function. Let $f_j \in \mathcal{D}$, $j = 1, 2$, and let the functions $g_j$ be defined by formula* (1-11). *Then $g_j \in C_0^\infty(\mathbb{R})$, and the representation*

$$\langle h, \bar{f}_1 \star f_2 \rangle = \langle b, (\mathcal{J}\bar{g}_1) * g_2 \rangle =: b[g_1, g_2] \tag{2-20}$$

*holds.*

Passing to the Fourier transforms on the right-hand side of (2-20) and using

$$\Phi^*((\mathcal{J}\bar{g}_1) * g_2) = (2\pi)^{1/2}\overline{\Phi^* g_1}\Phi^* g_2,$$

we obtain:

**Corollary 2.7.** *Let $s \in \mathcal{Z}'$ be the sign function of $h$, and let $u_j = \Phi^* g_j = \Phi^* \Xi f_j \in \mathcal{Z}$. Then*

$$\langle h, \bar{f}_1 \star f_2 \rangle = \langle s, u_1^* u_2 \rangle =: s[u_1, u_2]. \tag{2-21}$$

Loosely speaking, equalities (2-20) and (2-21) mean that

$$\langle h, \bar{f}_1 \star f_2 \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\xi - \eta) g_1(\eta) \overline{g_2(\xi)} \, d\xi \, d\eta = \int_{-\infty}^{\infty} s(x) u_1(x) \overline{u_2(x)} \, dx. \tag{2-22}$$

In the particular case $h(t) = t^{-1}$ when $H = \boldsymbol{C}$ is the Carleman operator, we have

$$\theta(x) = 1, \quad a(\xi) = (2\pi)^{1/2}\delta(\xi), \quad b(\xi) = \delta(\xi), \quad s(x) = 1, \tag{2-23}$$

and hence (2-22) yields

$$\langle h, \bar{f}_1 \star f_2 \rangle = \int_{-\infty}^{\infty} g_1(\xi) \overline{g_2(\xi)} \, d\xi = \int_{-\infty}^{\infty} \left|\Gamma\left(\tfrac{1}{2} + i\xi\right)\right|^2 \tilde{f}_1(\xi) \overline{\tilde{f}_2(\xi)} \, d\xi,$$

where

$$\left|\Gamma\left(\tfrac{1}{2} + i\xi\right)\right|^2 = \frac{\pi}{\cosh(\pi\xi)}. \tag{2-24}$$

This leads to the familiar diagonalization of the Carleman operator.

**2.3.** According to Proposition 2.5, the distribution $h^\natural$ determines uniquely the distribution $h$. Let us now obtain an explicit formula for the mapping $h^\natural \mapsto h$. This requires some auxiliary information.

Let $\boldsymbol{\Gamma}_\alpha : C_0^\infty(\mathbb{R}) \to C_0^\infty(\mathbb{R})$, $\alpha > 0$, be the operator of multiplication by the function $\Gamma(\alpha + i\xi)$. Making the change of variables $t = e^{-x}$ in the definition of the gamma function, we see that

$$\Gamma(\alpha + i\lambda) = \int_0^\infty e^{-t} t^{\alpha + i\lambda - 1} \, dt = \int_{-\infty}^{\infty} e^{-e^{-x}} e^{-\alpha x} e^{-ix\lambda} \, dx, \quad \alpha > 0,$$

and hence

$$(2\pi)^{-1} \int_{-\infty}^{\infty} e^{ix\lambda} \Gamma(\alpha + i\lambda) \, d\lambda = e^{-e^{-x}} e^{-\alpha x}. \tag{2-25}$$

It follows that

$$(\Phi^* \boldsymbol{\Gamma}_\alpha \Phi \Omega)(x) = \int_{-\infty}^{\infty} e^{\alpha(y-x)} e^{-e^{y-x}} \Omega(y) \, dy. \tag{2-26}$$

Let us also introduce the operator $\mathsf{L}_\alpha$:

$$(\mathsf{L}_\alpha F)(\lambda) = \lambda^\alpha \int_0^\infty e^{-t\lambda} t^{\alpha-1} F(t)\, dt, \quad \lambda > 0, \quad \alpha > 0. \tag{2-27}$$

Obviously, $\mathsf{L}_\alpha F \in C^\infty(\mathbb{R}_+)$ for all bounded functions $F(t)$ and, in particular, for $F \in \mathcal{Z}_+$. Note that $\mathsf{L}_\alpha$ is the Laplace operator $\mathsf{L}$,

$$(\mathsf{L} F)(\lambda) = \int_0^\infty e^{-t\lambda} F(t)\, dt, \tag{2-28}$$

sandwiched by the weights $\lambda^\alpha$ and $t^{\alpha-1}$.

The following result yields the whole scale of spaces where the Laplace operator $\mathsf{L}$ acts as an isomorphism. Recall that the operator $\mathcal{R}$ defined by (2-4) is a one-to-one mapping of $\mathcal{Z}_+$ onto $\mathcal{Z}$.

**Lemma 2.8.** *For all $\alpha > 0$, the identity*

$$\mathsf{L}_\alpha = \mathcal{R}^{-1} \mathcal{J} \Phi^* \mathbf{\Gamma}_\alpha \Phi \mathcal{R} \tag{2-29}$$

*holds. In particular, $\mathsf{L}_\alpha$ and its inverse are the one-to-one continuous mappings of $\mathcal{Z}_+$ onto itself.*

*Proof.* Putting $\Omega(y) = (\mathcal{R}F)(y) = F(e^y)$ in (2-26) and making the change of variables $t = e^y$, we find that

$$(\Phi^* \mathbf{\Gamma}_\alpha \Phi \mathcal{R} F)(x) = e^{-\alpha x} \int_0^\infty e^{-e^{-x} t} t^{\alpha-1} F(t)\, dt.$$

Now making the change of variables $\lambda = e^{-x}$, we arrive at the identity (2-29).

Consider the right-hand side of (2-29). All mappings $\mathcal{R} : \mathcal{Z}_+ \to \mathcal{Z}$, $\Phi : \mathcal{Z} \to C_0^\infty(\mathbb{R})$, $\mathbf{\Gamma}_\alpha : C_0^\infty(\mathbb{R}) \to C_0^\infty(\mathbb{R})$, $\Phi^* : C_0^\infty(\mathbb{R}) \to \mathcal{Z}$, $\mathcal{J} : \mathcal{Z} \to \mathcal{Z}$ are bijections. All of them as well as their inverses are continuous. Therefore the identity (2-29) ensures the same result for the operator $\mathsf{L}_\alpha : \mathcal{Z}_+ \to \mathcal{Z}_+$. $\square$

The adjoint operators $\mathsf{L}_\alpha^*$ are defined by the relation $\langle \mathsf{L}_\alpha F, \psi \rangle = \langle F, \mathsf{L}_\alpha^* \psi \rangle$, where $F \in \mathcal{Z}_+$ and $\psi \in \mathcal{Z}_+'$ are arbitrary. According to (2-27), they are formally given by the relation

$$(\mathsf{L}_\alpha^* \psi)(t) = t^{\alpha-1} \int_0^\infty e^{-t\lambda} \lambda^\alpha \psi(\lambda)\, d\lambda, \quad t > 0. \tag{2-30}$$

By duality, the next assertion follows from Lemma 2.8.

**Theorem 2.9.** *For all $\alpha > 0$, the operators $\mathsf{L}_\alpha^*$ as well as their inverses are the one-to-one continuous mappings of $\mathcal{Z}_+'$ onto itself.*

To recover $h(t)$, we proceed from formula (2-19). Passing to the Fourier transforms, we can write it as

$$\langle h, F \rangle = (2\pi)^{-1/2} \langle s, \Phi^* G \rangle,$$

where $G$ is defined by formulas (2-4), (2-14), that is, $G = (2\pi)^{1/2} \mathbf{\Gamma}_1 \Phi \mathcal{R} F$. Therefore, using the identity (2-29) for $\alpha = 1$, we see that

$$\langle h, F \rangle = \langle s, \mathcal{J} \mathcal{R} \mathsf{L}_1 F \rangle = \int_{-\infty}^\infty s(x) \overline{(\mathsf{L}_1 F)(e^{-x})}\, dx.$$

Making the change of variables $\lambda = e^{-x}$ and using notation (2-18), we obtain the identity

$$\langle h, F\rangle = \langle h^\natural, \mathsf{L}_1 F\rangle.$$

Passing here to adjoint operators and taking into account that $F \in \mathcal{X}_+$ is arbitrary, we find that

$$h = \mathsf{L}_1^* h^\natural \quad \text{or} \quad h = \mathsf{L}^*\sigma, \tag{2-31}$$

where $\sigma(\lambda) = \lambda h^\natural(\lambda)$. In view of (2-30), this gives the precise sense to formula (1-2).

Let us state the result obtained.

**Theorem 2.10.** *Let $h \in \mathcal{X}'_+$, and let $s \in \mathcal{X}'$ be the corresponding sign function* (*see Definition 2.4*). *Define the distribution $h^\natural$ by formula* (2-18). *Then $h^\natural \in \mathcal{X}'_+$ and $h$ can be recovered from $h^\natural$ or $\sigma$ by formulas* (2-31).

We emphasize that in the roundabout $h \mapsto h^\natural \mapsto h$, the mapping $h \mapsto h^\natural$ and its inverse $h^\natural \mapsto h$ are the one-to-one continuous mappings of the set $\mathcal{X}'_+$ onto itself.

Let us also give a direct expression of $u(x) = (\Phi^* g)(x)$ in terms of $f(t)$.

**Lemma 2.11.** *Suppose that $f \in \mathcal{D}$ and put $\varphi(t) = t^{1/2} f(t)$. Let $g(\xi)$ be defined by formula* (1-11) *and $u(x) = (\Phi^* g)(x)$. Then*

$$u(x) = (\mathsf{L}_{1/2}\varphi)(e^{-x}). \tag{2-32}$$

*Proof.* Since $(\mathcal{R}\varphi)(x) = (Uf)(x)$, it follows from formula (2-29) for $\alpha = \frac{1}{2}$ that

$$(\mathcal{R}^{-1}\mathcal{J}\Phi^*\mathbf{\Gamma}_{1/2}\Phi U f)(\lambda) = (\mathsf{L}_{1/2}\varphi)(\lambda).$$

The left-hand side here equals $(\mathcal{R}^{-1}\mathcal{J}u)(\lambda)$, which after the change of variables $\lambda = e^{-x}$ yields (2-32). □

Now we can rewrite identity (2-21) in a slightly different way.

**Corollary 2.12.** *Let $h \in \mathcal{X}'_+$, and let the distribution $h^\natural \in \mathcal{X}'_+$ be defined by formula* (2-18). *Then for arbitrary $f_j \in \mathcal{D}$, $j = 1, 2$, we have*

$$\langle h, \bar{f}_1 \star f_2\rangle = \langle h^\natural, \overline{\mathsf{L}_{1/2}\varphi_1}\mathsf{L}_{1/2}\varphi_2\rangle, \quad \textit{where } \varphi_j(t) = t^{1/2} f_j(t). \tag{2-33}$$

*Proof.* It suffices to make the change of variables $x = -\ln\lambda$ in the right-hand side of (2-22) and to take equality (2-32) into account. □

We emphasize that according to Lemma 2.8, $\mathsf{L}_{1/2}\varphi_j \in \mathcal{X}_+$, and hence $\overline{\mathsf{L}_{1/2}\varphi_1}\mathsf{L}_{1/2}\varphi_2 \in \mathcal{X}_+$. Thus the right-hand side of (2-33) is correctly defined.

**2.4.** Finally, we check that a distribution $h \in \mathcal{X}'_+$ is determined uniquely by the values $\langle h, \bar{f}_1 \star f_2\rangle$ on $f_1, f_2 \in \mathcal{D}$. First we consider convolution operators. Let us introduce the shift in the space $L^2(\mathbb{R})$:

$$(T(\tau)g)(\xi) = g(\xi - \tau), \quad \tau \in \mathbb{R}. \tag{2-34}$$

Since

$$(g_1 * g_2)(\xi) = \int_{-\infty}^{\infty} (T(\tau)g_1)(\xi) g_2(\tau)\, d\tau \quad \text{for all } g_1, g_2 \in C_0^\infty(\mathbb{R}),$$

we have the formula

$$\langle b, (\mathcal{J}\bar{g}_1) * g_2\rangle = \int_{-\infty}^{\infty} \langle b, T(\tau)\mathcal{J}\bar{g}_1\rangle \overline{g_2(\tau)}\, d\tau, \tag{2-35}$$

where for $b \in C_0^\infty(\mathbb{R})'$ the function $\langle b, T(\tau)\mathcal{J}\bar{g}_1\rangle$ is infinitely differentiable in $\tau \in \mathbb{R}$.

The following assertion is quite standard.

**Lemma 2.13.** *Let $b \in C_0^\infty(\mathbb{R})'$. Suppose that $\langle b, (\mathcal{J}\bar{g}_1) * g_2\rangle = 0$ for all $g_1, g_2 \in C_0^\infty(\mathbb{R})$. Then $b = 0$.*

*Proof.* If $\langle b, (\mathcal{J}\bar{g}_1) * g_2\rangle = 0$ for all $g_2 \in C_0^\infty(\mathbb{R})$, then $\langle b, T(\tau)\mathcal{J}\bar{g}_1\rangle = 0$ for all $\tau \in \mathbb{R}$ according to formula (2-35). In particular, for $\tau = 0$ we have $\langle b, \mathcal{J}\bar{g}_1\rangle = 0$, whence $b = 0$ because $g_1 \in C_0^\infty(\mathbb{R})$ is arbitrary. □

Next we pass to Hankel operators.

**Proposition 2.14.** *Let $h \in \mathcal{Z}'_+$. Suppose that $\langle\!\langle h, \bar{f}_1 \star f_2\rangle\!\rangle = 0$ for all $f_1, f_2 \in \mathcal{D}$. Then $h = 0$.*

*Proof.* Let $b \in C_0^\infty(\mathbb{R})'$ be the $b$-function of $h$ (see Definition 2.4). For arbitrary $g_1, g_2 \in C_0^\infty(\mathbb{R})$, we can construct $f_1, f_2 \in \mathcal{D}$ by formula (1-11). Since $\langle\!\langle h, \bar{f}_1 \star f_2\rangle\!\rangle = 0$, it follows from the identity (2-20) that $\langle b, (\mathcal{J}\bar{g}_1) * g_2\rangle = 0$. Therefore $b = 0$ according to Lemma 2.13. Now Proposition 2.5 implies that $h = 0$. □

## 3. Bounded Hankel operators

Our main goal here is to show that the condition $h \in \mathcal{Z}'_+$ is satisfied for all bounded Hankel operators $H$.

**3.1.** In this section we a priori only assume that $h \in C_0^\infty(\mathbb{R}_+)'$ and consider the Hankel form (2-3) on functions $f_1, f_2 \in C_0^\infty(\mathbb{R}_+)$. Let $T_+(\tau)$, where $\tau \geq 0$, be the restriction of the shift (2-34) on its invariant subspace $L^2(\mathbb{R}_+)$. Since

$$(\bar{f}_1 \star f_2)(t) = \int_0^\infty (T_+(\tau)\bar{f}_1)(t) f_2(\tau)\, d\tau \quad \text{for all } f_1, f_2 \in C_0^\infty(\mathbb{R}_+),$$

for all $h \in C_0^\infty(\mathbb{R}_+)'$ we have the formula

$$\langle\!\langle h, \bar{f}_1 \star f_2\rangle\!\rangle = \int_0^\infty \langle\!\langle h, T_+(\tau)\bar{f}_1\rangle\!\rangle \overline{f_2(\tau)}\, d\tau. \tag{3-1}$$

Here the function $\langle\!\langle h, T_+(\tau)\bar{f}_1\rangle\!\rangle$ is infinitely differentiable in $\tau \in \mathbb{R}_+$, and this function, as well as all its derivatives, has finite limits as $\tau \to 0$. In the theory of Hankel operators, formula (3-1) plays the role of formula (2-35) for convolution operators.

The proof of the following assertion is almost the same as that of Lemma 2.13.

**Proposition 3.1.** *Let $h \in C_0^\infty(\mathbb{R}_+)'$. Suppose that $\langle\!\langle h, \bar{f}_1 \star f_2\rangle\!\rangle = 0$ for all $f_1, f_2 \in C_0^\infty(\mathbb{R}_+)$. Then $h = 0$.*

*Proof.* If $\langle\!\langle h, \bar{f}_1 \star f_2\rangle\!\rangle = 0$ for all $f_2 \in C_0^\infty(\mathbb{R}_+)$, then $\langle\!\langle h, T_+(\tau)\bar{f}_1\rangle\!\rangle = 0$ for all $\tau \in [0, \infty)$ according to formula (3-1). In particular, for $\tau = 0$ we have $\langle\!\langle h, \bar{f}_1\rangle\!\rangle = 0$, which implies that $h = 0$ because $f_1 \in C_0^\infty(\mathbb{R}_+)$ is arbitrary. □

Of course Propositions 2.14 and 3.1 differ only by the set of functions on which the Hankel form is considered.

Assume now that

$$|\langle h, \bar{f} \star f \rangle| \le C \|f\|^2 \quad \text{for all } f \in C_0^\infty(\mathbb{R}_+). \tag{3-2}$$

Then there exists a bounded operator $H$ such that

$$(Hf_1, f_2) = \langle h, \bar{f}_1 \star f_2 \rangle \quad \text{for all } f_1, f_2 \in C_0^\infty(\mathbb{R}_+). \tag{3-3}$$

*We call $H$ the Hankel operator associated to the Hankel form $\langle h, \bar{f}_1 \star f_2 \rangle$.*

**3.2.** It is possible to characterize Hankel operators by some commutation relations. A presentation of such results for discrete Hankel operators acting in the space of sequences $l^2(\mathbb{Z}_+)$ can be found in [Power 1982, §1.1].

Let us define a bounded operator $Q$ in the space $L^2(\mathbb{R}_+)$ by the equality

$$(Qf)(t) = -2e^{-t} \int_0^t e^s f(s)\, ds.$$

Note that

$$Q = -2 \int_0^\infty T_+(\tau) e^{-\tau}\, d\tau. \tag{3-4}$$

**Lemma 3.2.** *Let* (3-2) *hold. Then the operator $H$ defined by formula* (3-3) *satisfies the commutation relations*

$$HT_+(\tau) = T_+(\tau)^* H \quad \textit{for all } \tau \ge 0 \tag{3-5}$$

*and*

$$HQ = Q^* H. \tag{3-6}$$

*Proof.* Since

$$(T_+(\tau)\bar{f}_1) \star f_2 = \bar{f}_1 \star (T_+(\tau) f_2) \quad \text{for all } \tau \ge 0,$$

relation (3-5) directly follows from (3-3). By virtue of formula (3-4), relation (3-6) is a consequence of (3-5). □

Below we need the Nehari theorem; see the original paper [1957] or [Peller 2003, Chapter 1, §1; Power 1982, Chapter 1, §2]. We formulate it in the Hardy space $\mathbb{H}^2_+(\mathbb{R})$ of functions analytic in the upper half-plane. We denote by $\widehat{Q}$ the operator of multiplication by the function $(\mu - i)/(\mu + i)$ in this space. Clearly, $\widehat{Q} = \Phi^* Q \Phi$.

**Theorem 3.3** [Nehari 1957]. *Let $\omega \in L^\infty(\mathbb{R})$, and let an operator $\widehat{H}$ in the space $\mathbb{H}^2_+(\mathbb{R})$ be defined by the relation*

$$(\widehat{H}\hat{f}_1, \hat{f}_2) = \int_{-\infty}^\infty \omega(\mu) \hat{f}_1(-\mu) \overline{\hat{f}_2(\mu)}\, d\mu \quad \textit{for all } \hat{f}_1, \hat{f}_2 \in \mathbb{H}^2_+(\mathbb{R}). \tag{3-7}$$

*Then $\widehat{H}$ is bounded and $\widehat{H}\widehat{Q} = \widehat{Q}^* \widehat{H}$. Conversely, if $\widehat{H}$ is a bounded operator in $\mathbb{H}^2_+(\mathbb{R})$ and $\widehat{H}\widehat{Q} = \widehat{Q}^* \widehat{H}$, then there exists a function $\omega \in L^\infty(\mathbb{R})$ such that representation* (3-7) *holds.*

The following assertion can be regarded as a translation of this theorem into the space $L^2(\mathbb{R}_+)$. Recall that, by the Paley–Wiener theorem, the Fourier transform $\Phi : \mathbb{H}^2_+(\mathbb{R}) \to L^2(\mathbb{R}_+)$ is the unitary operator. Since

$$\int_{-\infty}^{\infty} (\mu + i)^{-1} e^{-i\mu t}\, d\mu = -2\pi i e^{-t}$$

for $t > 0$ and this integral is zero for $t < 0$, we have the relation

$$I + Q = \Phi \widehat{Q} \Phi^*. \tag{3-8}$$

**Theorem 3.4.** *If $h = (2\pi)^{-1/2}\Phi\omega$, where $\omega \in L^\infty(\mathbb{R})$ (in this case $h \in \mathcal{S}' \subset C_0^\infty(\mathbb{R}_+)'$), then estimate* (3-2) *is true and the operator $H$ in the space $L^2(\mathbb{R}_+)$ defined by formula* (3-3) *satisfies the commutation relation* (3-5). *Conversely, if a bounded operator $H$ satisfies* (3-5), *then representation* (3-3) *holds with $h = (2\pi)^{-1/2}\Phi\omega$ for some $\omega \in L^\infty(\mathbb{R})$.*

*Proof.* Since

$$(\Phi^*(\bar{f}_1 \star f_2))(\mu) = \sqrt{2\pi}\, \overline{(\mathcal{J}\hat{f}_1)(\mu)}\, \hat{f}_2(\mu) \quad \text{for all } f_1, f_2 \in C_0^\infty(\mathbb{R}_+),$$

where $\hat{f}_1 = \Phi^* f_1$, $\hat{f}_2 = \Phi^* f_2$, and $(\mathcal{J}\hat{f}_1)(\mu) = \hat{f}_1(-\mu)$, we have

$$\langle h, \bar{f}_1 \star f_2 \rangle = \sqrt{2\pi}\, \langle \Phi^* h, \overline{(\mathcal{J}\hat{f}_1)}\, \hat{f}_2 \rangle \quad \text{for all } h \in \mathcal{S}'. \tag{3-9}$$

Therefore, estimate (3-2) is satisfied if $\Phi^* h \in L^\infty(\mathbb{R})$. Relation (3-5) for the corresponding Hankel operator $H$ follows from Lemma 3.2.

Conversely, if a bounded operator $H$ satisfies relation (3-5), then by virtue of (3-4), it also satisfies relation (3-6). Hence it follows from (3-8) that $\widehat{H}\widehat{Q} = \widehat{Q}^*\widehat{H}$, where $\widehat{H} = \Phi^* H \Phi$ is a bounded operator in the space $\mathbb{H}^2_+(\mathbb{R})$. Thus, by Theorem 3.3, there exists a function $\omega \in L^\infty(\mathbb{R})$ such that representation (3-7) holds. This means that

$$(Hf_1, f_2) = \int_{-\infty}^{\infty} \omega(\mu)\, \hat{f}_1(-\mu)\, \overline{\hat{f}_2(\mu)}\, d\mu \quad \text{for all } f_1, f_2 \in L^2(\mathbb{R}_+). \tag{3-10}$$

If $h = (2\pi)^{-1/2}\Phi\omega$, then the right-hand sides in (3-9) and (3-10) coincide. This yields representation (3-3). □

**Corollary 3.5.** *For a bounded operator $H$ in the space $L^2(\mathbb{R}_+)$, commutation relations* (3-5) *and* (3-6) *are equivalent.*

*Proof.* As was already noted, (3-6) follows from (3-5) according to formula (3-4). Conversely, if $H$ satisfies (3-5), then representation (3-3) holds according to Theorem 3.4. Thus it remains to use Lemma 3.2. □

Recall that *a function $\omega \in L^\infty(\mathbb{R})$ such that $\Phi\omega = \sqrt{2\pi}\, h$ is called the symbol of a bounded Hankel operator $H$ with kernel $h(t)$.* In view of formula (3-7), if $\omega \in \mathbb{H}^\infty_-(\mathbb{R})$, that is, $\omega$ admits an analytic continuation to a bounded function in the lower half-plane, then the Hankel operator $\widehat{H}$ equals 0, and hence $H = \Phi\widehat{H}\Phi^* = 0$. Therefore the symbol is defined up to a function in the class $\mathbb{H}^\infty_-(\mathbb{R})$.

**3.3.** Now we are in a position to check that the condition $h \in \mathscr{Z}'_+$ is satisfied for all bounded Hankel operators. By (2-7), it means that distribution (2-5) belongs to the class $\mathscr{Z}'$. We shall verify the stronger inclusion $\theta \in \mathscr{S}'$.

To that end, it suffices to check that, for some $N \in \mathbb{Z}_+$ and some $\kappa \in \mathbb{R}$,

$$|\langle \theta, \Omega \rangle| \le C \sum_{n=0}^{N} \max_{x \in \mathbb{R}} \big( (1+|x|)^{\kappa} |\Omega^{(n)}(x)| \big) \quad \text{for all } \Omega \in C_0^\infty(\mathbb{R}). \tag{3-11}$$

Putting $F(t) = \Omega(\ln t)$, we see that (3-11) is equivalent to the estimate

$$|\langle h, F \rangle| \le C \sum_{n=0}^{N} \max_{t \in \mathbb{R}_+} \big( (1+|\ln t|)^{\kappa} t^n |F^{(n)}(t)| \big), \quad F \in C_0^\infty(\mathbb{R}_+). \tag{3-12}$$

Let us make some comments on this condition. If $h \in L^1_{\mathrm{loc}}(\mathbb{R})$, then estimate (3-12) for $N = 0$ is equivalent to the convergence of integral (1-7) for the same values of $\kappa$. If $H$ is Hilbert–Schmidt, that is,

$$\int_0^\infty |h(t)|^2 t \, dt < \infty,$$

then integral (1-7) converges for any $\kappa > \frac{1}{2}$. Similarly, if $|h(t)| \le C t^{-1}$, then integral (1-7) converges for any $\kappa > 1$.

For the proof of (3-12) in the general case, we use the following elementary result. Its proof is given in Appendix A.

**Lemma 3.6.** *If $F \in C_0^\infty(\mathbb{R}_+)$, then for an arbitrary $\kappa > 2$, the estimate*

$$\|\Phi^* F\|_{L^1(\mathbb{R})} \le C(\kappa) \sum_{n=0}^{2} \max_{t \in \mathbb{R}_+} \big( (1+|\ln t|)^{\kappa} t^n |F^{(n)}(t)| \big) \tag{3-13}$$

*holds.*

**Corollary 3.7.** *If $h = \Phi\omega$, where $\omega \in L^\infty(\mathbb{R})$, then estimate* (3-12) *holds for $N = 2$ and an arbitrary $\kappa > 2$.*

*Proof.* It suffices to combine the estimates

$$|\langle h, F \rangle| = |\langle \omega, \Phi^* F \rangle| \le \|\omega\|_{L^\infty(\mathbb{R})} \|\Phi^* F\|_{L^1(\mathbb{R})}$$

and (3-13). □

Since, by Theorem 3.4, for a bounded Hankel operator $H$, its kernel $h$ equals $\Phi\omega$ for some $\omega \in L^\infty(\mathbb{R})$, we arrive at the following result.

**Theorem 3.8.** *Suppose that $h \in C_0^\infty(\mathbb{R}_+)'$ and that condition* (3-2) *is satisfied. Then estimate* (3-12) *holds for $N = 2$ and an arbitrary $\kappa > 2$; in particular, $h \in \mathscr{Z}'_+$.*

The following simple example shows that for $N = 0$, estimate (3-12) is in general violated (for all $\kappa$).

**Example 3.9.** Let $h(t) = e^{-it^2}$. Then the corresponding Hankel operator $H$ is bounded because according to the formula $e^{-i(t+s)^2} = e^{-it^2} e^{-i2ts} e^{-is^2}$, it is a product of three bounded operators. Since $h \in L^\infty(\mathbb{R}_+)$, estimate (3-12) for $N = 0$ is equivalent to the convergence of integral (1-7) for the same value of $\kappa$. However, this integral diverges at infinity for all $\kappa$.

Let us show that for $h(t) = e^{-it^2}$, condition (3-12) is satisfied for $N = 1$ and $\kappa = 0$. Integrating by parts, we see that

$$\int_0^\infty h(t)\overline{F(t)}\,dt = -\int_0^\infty h_1(t)\overline{F'(t)}\,dt, \tag{3-14}$$

where the function $h_1(t) = \int_0^t e^{-is^2} ds$ is bounded. Therefore, the integral on the right-hand side of (3-14) is bounded by $\max_{t \in \mathbb{R}_+}\big((1 + |\ln t|)^\kappa t|F'(t)|\big)$ for any $\kappa > 1$.

Note that for $h(t) = e^{-it^2}$, the symbol of $H$ equals $\omega(\mu) = \sqrt{\pi} e^{-\pi i/4} e^{i\mu^2/4}$. More generally, one can consider the class of symbols $\omega(\mu)$ such that $\omega \in C^\infty(\mathbb{R})$, $\omega(\mu) = e^{i\omega_0 \mu^\alpha}$, $\omega_0 > 0$ for large positive $\mu$ and $\omega(\mu) = 0$ for large negative $\mu$. Of course, Hankel operators with such symbols are bounded. Using the stationary phase method, we find that for $\alpha > 1$, the corresponding kernel $h(t)$ has the asymptotics

$$h(t) \sim h_0 t^\beta e^{i\sigma t^\gamma}, \quad t \to \infty, \tag{3-15}$$

where $\beta = (1 - \alpha/2)(\alpha - 1)^{-1}$, $\gamma = \alpha(\alpha - 1)^{-1}$, and $h_0$, $\sigma = \bar{\sigma}$ are some constants. Moreover, $h(t)$ is a bounded function on all finite intervals. Similarly to Example 3.9, it can be checked that for such kernels, condition (3-12) is satisfied for $N = 1$ but not for $N = 0$. The same conclusion is true for $\alpha \in (0, 1)$, because in this case the asymptotic relation (3-15) holds for $t \to 0$.

**3.4.** Here we shall show that, for bounded Hankel operators $H$, the representations (2-20) and (2-21) extend to all $f_1, f_2 \in L^2(\mathbb{R}_+)$. By Theorem 3.8, we have $h \in \mathfrak{X}'_+$. Let $b$ and $s$ be the corresponding $b$- and $s$-functions (see Definition 2.4). Recall that the operator $\Xi$ is defined by formula (1-11). We denote by $K$ the operator of multiplication by the function $\sqrt{\cosh(\pi\xi)/\pi}$ in the space $L^2(\mathbb{R})$. It follows from identity (2-24) and the unitarity of the Mellin transform (2-9) that

$$\|K\,\Xi f\| = \|f\|,$$

and hence the operator $K\,\Xi : L^2(\mathbb{R}_+) \to L^2(\mathbb{R})$ is unitary. Therefore, in view of the identities (2-20) and (2-21), we have the following result.

**Lemma 3.10.** *The inequalities* (3-2),

$$|\langle b, (\mathcal{J}\bar{g}) * g\rangle| \le C\|Kg\|^2 \quad \textit{for all } g \in C_0^\infty(\mathbb{R}), \tag{3-16}$$

*and*

$$|\langle s, u^* u\rangle| \le C\|K\Phi u\|^2 \quad \textit{for all } u \in \mathfrak{X} \tag{3-17}$$

*are equivalent. The Hankel operator corresponding to form* (2-3) *is bounded if and only if one of equivalent estimates* (3-2), (3-16), *or* (3-17) *is satisfied.*

These estimates can be formulated in a slightly different way. Let us introduce the space $\mathcal{E} \subset L^2(\mathbb{R})$ of exponentially decaying functions with the norm $\|g\|_{\mathcal{E}} = \|Kg\|$. Then the space $\mathcal{W} = \Phi^*\mathcal{E}$ consists of functions $u(x)$ admitting the analytic continuation $u(z)$ in the strip $\operatorname{Im} z \in (-\pi/2, \pi/2)$; moreover, functions $u(x+iy)$ have limits in $L^2(\mathbb{R})$ as $y \to \pm\pi/2$. The identity

$$\|\Phi u\|_{\mathcal{E}}^2 = (2\pi)^{-1} \int_{-\infty}^{\infty} \left( \left| u\left(x + i\frac{\pi}{2}\right)\right|^2 + \left| u\left(x - i\frac{\pi}{2}\right)\right|^2 \right) dx =: \|u\|_{\mathcal{W}}^2$$

defines the Hilbert norm on $\mathcal{W}$. We call $\mathcal{W}$ the exponential Sobolev space because it is contained in the standard Sobolev spaces $\mathsf{H}^l(\mathbb{R})$ for all $l$. The operators $\Xi : L^2(\mathbb{R}_+) \to \mathcal{E}$ and $\widehat{\Xi} := \Phi^*\Xi : L^2(\mathbb{R}_+) \to \mathcal{W}$ are of course unitary. Obviously, $\|Kg\|$ and $\|K\Phi u\|$ on the right-hand sides of (3-16) and (3-17) can be replaced by $\|g\|_{\mathcal{E}}$ and $\|u\|_{\mathcal{W}}$, respectively. Note that the inclusions $f \in L^2(\mathbb{R}_+)$, $g = \Xi f \in \mathcal{E}$, and $u = \widehat{\Xi} f \in \mathcal{W}$ are equivalent.

Recall that the operator $B$ is defined by formula (1-9) and $(Su)(x) = s(x)u(x)$. If one of the equivalent estimates (3-2), (3-16), or (3-17) is satisfied, then all operators $H : L^2(\mathbb{R}) \to L^2(\mathbb{R})$, $B : \mathcal{E} \to \mathcal{E}'$, and $S : \mathcal{W} \to \mathcal{W}'$ are bounded. Using that relations $f_n \to f$ in $L^2(\mathbb{R}_+)$, $g_n = \Xi f_n \to g = \Xi f$ in $\mathcal{E}$, and $u_n = \Phi^* g_n \to u = \Phi^* g$ in $\mathcal{W}$ are equivalent, we extend (2-20) and (2-21) to all $f \in L^2(\mathbb{R}_+)$. Thus we have obtained the following result.

**Proposition 3.11.** *If one of equivalent estimates* (3-2), (3-16), *or* (3-17) *is satisfied*, *then the identities*

$$(Hf_1, f_2) = (Bg_1, g_2) = (Su_1, u_2), \quad g_j = \Xi f_j, \quad u_j = \Phi^* g_j$$

*are true for all* $f_1, f_2 \in L^2(\mathbb{R}_+)$.

Let $K_l$ be the operator of multiplication by the function $(1+\xi^2)^{l/2}$. Then estimates (3-16) or (3-17) are satisfied provided

$$|\langle b, (\mathcal{J}\bar{g}) * g\rangle| \le C_l \|K_l g\|^2 \quad \text{or} \quad |\langle s, |u|^2\rangle| \le C_l \|u\|_{\mathsf{H}^l(\mathbb{R})}^2, \tag{3-18}$$

for some $l$; in this case

$$C = C_l \pi \max_{\xi \in \mathbb{R}} \big((1+\xi^2)^l (\cosh(\pi\xi))^{-1}\big).$$

**3.5.** In terms of the sign function, it is possible to give simple sufficient conditions for the boundedness and compactness of Hankel operators.

**Proposition 3.12.** *A Hankel operator $H$ is bounded if its sign function satisfies the condition*

$$s \in L^1(\mathbb{R}) + L^\infty(\mathbb{R}). \tag{3-19}$$

*If $s \in L^\infty(\mathbb{R})$ and $s(x) \to 0$ as $|x| \to \infty$, then $H$ is compact.*

*Proof.* The first statement is obvious because under (3-19), the second estimate (3-18) is satisfied with $l > \frac{1}{2}$. To prove the second statement, we observe that the operator $S\Phi^* K^{-1}$ is compact because both $S$

and $K^{-1}$ are operators of multiplication by bounded functions which tend to zero at infinity. Since the operator $K\,\Xi : L^2(\mathbb{R}_+) \to L^2(\mathbb{R})$ is bounded, it follows from the identity (1-12) that the operator

$$H = (\Xi^* \Phi) S (\Phi^* \Xi) = \Xi^* \Phi (S \Phi^* K^{-1})(K\,\Xi)$$

is also compact. □

Condition (3-19) is of course *not* necessary for the boundedness of $H$. For example, in view of formula (1-4) for Hankel operators $H$ of finite rank, the sign function is a singular distribution.

## 4. Criteria for sign-definiteness

In this section we suppose that $h(t) = \overline{h(t)}$ so that the operator $H$ is formally symmetric. The results of Section 2 allow us to give simple necessary and sufficient conditions for a Hankel operator $H$ to be positive or negative. Moreover, they also provide convenient tools for the calculation of the total multiplicity of the negative and positive spectra of $H$. We often state our results only for the negative spectrum. The corresponding results for the positive spectrum are obtained if $H$ is replaced by $-H$.

**4.1.** Actually, we consider the problem in terms of Hankel quadratic forms rather than Hankel operators. This is both more general and more convenient. As usual, we take a distribution $h \in \mathcal{Z}'_+$ and introduce the $b$-function $b \in C_0^\infty(\mathbb{R})'$ and the $s$-function $s \in \mathcal{Z}'$ as in Definition 2.4.

Below we use the following natural notation. Let $\mathsf{h}[\varphi, \varphi]$ be a real quadratic form defined on a linear set $\mathsf{D}$. We denote by $N_\pm(\mathsf{h}; \mathsf{D})$ the maximal dimension of linear sets $\mathcal{M}_\pm \subset \mathsf{D}$ such that $\pm\mathsf{h}[\varphi, \varphi] > 0$ for all $\varphi \in \mathcal{M}_\pm$, $\varphi \neq 0$. This means that there exists a linear set $\mathcal{M}_\pm$, $\dim \mathcal{M}_\pm = N_\pm(\mathsf{h})$, such that $\pm\mathsf{h}[\varphi, \varphi] > 0$ for all $\varphi \in \mathcal{M}_\pm$, $\varphi \neq 0$; and for every linear set $\mathcal{M}'_\pm$ with $\dim \mathcal{M}'_\pm > N_\pm(\mathsf{h})$ there exists $\varphi \in \mathcal{M}'_\pm$, $\varphi \neq 0$, such that $\pm\mathsf{h}[\varphi, \varphi] \leq 0$. We apply this definition to the forms $h[f, f] = \langle h, \bar{f} \star f \rangle$ defined on $\mathcal{D}$, to $b[g, g] = \langle b, (\mathcal{J}\bar{g}) * g \rangle$ defined on $C_0^\infty(\mathbb{R})$, and to $s[u, u] = \langle s, |u|^2 \rangle$ defined on $\mathcal{Z}$. Of course, if $\mathsf{D}$ is dense in a Hilbert space $\mathcal{H}$ and $\mathsf{h}[\varphi, \varphi]$ is semibounded and closed on $\mathsf{D}$, then for the self-adjoint operator $\mathsf{H}$ corresponding to $\mathsf{h}$, we have $N_\pm(\mathsf{H}) = N_\pm(\mathsf{h}; \mathsf{D})$.

Observe that formula (1-11) establishes a one-to-one correspondence between the sets $\mathcal{D}$ and $C_0^\infty(\mathbb{R})$. Moreover, the Fourier transform establishes a one-to-one correspondence between the sets $C_0^\infty(\mathbb{R})$ and $\mathcal{Z}$. Therefore the following assertion is a direct consequence of Theorem 2.6.

**Theorem 4.1.** *Let $h \in \mathcal{Z}'_+$. Then*

$$N_\pm(h; \mathcal{D}) = N_\pm(b; C_0^\infty(\mathbb{R})) = N_\pm(s; \mathcal{Z}).$$

In particular, we have:

**Theorem 4.2.** *Let $h \in \mathcal{Z}'_+$. Then $\pm\langle h, \bar{f} \star f \rangle \geq 0$ for all $f \in \mathcal{D}$ if and only if $\pm\langle b, (\mathcal{J}\bar{g}) * g \rangle \geq 0$ for all $g \in C_0^\infty(\mathbb{R})$, or $\pm\langle s, u^* u \rangle \geq 0$ for all $u \in \mathcal{Z}$.*

**4.2.** A calculation of the form $s[u, u]$ on analytic functions $u \in \mathcal{Z}$ is not always convenient. Therefore it is desirable to replace the class $\mathcal{Z}$, for example, by the class $C_0^\infty(\mathbb{R})$. Such a replacement is not obvious because for $u \in C_0^\infty(\mathbb{R})$ we only have $g = \Phi u \in \mathcal{Z}$. In this case $(Mf)(\xi) = \Gamma\left(\frac{1}{2} + i\xi\right)^{-1} g(\xi)$ need not even belong to $L^2(\mathbb{R})$, so that $f \notin L^2(\mathbb{R}_+)$. Nevertheless, under the additional assumption $s \in \mathcal{S}'$, we have the following assertion.

**Lemma 4.3.** *If $s \in \mathcal{S}'$, then $N_\pm(s; \mathcal{Z}) = N_\pm(s; C_0^\infty(\mathbb{R})) = N_\pm(s; \mathcal{S})$.*

*Proof.* Since $\Phi : \mathcal{Z} \to C_0^\infty(\mathbb{R})$, $\Phi^* : C_0^\infty(\mathbb{R}) \to \mathcal{Z}$ and $\mathcal{S}$ is invariant with respect to the Fourier transform $\Phi$, it suffices, for example, to show that $N_\pm(s; \mathcal{Z}) = N_\pm(s; \mathcal{S})$. The inequality $N_\pm := N_\pm(s; \mathcal{S}) \geq N_\pm(s; \mathcal{Z})$ is obvious because $\mathcal{Z} \subset \mathcal{S}$.

Let us prove the opposite inequality. Consider for definiteness the sign "+". Let $\mathcal{L}_+ \subset \mathcal{S}$, and let $s[u, u] > 0$ for all $u \in \mathcal{L}_+$, $u \neq 0$. Suppose first that $N := \dim \mathcal{L}_+ < \infty$ and choose elements $u_1, \dots, u_N \in \mathcal{L}_+$ such that $s[u_j, u_k] = \delta_{j,k}$ for all $j, k = 1, \dots, N$. Let us construct elements $u_j^{(\epsilon)} \in \mathcal{Z}$ such that $u_j^{(\epsilon)} \to u_j$ and hence $u_j^{(\epsilon)} \bar{u}_k^{(\epsilon)} \to u_j \bar{u}_k$ in $\mathcal{S}$ as $\epsilon \to 0$ for $j, k = 1, \dots, N$. Since $s \in \mathcal{S}'$, we see that $s[u_j^{(\epsilon)}, u_k^{(\epsilon)}] \to \delta_{j,k}$ as $\epsilon \to 0$. For an arbitrary $\sigma > 0$, we can choose $\epsilon$ such that $\left| s[u_j^{(\epsilon)}, u_k^{(\epsilon)}] - \delta_{j,k} \right| \leq \sigma$. Then for arbitrary $\lambda_1, \dots, \lambda_N \in \mathbb{C}$, we have

$$s\Big[\sum_{j=1}^N \lambda_j u_j^{(\epsilon)}, \sum_{j=1}^N \lambda_j u_j^{(\epsilon)}\Big] = \sum_{j=1}^N |\lambda_j|^2 s[u_j^{(\epsilon)}, u_j^{(\epsilon)}] + 2\,\mathrm{Re} \sum_{\substack{j,k=1 \\ j \neq k}}^N \lambda_j \bar{\lambda}_k s[u_j^{(\epsilon)}, u_k^{(\epsilon)}]$$
$$\geq (1 - \sigma) \sum_{j=1}^N |\lambda_j|^2 - 2\sigma \sum_{\substack{j,k=1 \\ j \neq k}}^N \lambda_j \bar{\lambda}_k \geq \big(1 - (2N-1)\sigma\big) \sum_{j=1}^N |\lambda_j|^2.$$

Thus elements $u_1^{(\epsilon)}, \dots, u_N^{(\epsilon)}$ are linearly independent if $(2N-1)\sigma < 1$. The same inequality shows that $s[u, u] > 0$ on all vectors $u \neq 0$ in the space $\mathcal{L}_+^{(\epsilon)}$ spanned by $u_1^{(\epsilon)}, \dots, u_N^{(\epsilon)}$.

If $N_+ = \infty$, then the same construction works on every finite-dimensional subspace of $\mathcal{L}_+$ where $s[u, u] > 0$. This yields a space $\mathcal{L}_+^{(\epsilon)} \subset \mathcal{Z}$ of arbitrarily large dimension where $s[u, u] > 0$. ☐

Putting together this lemma with Theorem 4.1, we obtain the following result.

**Theorem 4.4.** *Let $h \in \mathcal{Z}'_+$. Suppose that $b \in \mathcal{S}'$ or, equivalently, that $s \in \mathcal{S}'$. Then $N_\pm(h; \mathcal{D}) = N_\pm(s; C_0^\infty(\mathbb{R}))$.*

In many cases the following consequence of Theorem 4.4 is convenient. According to Proposition 3.12, under the assumptions of Theorem 4.5, $H$ is defined as the bounded self-adjoint operator corresponding to the form $\langle h, \bar{f} \star f \rangle$. Therefore $N_\pm(h; \mathcal{D}) = N_\pm(H)$ is the total multiplicity of the (strictly) positive spectrum for the sign "+" and of the (strictly) negative spectrum for the sign "−" of the operator $H$. For definiteness, we consider the negative spectrum.

**Theorem 4.5.** *Let $h \in \mathcal{Z}'_+$, and let the corresponding sign function satisfy condition* (3-19). *If $s(x) \geq 0$, then the operator $H$ is positive. If $s(x) \leq -s_0 < 0$ for almost all $x$ in some interval $\Delta \subset \mathbb{R}$, then $N_-(H) = \infty$.*

*Proof.* If $s(x) \geq 0$, then $H \geq 0$ according to the second relation in (2-22).

Let $s(x) \leq -s_0 < 0$ for $x \in \Delta$, and let $N$ be arbitrary. Choose a function $\varphi \in C_0^\infty(\mathbb{R})$ such that $\varphi(x) = 1$ for $x \in [-\delta, \delta]$ and $\varphi(x) = 0$ for $x \notin [-2\delta, 2\delta]$, where $\delta = \delta_N$ is a sufficiently small number. Let points

$\alpha_j \in \Delta$, $j = 1, \dots, N$, be such that $\alpha_{j+1} - \alpha_j = \alpha_j - \alpha_{j-1}$ for $j = 2, \dots, N-1$. Set $\Delta_j = (\alpha_j - \delta, \alpha_j + \delta)$, $\widetilde{\Delta}_j = (\alpha_j - 2\delta, \alpha_j + 2\delta)$. For a sufficiently small $\delta$, we may suppose that $\widetilde{\Delta}_j \subset \Delta$ for all $j = 1, \dots, N$ and that $\widetilde{\Delta}_{j+1} \cap \widetilde{\Delta}_j = \varnothing$ for $j = 1, \dots, N-1$. We set $\varphi_j(x) = \varphi(x - \alpha_j)$. Since $s(x) \le -s_0 < 0$ for $x \in \Delta$, we have

$$\langle s, |\varphi_j|^2 \rangle = \int_{-\infty}^{\infty} s(x) |\varphi_j(x)|^2 \, dx \le -2\delta s_0 < 0. \tag{4-1}$$

The functions $\varphi_1, \dots, \varphi_N$ have disjoint supports, and hence $\langle s, |u|^2 \rangle < 0$ for an arbitrary nontrivial linear combination $u$ of the functions $\varphi_j$. Therefore, combining Theorem 4.1 and Theorem 4.4, we obtain the second statement of the theorem. □

Theorem 4.5 can be reformulated, although in a weaker form, in terms of the functions $b(\xi)$ and even $h(t)$. Suppose, for example, that

$$b \in L^1(\mathbb{R}). \tag{4-2}$$

Then $b(\xi)$'s Fourier transform $s(x)$ is a continuous function which tends to 0 as $|x| \to \infty$. The convolution operator $B$ defined by formula (1-9) is bounded in $L^2(\mathbb{R})$ and self-adjoint, and

$$\operatorname{spec}(B) = [\min_{x \in \mathbb{R}} s(x), \max_{x \in \mathbb{R}} s(x)].$$

The result below follows directly from Theorem 4.5. Note that by Proposition 3.12, under (4-2) the operator $H$ is compact.

**Proposition 4.6.** *Under* (4-2), *the Hankel operator $H$ is positive if and only if $s(x) \ge 0$. If $\min_{x \in \mathbb{R}} s(x) < 0$, then $H$ necessarily has an infinite number of negative eigenvalues.*

In particular, condition (4-2) is satisfied if

$$h(t) = \frac{\theta(\ln t)}{t}, \quad \text{where } \theta \in \mathcal{Z}.$$

In this case $a = \Phi\theta \in C_0^\infty(\mathbb{R})$, and hence $b \in C_0^\infty(\mathbb{R})$.

**4.3.** For the proof that a Hankel operator is not sign-definite, it is sometimes not even necessary to calculate the sign function $s(x)$ (the Fourier transform of $b(\xi)$). It turns out that if $b(\xi)$ grows as $|\xi| \to \infty$, then the form $b[g, g] = \langle b, \mathcal{J}\bar{g} * g \rangle$ cannot be sign-definite. More precisely, we have the following statement about convolutions with growing kernels $b(-\xi) = \overline{b(\xi)}$.

**Theorem 4.7.** *Let $b = b_0 + b_\infty$, where $b_0 \in C^p(\mathbb{R})'$ for some $p \in \mathbb{Z}_+$ and $b_\infty \in L^\infty_{\mathrm{loc}}(\mathbb{R})$. Suppose that there exists a sequence of intervals $\Delta_n = (r_n - \sigma_n, r_n + \sigma_n)$, where $r_n \to \infty$ (or equivalently $r_n \to -\infty$) and the sequence $\sigma_n$ is bounded such that*

$$\lim_{n \to \infty} \sigma_n^l \min_{\xi \in \Delta_n} \operatorname{Re} b_\infty(\xi) = \infty \quad \text{or} \quad \lim_{n \to \infty} \sigma_n^l \max_{\xi \in \Delta_n} \operatorname{Re} b_\infty(\xi) = -\infty, \tag{4-3}$$

*where $l = 2$ if $p = 0$ or $p = 1$ and $l = p + 1$ if $p \ge 2$. Then for both signs, $N_\pm(b; C_0^\infty(\mathbb{R})) \ge 1$.*

*Proof.* Since $b$ can be replaced by $-b$, we can assume that, for example, the first condition (4-3) is satisfied. Pick a real even function $\varphi \in C_0^\infty(\mathbb{R})$ such that $\varphi(\xi) \geq 0$, $\varphi(\xi) = 1$ for $|\xi| \leq \frac{1}{4}$, and $\varphi(\xi) = 0$ for $|\xi| \geq \frac{1}{2}$, and set

$$g_n(\xi) = \varphi\left(\frac{\xi - \frac{r_n}{2}}{\sigma_n}\right) \pm \varphi\left(\frac{\xi + \frac{r_n}{2}}{\sigma_n}\right). \tag{4-4}$$

An easy calculation shows that

$$((\mathcal{J} g_n) * g_n)(\xi) = 2\sigma_n \psi\left(\frac{\xi}{\sigma_n}\right) \pm \sigma_n \psi\left(\frac{\xi - r_n}{\sigma_n}\right) \pm \sigma_n \psi\left(\frac{\xi + r_n}{\sigma_n}\right), \tag{4-5}$$

where $\psi = (\mathcal{J}\varphi) * \varphi \in C_0^\infty(\mathbb{R})$. The function $\psi(\xi)$ is also even, with $\psi(\xi) \geq 0$, $\psi(\xi) \geq \frac{1}{8}$ for $|\xi| \leq \frac{1}{8}$, and $\psi(\xi) = 0$ for $|\xi| \geq 1$.

Since $|\langle b_0, g\rangle| \leq C\|g\|_{C^p}$, it follows from (4-5) that

$$|\langle b_0, (\mathcal{J} g_n) * g_n\rangle| \leq C\sigma_n^{1-p}. \tag{4-6}$$

Moreover, again according to (4-5), we have

$$\langle b_\infty, (\mathcal{J} g_n) * g_n\rangle = 2\sigma_n^2 \int_{-\infty}^{\infty} b_\infty(\sigma_n \eta)\psi(\eta)\, d\eta \pm 2\sigma_n^2 \int_{-\infty}^{\infty} \operatorname{Re} b_\infty(\sigma_n \eta + r_n)\psi(\eta)\, d\eta. \tag{4-7}$$

The first term on the right-hand side is $O(\sigma_n^2)$. For the second one, we use the estimate

$$32 \int_{-\infty}^{\infty} \operatorname{Re} b_\infty(\sigma_n \eta + r_n)\psi(\eta)\, d\eta \geq \min_{|\xi - r_n| \leq \sigma_n} \operatorname{Re} b_\infty(\xi). \tag{4-8}$$

Let us first choose the sign "+" in (4-4). Then using representation (4-7) and putting together estimates (4-6) and (4-8), we obtain the lower bound

$$\langle b, (\mathcal{J} g_n) * g_n\rangle \geq -c(\sigma_n^{1-p} + \sigma_n^2) + \frac{\sigma_n^2}{16} \min_{|\xi - r_n| \leq \sigma_n} \operatorname{Re} b_\infty(\xi).$$

If $p = 0$ or $p = 1$, then under the first condition in (4-3), the right-hand side here tends to $+\infty$ as $n \to \infty$. If $p \geq 2$, it is bounded from below by

$$\sigma_n^{1-p}\left(-c + \frac{\sigma_n^l}{16} \min_{|\xi - r_n| \leq \sigma_n} \operatorname{Re} b_\infty(\xi)\right),$$

where the expression in the brackets tends again to $+\infty$. Therefore $\langle b, (\mathcal{J} g_n) * g_n\rangle > 0$ for sufficiently large $n$. Similarly choosing the sign "−" in (4-4), we see that $\langle b, (\mathcal{J} g_n) * g_n\rangle < 0$ for sufficiently large $n$. □

**Corollary 4.8.** *Instead of condition* (4-3), *assume that*

$$\lim_{|\xi| \to \infty} \operatorname{Re} b_\infty(\xi) = \infty \quad \text{or} \quad \lim_{|\xi| \to \infty} \operatorname{Re} b_\infty(\xi) = -\infty.$$

*Then for both signs*, $N_\pm(b; C_0^\infty(\mathbb{R})) \geq 1$.

In contrast to Theorem 4.7, there are no restrictions in Corollary 4.8 on the parameter $p$ in the assumption $b_0 \in C^p(\mathbb{R})'$. On the other hand, condition (4-3) permits $\operatorname{Re} b(\xi)$ to tend to $\pm\infty$ only on some system of intervals. Moreover, the lengths of these intervals may tend to zero. In this case, however, the growth of $\operatorname{Re} b(\xi)$ and the decay of these lengths should be correlated and there are restrictions on admissible values of the parameters $p$ and $l$.

Unlike Theorem 4.5, Theorem 4.7 does not guarantee that $N = \infty$; see Section 5.4 for a discussion of various possible cases.

**4.4.** Theorem 4.2 can be combined with the Bochner–Schwartz theorem (see, e.g., Theorem 3 in [Gel'fand and Vilenkin 1964, Chapter II, §3]). It states that a distribution $b \in C_0^\infty(\mathbb{R})'$ satisfying the condition $\langle b, \mathcal{J}\bar{g} * g\rangle \geq 0$ for all $g \in C_0^\infty(\mathbb{R})$ (such $b$ are sometimes called distributions of positive type) is the Fourier transform

$$b(\xi) = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{-ix\xi}\, d\mathsf{M}(x)$$

of a positive measure $d\mathsf{M}(x)$ such that

$$\int_{-\infty}^{\infty} (1+|x|)^{-\varkappa}\, d\mathsf{M}(x) < \infty \tag{4-9}$$

for some $\varkappa$ (that is, of at most polynomial growth at infinity). In particular, this ensures that $b \in \mathcal{S}'$.

Theorem 4.2 implies that if $\langle h, \bar{f} \star f\rangle \geq 0$ for all $f \in \mathcal{D}$, then the distribution $b$ related to $h$ by Definition 2.4 is of positive type. This means that the sign function $s(x)$ of $h(t)$ is determined by the measure $d\mathsf{M}(x)$:

$$\langle s, \varphi\rangle = \int_{-\infty}^{\infty} \overline{\varphi(x)}\, d\mathsf{M}(x), \quad \varphi \in \mathcal{S},$$

that is, $s(x)\, dx = d\mathsf{M}(x)$. Let us define the measure

$$dm(\lambda) = \lambda\, d\mathsf{M}(-\ln\lambda), \quad \lambda \in \mathbb{R}_+. \tag{4-10}$$

It is easy to see that condition (4-9) is equivalent to condition (1-14) on measure (4-10). In terms of distribution (1-6), we have $\sigma(\lambda)\, d\lambda = dm(\lambda)$. Therefore, Theorem 2.10 leads to the following result.

**Theorem 4.9.** *Let $h \in \mathcal{Z}'_+$ and $\langle h, \bar{f} \star f\rangle \geq 0$ for all $f \in \mathcal{D}$. Then $h(t)$ admits the representation* (1-3) *with a positive measure $dm(\lambda)$ on $\mathbb{R}_+$ satisfying for some $\varkappa$ condition* (1-14).

The representation (1-3) is of course a particular case of (1-2). It is much more precise than (1-2), but requires the positivity of $\langle h, \bar{f} \star f\rangle$. Theorem 4.9 shows that the positivity of $\langle h, \bar{f} \star f\rangle$ imposes very strong conditions on $h(t)$. Actually, we have:

**Corollary 4.10.** *Let $h \in \mathcal{Z}'_+$ and $\langle h, \bar{f} \star f\rangle \geq 0$ for all $f \in \mathcal{D}$. Then $h \in C^\infty(\mathbb{R}_+)$ and*

$$(-1)^n h^{(n)}(t) \geq 0 \tag{4-11}$$

*for all $t > 0$ and all $n = 0, 1, 2, \dots$ (such functions are called completely monotonic). The function $h(t)$ admits an analytic continuation in the right half-plane* $\operatorname{Re} t > 0$, *and it is uniformly bounded in every strip*

$\operatorname{Re} t \in (t_1, t_2)$, *where* $0 < t_1 < t_2 < \infty$. *Moreover, for some* $\varkappa \in \mathbb{R}$ *and* $C > 0$, *we have the estimate*

$$h(t) \le C t^{-1}(1 + |\ln t|)^{\varkappa}, \quad t > 0. \tag{4-12}$$

All these assertions are direct consequences of the representation (1-3). In particular, under condition (1-14), we have

$$h(t) \le C \max_{\lambda \ge 0} \bigl(e^{-t\lambda} \lambda (1 + |\ln \lambda|)^{\varkappa}\bigr),$$

which yields (4-12).

Recall that according to the Bernstein theorem (see, e.g., Theorems 5.5.1 and 5.5.2 in [Akhiezer 1965]), condition (4-11) implies that the function $h(t)$ admits the representation (1-3) with some positive measure $dm(\lambda)$ on $[0, \infty)$. Note that condition (4-11) does not impose any restrictions on the measure $dm(\lambda)$ (except that the integral (1-3) is convergent for all $t > 0$).

Under the positivity assumption, the identity (2-21) takes a more precise form.

**Proposition 4.11.** *Let* $h \in \mathcal{Z}'_+$ *and* $\langle h, \bar{f} \star f \rangle \ge 0$ *for all* $f \in \mathcal{D}$. *Then there exists a positive measure* $d\mathsf{M}(x)$ *satisfying condition* (4-9) *for some* $\varkappa$ *such that*

$$\langle h, \bar{f}_1 \star f_2 \rangle = \int_{-\infty}^{\infty} u_1(x) \overline{u_2(x)} \, d\mathsf{M}(x)$$

*for all* $f_j \in \mathcal{D}$, $j = 1, 2$, *and* $u_j = \Phi^* \Xi f_j \in \mathcal{Z}$, *where the mapping* $\Xi$ *is defined by* (1-11).

## 5. Applications and examples

**5.1.** Consider first self-adjoint Hankel operators $H$ of finite rank. Recall that integral kernels of Hankel operators of finite rank are given (this is the Kronecker theorem — see, e.g., Sections 1.3 and 1.8 of [Peller 2003]) by the formula

$$h(t) = \sum_{m=1}^{M} P_m(t) e^{-\alpha_m t}, \tag{5-1}$$

where $\operatorname{Re} \alpha_m > 0$ and $P_m(t)$ are polynomials of degree $K_m$. If $H$ is self-adjoint, that is, $h(t) = \overline{h(t)}$, then the set $\{\alpha_1, \dots, \alpha_M\}$ consists of points lying on the real axis and pairs of points symmetric with respect to it. Let $\operatorname{Im} \alpha_m = 0$ for $m = 1, \dots, M_0$ and $\operatorname{Im} \alpha_m > 0$, $\alpha_{M_1+m} = \bar{\alpha}_m$ for $m = M_0 + 1, \dots, M_0 + M_1$. Thus $M = M_0 + 2M_1$; of course the cases $M_0 = 0$ or $M_1 = 0$ are not excluded. The condition $h(t) = \overline{h(t)}$ also requires that $P_m(t) = \overline{P_m(t)}$ for $m = 1, \dots, M_0$ and $P_{M_1+m}(t) = \overline{P_m(t)}$ for $m = M_0 + 1, \dots, M_0 + M_1$. As is well known and as we shall see below,

$$\operatorname{rank} H = \sum_{m=1}^{M} K_m + M =: r.$$

For $m = 1, \dots, M_0$, we denote by $\mathsf{p}_m = \bar{\mathsf{p}}_m$ the coefficient at $t^{K_m}$ in the polynomial $P_m(t)$.

The following assertion yields an explicit formula for the numbers $N_\pm(H)$. Its proof relies on formula (1-4) for the sigma function of the kernel $h(t) = t^k e^{\alpha t}$ and on the identity $N_\pm(H) = N_\pm(S)$. The detailed proof is given in [Yafaev 2015].

**Theorem 5.1.** *For $m = 1, \dots, M_0$, set*

$$\begin{cases} \mathcal{N}_+^{(m)} = \mathcal{N}_-^{(m)} = \dfrac{K_m+1}{2} & \text{if } K_m \text{ is odd,} \\ \mathcal{N}_+^{(m)} - 1 = \mathcal{N}_-^{(m)} = \dfrac{K_m}{2} & \text{if } K_m \text{ is even and } \mathsf{p}_m > 0, \\ \mathcal{N}_+^{(m)} = \mathcal{N}_-^{(m)} - 1 = \dfrac{K_m}{2} & \text{if } K_m \text{ is even and } \mathsf{p}_m < 0. \end{cases} \tag{5-2}$$

*Then the total numbers $N_\pm(H)$ of (strictly) positive and negative eigenvalues of the operator $H$ are given by the formula*

$$N_\pm(H) = \sum_{m=1}^{M_0} \mathcal{N}_\pm^{(m)} + \sum_{m=M_0+1}^{M_0+M_1} K_m + M_1. \tag{5-3}$$

Formula (5-2) shows that every pair

$$P_m(t)e^{-\alpha_m t} + P_{m+M_1}(t)e^{-\alpha_{m+M_1} t}, \quad m = M_0+1, \dots, M_0+M_1, \tag{5-4}$$

of complex conjugate terms in (5-1) yields $K_m + 1$ positive and $K_m + 1$ negative eigenvalues. The contribution of every real term $P_m(t)e^{-\alpha_m t}$, where $m = 1, \dots, M_0$, also consists of equal numbers $(K_m+1)/2$ of positive and negative eigenvalues if the degree $K_m$ of the polynomial $P_m(t)$ is odd. If $K_m$ is even, then there is one more positive (negative) eigenvalue if $\mathsf{p}_m > 0$ ($\mathsf{p}_m < 0$). In particular, in the question considered, there is no "interference" between different terms $P_m(t)e^{-\alpha_m t}$, $m = 1, \dots, M_0$, and pairs (5-4) in representation (5-1) of the kernel $h(t)$.

According to (5-3), the operator $H$ cannot be sign-definite if $M_1 > 0$. Moreover, according to (5-2), $\mathcal{N}_\pm^{(m)} = 0$ for $m = 1, \dots, M_0$ if and only if $K_m = 0$ and $\mp\mathsf{p}_m > 0$. Therefore we have the following result.

**Corollary 5.2.** *A Hankel operator $H$ of finite rank is positive (negative) if and only if its kernel is given by the formula*

$$h(t) = \sum_{m=1}^{M_0} \mathsf{p}_m e^{-\alpha_m t},$$

*where $\alpha_m > 0$ and $\mathsf{p}_m > 0$ ($\mathsf{p}_m < 0$) for all $m = 1, \dots, M_0$.*

Corollary 5.2 admits different proofs which avoid formula (5-3). For example, one can use that although the functions $P_m(t)e^{-\alpha_m t}$ are analytic in the right half-plane $\operatorname{Re} t > 0$, they are bounded for $t = \tau + i\sigma$ as $\sigma \to \infty$ for a constant $P_m(t)$ only. Therefore, according to Corollary 4.10, such Hankel operators cannot be positive. Alternatively, using formula (5-15) below for the $b$-function of the kernel $t^k e^{-\alpha t}$, one can deduce Corollary 5.2 from Theorem 4.7.

Let us compare formula (5-3) with the result of [Megretskiĭ et al. 1995]. In application to finite-rank operators $H$, this general result implies that the spectra of Hankel operators are characterized by the following condition: the multiplicities of eigenvalues $\lambda \neq 0$ and $-\lambda$ of $H$ do not differ by more than 1. This condition and formula (5-3) mean that there is a certain balance between positive and negative spectra of finite-rank Hankel operators. Nevertheless, neither of these results ensures another one.

**5.2.** Consider now Hankel operators $H$ with kernels (1-15). Since the case $k = 0, 1, \dots$ (finite-rank Hankel operators) has been discussed in the previous subsection, here we suppose that $k \neq 0, 1, \dots$. If $k > -1$, condition (1-7) is satisfied for all $\kappa$, and the operators $H$ are compact (actually, they belong to much better classes of operators). If $k = -1$, then condition (1-7) is satisfied for $\kappa > 1$, and the operators $H$ are bounded but not compact.

Let us calculate the $b$- and $s$-functions of kernels (1-15). If $k > -1$, then function (2-15) equals

$$a(\xi) = (2\pi)^{-1/2} \int_0^\infty t^k e^{-\alpha t} t^{-i\xi}\, dt = (2\pi)^{-1/2} \alpha^{-1-k+i\xi} \Gamma(1+k-i\xi), \tag{5-5}$$

and hence function (2-17) equals

$$b(\xi) = \alpha^{-1-k+i\xi} \frac{\Gamma(1+k-i\xi)}{2\pi\Gamma(1-i\xi)}. \tag{5-6}$$

If $k = -1$, then in accordance with formulas (5-5) and (5-6), we have

$$a(\xi) = (2\pi)^{-1/2}\alpha^{i\xi} \lim_{\varepsilon\to+0} \Gamma(\varepsilon - i\xi), \quad b(\xi) = (2\pi)^{-1}\alpha^{i\xi} i(\xi + i0)^{-1}.$$

This yields the expression

$$\begin{cases} s(x) = 0 & \text{if } x > \beta, \\ s(x) = 1 & \text{if } x < \beta, \end{cases} \quad \text{where } \beta = -\ln\alpha, \tag{5-7}$$

for the function $s = \sqrt{2\pi}\,\Phi^* b$. Formula (5-7) remains true for the Carleman operator $\boldsymbol{C}$ (the Hankel operator with kernel $h(t) = t^{-1}$) when $\alpha = 0$. Indeed, in this case, according to (2-23), the sign function $s(x)$ equals 1.

Next, we calculate the Fourier transform of function (5-6). Assume first that $k \in (-1, 0)$. Then (see, e.g., formula (1.5.12) in [Erdélyi et al. 1953])

$$\int_0^\infty t^{-k-1}(t+1)^{-1+i\xi}\, dt = \frac{\Gamma(-k)\Gamma(1+k-i\xi)}{\Gamma(1-i\xi)}.$$

Making here the change of variables $t + 1 = \alpha^{-1}e^{-x}$, we find that

$$\frac{1}{\Gamma(-k)} \int_{-\infty}^{\infty} (e^{-x} - \alpha)_+^{-k-1} e^{-ix\xi}\, dx = \alpha^{-1-k-i\xi} \frac{\Gamma(1+k-i\xi)}{\Gamma(1-i\xi)}.$$

Passing now to the inverse Fourier transform, we see that for $k \in (-1, 0)$ the sign function $s(x) = s_k(x)$ of kernel (1-15) equals

$$s(x) = \frac{1}{\Gamma(-k)} (e^{-x} - \alpha)_+^{-k-1}. \tag{5-8}$$

Let us verify that this formula remains true for all noninteger $k$. To that end, we assume that (5-8) holds for some noninteger $k > -1$ and check it for $k_1 = k + 1$. Since

$$\Gamma(1 + k_1 - i\xi) = (k_1 - i\xi)\Gamma(1 + k - i\xi),$$

we have

$$s_{k_1}(x) = \alpha^{-1}(k_1 - \partial)s_k(x).$$

Substituting here formula (5-8) for $s_k(x)$ and differentiating this expression, we obtain formula (5-8) for $s_{k_1}(x)$. This concludes the proof of relation (5-8) for all $k \geq -1$.

**Lemma 5.3.** *Let $h(t)$ be given by formula* (1-15), *where $k \notin \mathbb{Z}_+$. Then the sign function is determined by relation* (5-8).

Actually, relation (5-8) remains true for $k \in \mathbb{Z}_+$ if one takes into account that the distribution $(e^{-x} - \alpha)_+^{-k-1}$ has poles at integer points. For example, for $k = 0$ we have $s(x) = \alpha^{-1}\delta(x - \beta)$.

Obviously, $s(x) = 0$ for $x > \beta = -\ln \alpha$. If $k = -1$, then $s(x) = 1$ for $x < \beta$. If $k \in (-1, 0)$, then $s(x) \geq 0$ and $s \in L^1(\mathbb{R})$. Therefore it follows from Theorem 4.5 that $H \geq 0$.

If $k > 0$, then distribution (5-8) does not have a definite sign. Therefore it can be deduced from Theorem 4.2 that the corresponding Hankel operator also is not sign-definite.

Alternatively, for the proof of this result we can use Corollary 4.8. Formula (2-11) implies that function (5-6) has the asymptotics

$$b(\xi) = (2\pi)^{-1}\alpha^{-1-k-i\xi}(-i\xi)^k(1 + O(|\xi|^{-1})), \quad |\xi| \to \infty. \tag{5-9}$$

Making the dilation transformation in (1-15), we can suppose that $\alpha = 1$. Then we have

$$\operatorname{Re} b(\xi) = (2\pi)^{-1} \cos\left(\frac{\pi k}{2}\right)\xi^k + O(\xi^{k-1}), \quad \xi \to +\infty. \tag{5-10}$$

Since $\cos(\pi k/2) \neq 0$ unless $k$ is an integer odd number, this expression tends to $\pm\infty$ if $\pm\cos(\pi k/2) > 0$. Thus Corollary 4.8 for the case $b = b_\infty$ ensures that the Hankel operator $H$ is not sign-definite.

Let us summarize the results obtained.

**Proposition 5.4.** *The Hankel operator with kernel* (1-15) *is positive for $k \in [-1, 0]$, and it is not sign-definite for $k > 0$.*

Actually, using relation (5-8), one can calculate explicitly the numbers $N_\pm(H)$ for all values of $k$ (see [Yafaev 2014b]).

Explicit formulas for the sign functions can also be used to treat more complicated Hankel operators. For example, in view of (5-7), the following assertion directly follows from Theorem 4.5.

**Example 5.5.** The Hankel operator with kernel

$$h(t) = t^{-1}(e^{-\alpha_1 t} - \gamma e^{-\alpha_2 t}), \quad \gamma \geq 0,$$

is positive if and only if $\alpha_2 \geq \alpha_1 \geq 0$ and $\gamma \leq 1$.

**5.3.** In this subsection, we consider the Hankel operator $H$ with kernel (1-16). Condition (1-7) is now fulfilled for all $\kappa$, and the operator $H$ belongs of course to the Hilbert–Schmidt class (actually, to much better classes). Observe that

$$a(\xi) = (2\pi)^{-1/2} \int_0^\infty e^{-t^r} t^{-i\xi}\, dt = (2\pi)^{-1/2} r^{-1} \Gamma\Big(\frac{1-i\xi}{r}\Big)$$

and define, as usual, the function $b(\xi)$ by formula (2-17) so that

$$b(\xi) = (2\pi r)^{-1} \frac{\Gamma\Big(\dfrac{1-i\xi}{r}\Big)}{\Gamma(1-i\xi)}. \tag{5-11}$$

Consider first the case $r > 1$. It follows from the Stirling formula (2-11) that for all $r > 1$, the modulus of function (5-11) exponentially grows and the periods of its oscillations tend to zero only logarithmically as $|\xi| \to \infty$. Therefore, Theorem 4.7 implies that the Hankel operator with kernel (1-16) is not sign-definite.

The Hankel operator $H$ with kernel $h(t) = e^{-t^2}$ can also be treated (see Appendix B) in a completely different way, which is perhaps also of some interest. This method shows that both positive and negative spectra of the operator $H$ are infinite.

If $r = 1$, then $h(t) = e^{-t}$ yields a positive Hankel operator of rank 1.

Let us now consider the case $r < 1$. Again according to the Stirling formula (2-11), function (5-11) belongs to $L^1(\mathbb{R})$, so that its Fourier transform

$$s(x) = (2\pi r)^{-1} \int_{-\infty}^{\infty} \frac{\Gamma\Big(\dfrac{1-i\xi}{r}\Big)}{\Gamma(1-i\xi)} e^{ix\xi}\, d\xi =: I_r(x) \tag{5-12}$$

is a continuous function which tends to 0 as $|x| \to \infty$. Therefore, by Proposition 4.6, the corresponding Hankel operator $H$ is nonnegative if and only if $I_r(x) \geq 0$ for all $x \in \mathbb{R}$.

It turns out that $I_r(x) \geq 0$. Surprisingly, we have not found a proof of this fact in the literature, but it follows from our results. Only for $r = \frac{1}{2}$, integral (5-12) can be explicitly calculated. Indeed, according to formula (1.2.15) of [Erdélyi et al. 1953],

$$\frac{\Gamma(2(1-i\xi))}{\Gamma(1-i\xi)} = 2^{1-2i\xi} \pi^{-1/2} \Gamma\big(\tfrac{3}{2} - i\xi\big).$$

Therefore it follows from formula (2-25) that

$$I_{1/2}(x) = 2^{-1} \pi^{-1/2} e^{3x/2} e^{-e^x/4}, \tag{5-13}$$

which is of course positive.

For an arbitrary $r \in (0, 1)$, one can proceed from the Bernstein theorem on completely monotonic functions (see Section 4.4). Observe that if

$$\psi(t) = t^{-p} e^{-t^r}, \quad p \geq 0, \tag{5-14}$$

then

$$\psi'(t) = -pt^{-p-1}e^{-t^r} - rt^{-p+r-1}e^{-t^r}.$$

Further differentiations of $\psi(t)$ change the sign and yield sums of terms having the form (5-14). Thus the function $h(t) = e^{-t^r}$ satisfies, for all $n$, condition (4-11), and hence admits the representation (1-3) with some positive measure $dm(\lambda)$. It follows from (1-3) that

$$(Hf, f) = \int_0^\infty |(\mathsf{L}f)(\lambda)|^2 \, dm(\lambda) \geq 0, \quad \text{for all } f \in C_0^\infty(\mathbb{R}_+),$$

where $\mathsf{L}$ is the Laplace transform (2-28). Since the operator $H$ is bounded, this implies that $H \geq 0$.

Thus we have obtained the following result.

**Proposition 5.6.** *The Hankel operator with kernel* (1-16) *is positive for* $r \in (0, 1]$, *and it is not sign-definite for* $r > 1$.

Putting together this result with Theorem 4.5, we see that integral (5-12) is positive for all $r \in (0, 1)$. Our indirect proof of this fact looks curiously enough.

**5.4.** Let us now discuss convolution operators with growing kernels $b(\xi)$. We emphasize that condition (4-3) does not guarantee that the numbers $N_\pm(b; C_0^\infty(\mathbb{R}))$ are infinite. Indeed, consider the kernel $h(t) = t^k e^{-\alpha t}$, where $k$ is a positive integer. Formula (5-6) shows that for $\operatorname{Im} \alpha = 0$, the corresponding $b$-function

$$b(\xi) = (2\pi)^{-1} \alpha^{-1-k+i\xi} (1 - i\xi) \cdots (k - i\xi) \tag{5-15}$$

has power asymptotics as $|\xi| \to \infty$. According to Theorem 5.1, the positive and negative spectra of the Hankel operator $H$ with the kernel $h(t)$ are finite; for example, $H$ has exactly $(k+1)/2$ positive and negative eigenvalues if $k$ is odd. Moreover, if $\operatorname{Im} \alpha \neq 0$, then in view of (5-15), the function $b(\xi)$ exponentially grows as $\xi \to +\infty$ or $\xi \to -\infty$. Nevertheless, the Hankel operator $H$ with kernel $h(t) = t^k (e^{-\alpha t} + e^{-\bar{\alpha} t})$ has exactly $k+1$ positive and negative eigenvalues.

On the other hand, for kernel (5-11), where $r = 2$, we have $N_\pm(b; C_0^\infty(\mathbb{R})) = \infty$. This follows from Theorem 4.1 because, by Proposition B.1, the Hankel operator with kernel $h(t) = e^{-t^2}$ has an infinite number of positive and negative eigenvalues.

A similar phenomenon occurs for Hankel operators with nonsmooth kernels. This is discussed in the next section. However, in general, the calculation of the numbers $N_\pm(b; C_0^\infty(\mathbb{R}))$ for convolutions with kernels $b(\xi)$ growing and oscillating at infinity looks like an open problem.

## 6. Hankel operators with nonsmooth kernels

According to Corollary 4.10, a Hankel operator $H$ can be sign-definite only for kernels $h \in C^\infty(\mathbb{R}_+)$. Here we show that if $h(t)$ or one of its derivatives $h^{(l)}(t)$ has a jump discontinuity at some point $t_0 > 0$, then $H$ has an infinite number of both positive and negative eigenvalues accumulating to zero. Moreover, we calculate their asymptotic behavior.

**6.1.** We start with a distributional kernel. Let the symbol (see the definition in Section 3.2) of the Hankel operator $H$ be defined by the formula $\omega(\mu) = e^{it_0\mu}$. Then $h(t) = (2\pi)^{-1/2}(\Phi\omega)(t) = \delta(t - t_0)$. It follows from (1-1) that $H = 0$ for $t_0 \le 0$ and

$$(Hf)(t) = f(t_0 - t)$$

for $t_0 > 0$, which we suppose from now on. For such $h(t)$, condition (3-12) is satisfied for $N = 0$ and $\kappa = 0$.

The operator $H$ admits an explicit spectral analysis. Indeed, observe first that $(Hf)(t) = 0$ for $t > t_0$ and hence $L^2(t_0, \infty) \subset \operatorname{Ker} H$. Since $H^2 f = f$ for $f \in L^2(0, t_0)$, the restriction of $H$ on its invariant subspace $L^2(0, t_0)$ may have only $\pm 1$ as eigenvalues. Obviously, the eigenspace $\mathcal{H}_\pm$ of $H$ corresponding to the eigenvalue $\pm 1$ consists of all functions $f(t)$ such that $f(t) = \pm f(t_0 - t)$. Since

$$\mathcal{H}_+ \oplus \mathcal{H}_- \oplus L^2(t_0, \infty) = L^2(\mathbb{R}_+),$$

the spectrum of $H$ consists of the eigenvalues 0, 1, $-1$ of infinite multiplicity each.

**6.2.** For a compact operator $H$, let us denote by $\lambda_n^{(+)}$ $(-\lambda_n^{(-)})$ its positive (negative) eigenvalues. Positive (negative) eigenvalues are of course enumerated in decreasing (increasing) order with multiplicities taken into account.

Let us start with the explicit kernel

$$h(t) = (t_0 - t)^l \text{ for } t \le t_0, \quad h(t) = 0 \text{ for } t > t_0, \tag{6-1}$$

where $l$ is one of the numbers $l = 0, 1, \dots$. Then

$$(Hf)(t) = \int_0^{t_0 - t} (t_0 - t - s)^l f(s)\, ds, \quad t \in (0, t_0),$$

and $(Hf)(t) = 0$ for $t \ge t_0$. For such $h(t)$, the symbol equals

$$\omega(\mu) = \int_0^{t_0} e^{i\mu t}(t_0 - t)^l\, dt = l!(i\mu)^{-l-1}\Big(e^{i\mu t_0} - \sum_{k=0}^{l} \frac{1}{k!}(i\mu t_0)^k\Big).$$

It is a smooth function oscillating as $|\mu| \to \infty$.

It follows from (1-5) that the $b$-function of the operator $H$ equals

$$b(\xi) = \frac{l!\, t_0^{l+1-i\xi}}{2\pi\Gamma(l + 2 - i\xi)} \tag{6-2}$$

(if $h(t) = \delta(t - t_0)$, then this formula is true with $l = -1$). So according to Theorem 4.7, we have $N_\pm(H) > 0$. Actually, the spectrum of $H$ consists of an infinite number of positive and negative eigenvalues denoted by $\lambda_n^{(\pm)}$, and we will find their asymptotic behavior as $n \to \infty$.

Let us consider the spectral problem $Hf = \lambda f$, that is,

$$\int_0^{t_0 - t} (t_0 - t - s)^l f(s)\, ds = \lambda f(t), \quad t \in (0, t_0). \tag{6-3}$$

Differentiating this equation $k$ times, we find that

$$(-1)^k l(l-1)\cdots(l-k+1)\int_0^{t_0-t}(t_0-t-s)^{l-k}f(s)\,ds=\lambda f^{(k)}(t) \tag{6-4}$$

for $k=1,\dots,l$. Differentiating (6-4), where $k=l$ once more, we see that

$$l!\,f(t)=\lambda(-1)^{l+1}f^{(l+1)}(t_0-t),\quad t\in(0,t_0). \tag{6-5}$$

Setting $t=t_0$ in (6-3) and (6-4), we obtain the boundary conditions

$$f(t_0)=f'(t_0)=\cdots=f^{(l)}(t_0)=0. \tag{6-6}$$

Conversely, if a function $f(t)$ satisfies (6-5) and boundary conditions (6-6), it satisfies also (6-3). This leads to the following intermediary result.

**Lemma 6.1.** *Let the operator $A$ be defined on the Sobolev class $\mathsf{H}^{l+1}(0,t_0)$ by the equation*

$$(Af)(t)=(-1)^{l+1}f^{(l+1)}(t_0-t). \tag{6-7}$$

*Considered with boundary conditions* (6-6), *it is self-adjoint in the space $L^2(0,t_0)$, and its eigenvalues $\alpha_n^{(\pm)}$ are linked to eigenvalues $\lambda_n^{(\pm)}$ of the Hankel operator $H$ with kernel* (6-1) *by the equation $\alpha_n^{(\pm)}=l!\,(\lambda_n^{(\pm)})^{-1}$.*

**6.3.** Clearly, $A^2$ is a differential operator and the asymptotic behavior of its eigenvalues is described by the Weyl formula. However, to find the eigenvalue asymptotics of the operator $A$, we have to distinguish between positive and negative eigenvalues. For this reason, it is convenient to introduce an auxiliary operator $\widetilde{A}$ with symmetric (with respect to the point 0) spectrum having the same eigenvalue asymptotics as $A$.

We define $\widetilde{A}$ by the same formula (6-7) as $A$ but consider it on functions in $\mathsf{H}^{l+1}(0,t_0/2)\oplus\mathsf{H}^{l+1}(t_0/2,t_0)$ satisfying the boundary conditions

$$f^{(k)}(0)=f^{(k)}\Big(\frac{t_0}{2}-0\Big),\quad f^{(k)}\Big(\frac{t_0}{2}+0\Big)=f^{(k)}(t_0), \tag{6-8}$$

where $k=0,\dots,l$ for $l$ even, and

$$f^{(k)}(0)=f^{(k)}\Big(\frac{t_0}{2}-0\Big)=0,\quad f^{(k)}\Big(\frac{t_0}{2}+0\Big)=f^{(k)}(t_0)=0, \tag{6-9}$$

where $k=0,\dots,(l-1)/2$ for $l$ odd. The operator $\widetilde{A}$ is self-adjoint in the space $L^2(0,t_0/2)\oplus L^2(t_0/2,t_0)$, and it is determined by the matrix

$$\widetilde{A}=\begin{pmatrix}0 & A_{1,2}\\ A_{2,1} & 0\end{pmatrix},\quad A_{1,2}=A_{2,1}^*, \tag{6-10}$$

where $A_{2,1}:L^2(0,t_0/2)\to L^2(t_0/2,t_0)$. The operator $A_{2,1}$ is again given by relation (6-7) on functions in $\mathsf{H}^{l+1}(0,t_0/2)$ satisfying conditions (6-8) or (6-9) at the points 0 and $t_0/2-0$. It follows from formula

(6-10) that the spectrum of the operator $\widetilde{A}$ is symmetric with respect to the point 0 and consists of eigenvalues $\pm a_n$, where $a_n^2$ are eigenvalues of the operator $A_{2,1}^* A_{2,1} =: \boldsymbol{A}$.

An easy calculation shows that $\boldsymbol{A}$ is the differential operator $\boldsymbol{A} = (-1)^{l+1}\partial^{2l+2}$ in the space $L^2(0, t_0/2)$ defined on functions in the class $\mathsf{H}^{2l+2}(0, t_0/2)$ satisfying the boundary conditions $f^{(k)}(0) = f^{(k)}(t_0/2)$, where $k = 0, \dots, 2l+1$ for $l$ even, and the boundary conditions

$$f^{(k)}(0) = f^{(k)}\left(\frac{t_0}{2}\right) = f^{(l+1+k)}(0) = f^{(l+1+k)}\left(\frac{t_0}{2}\right) = 0,$$

where $k = 0, \dots, (l-1)/2$ for $l$ odd. The asymptotic formula for the eigenvalues $a_n^2$ of $\boldsymbol{A}$ is given by the Weyl formula, that is,

$$a_n = (2\pi t_0^{-1} n)^{l+1}(1 + O(n^{-1})).$$

Let us now observe that the operators $A$ and $\widetilde{A}$ are self-adjoint extensions of a symmetric operator $A_0$ with finite deficiency indices $(2l+2, 2l+2)$. For example, $A_0$ can be defined by formula (6-7) on $C^\infty$-functions vanishing in some neighborhoods of the points 0, $t_0/2$, and $t_0$. Therefore, the operators $A$ and $\widetilde{A}$ have the same spectral asymptotics. Taking Lemma 6.1 into account, we obtain the following result.

**Lemma 6.2.** *The eigenvalues of the Hankel operator $H = H(t_0)$ with kernel* (6-1) *have the asymptotic behavior*

$$\lambda_n^{(\pm)} = l!\,(2\pi)^{-l-1} t_0^{l+1} n^{-l-1}(1 + O(n^{-1})). \tag{6-11}$$

**Remark 6.3.** It is interesting that the asymptotic coefficient in (6-11) is proportional to $t_0^{l+1}$, where $t_0$ is the jump point. However, this fact is not surprising, because the operators $H(t_0)$ are related by the equation $H(t_0) = t_0^{l+1}\mathsf{D}(t_0)^* H(1)\mathsf{D}(t_0)$, where $\mathsf{D}(t_0)$, $(\mathsf{D}(t_0)f)(t) = \sqrt{t_0}\, f(t_0 t)$, is the unitary operator of dilations.

**Remark 6.4.** In the case $l = 0$ we have the explicit formulas

$$\lambda_n^{(+)} = (2\pi)^{-1} t_0\big(n - \tfrac{3}{4}\big)^{-1}, \quad \lambda_n^{(-)} = (2\pi)^{-1} t_0\big(n - \tfrac{1}{4}\big)^{-1}, \quad n = 1, 2, \dots.$$

**6.4.** Now we extend the asymptotics (6-11) to general Hankel operators whose kernels (or their derivatives) have jumps of continuity at a single positive point. To that end, we combine Lemma 6.2 with Theorem 7.4 in Chapter 6 of [Peller 2003]. This theorem implies that singular values $s_n(V)$ of a Hankel operator $V$ satisfy the bound

$$s_n(V) = o(n^{-l-1})$$

if $V$ has a symbol belonging to the Besov class $\boldsymbol{B}_{p,p}^{l+1}(\mathbb{R})$, where $p = (l+1)^{-1}$. By the Weyl theorem on the stability of the power asymptotics of eigenvalues, adding such an operator $V$ to the Hankel operator with kernel (6-1) cannot change the leading asymptotic term in formula (6-11). This yields the following result.

**Theorem 6.5.** *Let $l \in \mathbb{Z}_+$, and let $v(t)$ be the Fourier transform of a function in the Besov class $\boldsymbol{B}^{l+1}_{p,p}(\mathbb{R})$, where $p = (l+1)^{-1}$. Set*

$$h(t) = h_0(t_0 - t)^l + v(t)$$

*for $t \le t_0$ and $h(t) = v(t)$ for $t > t_0$. Then eigenvalues of the Hankel operator $H$ have the asymptotics*

$$\lambda_n^{(\pm)} = |h_0| l! \, (2\pi)^{-l-1} t_0^{l+1} n^{-l-1} (1 + o(1))$$

*as $n \to \infty$.*

We emphasize that under the assumptions of this theorem, the leading terms in the asymptotics of the positive and negative eigenvalues are the same. Of course if $h(t)$ becomes smoother ($l$ increases), then eigenvalues of the Hankel operator $H$ decrease faster as $n \to \infty$. Observe that for $l = 0$ (when the kernel itself is discontinuous), the Hankel operator $H$ does not belong to the trace class.

We finally note that, under assumptions close to those of Theorem 6.5, the asymptotic behavior of the singular values of the Hankel operator $H$ was obtained long ago in [Glover et al. 1990] by a completely different method.

## 7. Perturbations of the Carleman operator

In this section we consider operators $H = H_0 + V$, where $H_0$ is the Carleman operator $\boldsymbol{C}$ (or a more general operator) and the perturbation $V$ belongs to one of the classes introduced in Section 5. Various objects related to the operator $H_0$ will be endowed with the index "0", and objects related to the operator $V$ will be endowed with the index "$v$".

**7.1.** For perturbations $V$ of finite rank, we have the following result.

**Theorem 7.1.** *Let the sign function $s_0(x)$ of a Hankel operator $H_0$ be bounded and positive. Let the kernel $v(t)$ of $V$ be given by the formula*

$$v(t) = \sum_{m=1}^{M} P_m(t) e^{-\alpha_m t},$$

*where $P_m(t)$ is a polynomial of degree $K_m$. Put $H = H_0 + V$ and define the numbers $\mathcal{N}_-^{(m)}$ by formula* (5-2). *Then $N_-(H)$ is given by formula* (5-3).

**Corollary 7.2.** *Under the assumptions of Theorem 7.1, we have $N_-(H) = N_-(V)$. In particular, $H \ge 0$ if and only if $V \ge 0$.*

Of course, in the case $H_0 = 0$, Theorem 7.1 reduces to Theorem 5.1. Since for the Carleman operator $\boldsymbol{C}$ the sign function equals 1, Theorem 7.1 applies to $H_0 = \boldsymbol{C}$.

The inequality $N_-(H) \le N_-(V)$ is obvious because $H_0 \ge 0$. On the other hand, the opposite inequality $N_-(H) \ge N_-(V)$ looks surprising because the operator $H_0$, which may have the continuous spectrum, is much "stronger" than the operator $V$ of finite rank. At a heuristic level, the equality $N_-(H) = N_-(V)$ can be explained by the fact that the supports of the sign functions $s_0(x)$ and $s_v(x)$ are essentially disjoint.

Very loosely speaking, this means that the operators $H_0$ and $V$ "live in orthogonal subspaces", and hence the positive operator $H_0$ does not affect the negative spectrum of $V$. The detailed proof of Theorem 7.1, as well as that of Theorem 5.1, is given in [Yafaev 2015].

**7.2.** Let $\boldsymbol{C}$ be the Carleman operator, and let $V$ be the Hankel operator with kernel

$$v(t) = t^k e^{-\alpha t}, \quad \alpha > 0,\ k > -1. \tag{7-1}$$

The operator $V$ is compact, and hence the essential spectrum $\operatorname{spec}_{\mathrm{ess}}(H_\gamma)$ of the operator

$$H_\gamma = \boldsymbol{C} - \gamma V, \quad \gamma \in \mathbb{R}, \tag{7-2}$$

coincides with the interval $[0, \pi]$. Since the sign function of the operator $\boldsymbol{C}$ equals 1, the sign function $s_\gamma$ of the operator $H_\gamma$ equals

$$s_\gamma(x) = 1 - \gamma s_v(x),$$

where the function $s_v(x)$ is given by formula (5-8).

Let first $k \in (-1, 0)$. Observe that $s_v(x)$ is continuous for $x < \beta = -\ln\alpha$ and $s_v(x) \to +\infty$ as $x \to \beta - 0$ but $s_v \in L^1(\mathbb{R})$. Thus the function $s_\gamma(x)$ goes to $-\infty$ as $x \to \beta - 0$ for all $\gamma > 0$, and hence it follows from Theorem 4.5 that the operator $H_\gamma$ has infinite negative spectrum for all $\gamma > 0$.

In the case $k > 0$, we use the formula

$$b(\xi) = \delta(\xi) + b_v(\xi) \tag{7-3}$$

and apply Theorem 4.7 (more precisely, Corollary 4.8) with $b_0(\xi) = \delta(\xi)$ and $b_\infty(\xi) = b_v(\xi)$. Since $b_0 \in C(\mathbb{R})'$ and $b_\infty$ has the asymptotic behavior (5-9), the operator $H_\gamma$ has a negative spectrum for all $\gamma \neq 0$.

Let us summarize the results obtained.

**Proposition 7.3.** *Let $H_\gamma = \boldsymbol{C} - \gamma V$, where $V$ is the Hankel operator with kernel* (7-1). *Then*:

(1) *If $k \in (-1, 0)$ and $\gamma > 0$, then the operator $H_\gamma$ has an infinite number of negative eigenvalues.*

(2) *If $k > 0$, then the operator $H_\gamma$ has at least one negative eigenvalue for all $\gamma \neq 0$.*

**7.3.** The result below directly follows from Theorem 4.5.

**Proposition 7.4.** *Suppose that the sign function $s_v(x)$ of a Hankel operator $V$ is continuous and $s_v(x) \to 0$ as $|x| \to \infty$. Then the operator $H_\gamma$ defined by formula* (7-2) *is positive if and only if*

$$\gamma s_v(x) \le 1 \quad \textit{for all } x \in \mathbb{R}.$$

*If this condition is not satisfied, then $H_\gamma$ has infinite negative spectrum.*

We note that by Proposition 3.12, under the assumption of Proposition 7.4 on the sign function $s_v$, the operator $V$ is compact, and hence $\operatorname{spec}_{\mathrm{ess}}(H_\gamma) = [0, \infty)$. Of course, this assumption on $s_v$ is satisfied if $b_v \in L^1(\mathbb{R})$.

**Example 7.5.** Let $v(t) = e^{-t^r}$, where $r < 1$. We have seen in Section 5.3 that its sign function $s_v(x)$ equals $I_r(x)$, where $I_r(x)$ is integral (5-12). Recall that $I_r(x)$ is a nonnegative continuous function of $x \in \mathbb{R}$ and $I_r(x) \to 0$ as $|x| \to \infty$. Set

$$\nu_r = \max_{x \in \mathbb{R}} I_r(x).$$

Then $H_\gamma \geq 0$ if $\gamma \leq \nu_r^{-1}$, and the operator $H_\gamma$ has infinite negative spectrum for all $\gamma > \nu_r^{-1}$. Using the explicit formula (5-13), it is easy to calculate $\nu_{1/2} = 3\sqrt{6/\pi}\, e^{-3/2}$.

In the case $r > 1$ we use formula (7-3). As shown in Section 5.3, the modulus of the function $b_v(\xi)$ exponentially grows and the periods of its oscillations tend to zero only logarithmically as $|\xi| \to \infty$. Therefore Theorem 4.7 yields the following result.

**Proposition 7.6.** *Let* $v(t) = e^{-t^r}$, *where* $r > 1$. *Then the operator* (7-2) *has at least one negative eigenvalue for all* $\gamma \neq 0$.

Thus the results on the negative spectrum of the operator $H_\gamma = \boldsymbol{C} - \gamma V$, where $v(t) = e^{-t^r}$, are qualitatively different for $r < 1$, $r = 1$, and $r > 1$.

## Appendix A: Proof of Lemma 3.6

Set

$$F_\kappa^{(n)} = \max_{t \in \mathbb{R}_+} \big( \langle \ln t \rangle^\kappa t^n |F^{(n)}(t)| \big),$$

where for shortness we use the notation $\langle x \rangle = (1 + |x|)$.

Let us first consider $(\Phi^* F)(\lambda)$ for $\lambda \in (-1, 1) =: I$. We have

$$\sqrt{2\pi}\,(\Phi^* F)(\lambda) = \int_0^a F(t) e^{i\lambda t}\, dt + \int_a^\infty F(t) e^{i\lambda t}\, dt, \quad a = |\lambda|^{-1/2}.$$

The first integral on the right-hand side is bounded by $F_0^{(0)} |\lambda|^{-1/2}$, which belongs to $L^1(I)$. In the second integral, we integrate by parts:

$$\int_a^\infty F(t) e^{i\lambda t}\, dt = i\lambda^{-1} F(a) e^{i\lambda a} + i\lambda^{-1} \int_a^\infty F'(t) e^{i\lambda t}\, dt. \tag{A-1}$$

The first term here is bounded by $C|\lambda|^{-1} \langle \ln \lambda \rangle^{-\kappa} F_\kappa^{(0)}$, which belongs to $L^1(I)$ if $\kappa > 1$. The second term is bounded by

$$|\lambda|^{-1} \int_a^\infty t^{-1} \langle \ln t \rangle^{-\kappa}\, dt\, F_\kappa^{(1)} \leq C |\lambda|^{-1} \langle \ln \lambda \rangle^{-\kappa+1} F_\kappa^{(1)}.$$

It belongs to $L^1(I)$ if $\kappa > 2$. Thus, for all $\kappa > 2$, we have

$$\|\Phi^* F\|_{L^1(I)} \leq C(\kappa) (F_0^{(0)} + F_\kappa^{(1)}). \tag{A-2}$$

Next, we consider $(\Phi^* F)(\lambda)$ for $|\lambda| \geq 1$. Integrating by parts, we see that

$$\sqrt{2\pi}\,(\Phi^* F)(\lambda) = i\lambda^{-1} \int_0^a F'(t) e^{i\lambda t}\, dt + i\lambda^{-1} \int_a^\infty F'(t) e^{i\lambda t}\, dt. \tag{A-3}$$

The first term here is bounded by

$$|\lambda|^{-1}\int_0^a t^{-1}\langle\ln t\rangle^{-\kappa}\,dt\,F_\kappa^{(1)} \le C|\lambda|^{-1}\langle\ln\lambda\rangle^{-\kappa+1}F_\kappa^{(1)}.$$

It belongs to $L^1(\mathbb{R}\setminus I)$ if $\kappa > 2$. In the second integral in (A-3) we once more integrate by parts, that is, we use formula (A-1) with $F(t)$ replaced by $F'(t)$. The function $\lambda^{-2}F'(a)$ is bounded by $|\lambda|^{-3/2}F_0^{(1)}$. For the second term, we use the estimate

$$\left|\lambda^{-2}\int_a^\infty F''(t)e^{i\lambda t}\,dt\right| \le \lambda^{-2}\int_a^\infty t^{-2}\,dt\,F_0^{(2)} = |\lambda|^{-3/2}F_0^{(2)}.$$

Therefore the second term in (A-3) also belongs to $L^1(\mathbb{R}\setminus I)$. Thus, for all $\kappa > 2$, we have

$$\|\Phi^* F\|_{L^1(\mathbb{R}\setminus I)} \le C(\kappa)(F_\kappa^{(1)} + F_0^{(2)}). \tag{A-4}$$

Putting together (A-2) and (A-4), we obtain estimate (3-13).

## Appendix B: The Gaussian kernel

Here we return to the Hankel operator $H$ with kernel $h(t) = e^{-t^2}$ considered in Section 5.3. Since $e^{-(t+s)^2} = e^{-t^2}e^{-2ts}e^{-s^2}$, we have the identity

$$(Hf, f) = (\mathsf{L}\psi, \psi), \tag{B-1}$$

where $\psi(t) = e^{-t^2/2}f(t/\sqrt{2})/\sqrt{2}$ and $\mathsf{L}$ is the Laplace transform defined in the space $L^2(\mathbb{R}_+)$ by formula (2-28). We shall use (B-1) essentially in the same way as the main identity (1-10). It follows from equality (2-29) for $\alpha = \frac{1}{2}$ that $\mathsf{L} = M^*\mathcal{J}\mathbf{\Gamma}_{1/2}M$, where $M$ is the Mellin transform. Therefore the spectrum of $\mathsf{L}$ consists of the interval $[-\gamma, \gamma]$, where, according to (2-24),

$$\gamma = \max_{\xi\in\mathbb{R}}\left|\Gamma\left(\tfrac{1}{2} + i\xi\right)\right| = \sqrt{\pi}\max_{\xi\in\mathbb{R}}(\cosh(\pi\xi))^{-1} = \sqrt{\pi}.$$

This allows us to check the following assertion.

**Proposition B.1.** *The Hankel operator $H$ with kernel $h(t) = e^{-t^2}$ has an infinite number of positive and negative eigenvalues.*

*Proof.* Fix some $\mu \in (0, \sqrt{\pi})$. For an arbitrary $N$, let $\Delta_1^{(+)}, \dots, \Delta_N^{(+)} \subset (\mu, \sqrt{\pi})$ and $\Delta_1^{(-)}, \dots, \Delta_N^{(-)} \subset (-\sqrt{\pi}, -\mu)$ be closed mutually disjoint intervals. Choose functions $\varphi_j^{(\pm)}$ in the spectral intervals $\Delta_j^{(\pm)}$ of the operator $\mathsf{L}$ and such that $\|\varphi_j^{(\pm)}\| = 1$, $j = 1, \dots N$. Let $\varphi^{(\pm)} = \sum_{j=1}^N \alpha_j\varphi_j^{(\pm)}$ be a linear combination of the functions $\varphi_1^{(\pm)}, \dots, \varphi_N^{(\pm)}$. Then

$$\pm(\mathsf{L}\varphi^{(\pm)}, \varphi^{(\pm)}) = \pm\sum_{j=1}^N |\alpha_j|^2(\mathsf{L}\varphi_j^{(\pm)}, \varphi_j^{(\pm)}) \ge \mu\sum_{j=1}^N |\alpha_j|^2\|\varphi_j^{(\pm)}\|^2 = \mu\|\varphi^{(\pm)}\|^2. \tag{B-2}$$

For an arbitrary $\varepsilon > 0$, we can choose $\psi_j^{(\pm)} \in C_0^\infty(\mathbb{R}_+)$ such that $\|\psi_j^{(\pm)} - \varphi_j^{(\pm)}\| < \varepsilon$ for all $j = 1, \dots, N$. Since the functions $\varphi_j^{(\pm)}$ are orthogonal, the functions $\psi_j^{(\pm)}$ are linearly independent if $\varepsilon$ is small enough.

Moreover, it follows from (B-2) that

$$\pm(\mathsf{L}\psi^{(\pm)}, \psi^{(\pm)}) \geq 2^{-1}\mu\|\psi^{(\pm)}\|^2 \tag{B-3}$$

if $\psi^{(\pm)} = \sum_{j=1}^{N} \alpha_j \psi_j^{(\pm)}$ and $\varepsilon$ is small.

Set now $f^{(\pm)}(t) = \sqrt{2}\, e^{t^2} \psi^{(\pm)}(\sqrt{2}t)$. Then $f^{(\pm)} \in L^2(\mathbb{R}_+)$, and according to the identity (B-1), inequality (B-3) implies that $\pm(Hf^{(\pm)}, f^{(\pm)}) > 0$ on the linear subspace of such functions $f^{(\pm)}$ (except $f^{(\pm)} = 0$). This subspace has dimension $N$. Hence the operator $H$ has at least $N$ positive and $N$ negative eigenvalues. Since $N$ is arbitrary, this concludes the proof. □

We emphasize that the operator $H$ is compact while the operator $\mathsf{L}$ has the continuous spectrum. Nevertheless the total multiplicities of their positive and negative spectra are the same (infinite).

## References

[Akhiezer 1965] N. I. Akhiezer, *The classical moment problem and some related questions in analysis*, Hafner, New York, 1965. MR 32 #1518 Zbl 0135.33803

[Bernstein 1929] S. N. Bernstein, “Sur les fonctions absolument monotones”, *Acta Math.* **52**:1 (1929), 1–66. MR 1555269 JFM 55.0142.07

[Carleman 1923] T. Carleman, *Sur les équations intégrales singulières à noyau réel et symétrique*, Uppsala Universitets Årsskrift Matematik och Naturvetenskap (8) **3**, Almqvist and Wiksell, Uppsala, 1923. JFM 49.0272.01

[Erdélyi et al. 1953] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher transcendental functions*, vol. 1, McGraw-Hill, New York, 1953. MR 15,419i Zbl 0051.30303

[Gel’fand and Vilenkin 1964] I. M. Gel’fand and N. Y. Vilenkin, *Generalized functions, 4: Applications of harmonic analysis*, Academic Press, New York, 1964. MR 30 #4152 Zbl 0136.11201

[Glover et al. 1990] K. Glover, J. Lam, and J. R. Partington, “Rational approximation of a class of infinite-dimensional systems, I: Singular values of Hankel operators”, *Math. Control Signals Systems* **3**:4 (1990), 325–344. MR 91e:93054 Zbl 0727.41020

[Howland 1971] J. S. Howland, “Trace class Hankel operators”, *Quart. J. Math. Oxford Ser.* (2) **22** (1971), 147–159. MR 44 #5826 Zbl 0216.41903

[Howland 1992] J. S. Howland, “Spectral theory of operators of Hankel type. I, II”, *Indiana Univ. Math. J.* **41**:2 (1992), 409–426, 427–434. MR 94a:47041 Zbl 0773.47012

[Magnus 1950] W. Magnus, “On the spectrum of Hilbert’s matrix”, *Amer. J. Math.* **72** (1950), 699–704. MR 12,836e Zbl 0041.23805

[Megretskiĭ et al. 1995] A. V. Megretskiĭ, V. V. Peller, and S. R. Treil, “The inverse spectral problem for self-adjoint Hankel operators”, *Acta Math.* **174**:2 (1995), 241–309. MR 96g:47010 Zbl 0865.47015

[Mehler 1881] F. G. Mehler, “Ueber eine mit den Kugel- und Cylinderfunctionen verwandte Function und ihre Anwendung in der Theorie der Elektricitätsvertheilung”, *Math. Ann.* **18**:2 (1881), 161–194. MR 1510098 JFM 13.0779.02

[Nehari 1957] Z. Nehari, “On bounded bilinear forms”, *Ann. of Math.* (2) **65** (1957), 153–162. MR 18,633f Zbl 0077.10605

[Paley and Wiener 1934] R. E. A. C. Paley and N. Wiener, *Fourier transforms in the complex domain*, American Mathematical Society Colloquium Publications **19**, American Mathematical Society, Providence, RI, 1934. MR 98a:01023 Zbl 0011.01601

[Peller 2003] V. V. Peller, *Hankel operators and their applications*, Springer, New York, 2003. MR 2004e:47040 Zbl 1030.47002

[Power 1982] S. C. Power, *Hankel operators on Hilbert space*, Research Notes in Mathematics **64**, Pitman, Boston, 1982. MR 84e:47037 Zbl 0489.47011

[Rosenblum 1958a] M. Rosenblum, “On the Hilbert matrix, I”, *Proc. Amer. Math. Soc.* **9**:1 (1958), 137–140. MR 20 #1139 Zbl 0080.10504

[Rosenblum 1958b] M. Rosenblum, “On the Hilbert matrix, II”, *Proc. Amer. Math. Soc.* **9**:4 (1958), 581–585. MR 20 #6038 Zbl 0090.09202

[Widder 1941] D. V. Widder, *The Laplace transform*, Princeton Mathematical Series **6**, Princeton University Press, 1941. MR 3,232d Zbl 0063.08245

[Widom 1966] H. Widom, "Hankel matrices", *Trans. Amer. Math. Soc.* **121** (1966), 1–35. MR 32 #4553 Zbl 0148.12303

[Yafaev 2010] D. R. Yafaev, "Коммутаторный метод диагонализации операторов Ганкеля", *Funkts. Anal. Prilozh.* **44**:4 (2010), 65–79. Translated as "A commutator method for the diagonalization of Hankel operators" in *Funct. Anal. Appl.* **44**:4 (2010), 295–306. MR 2011m:47056 Zbl 1271.47019

[Yafaev 2013] D. R. Yafaev, "Spectral and scattering theory for perturbations of the Carleman operator", *Algebra i Analiz* **25**:2 (2013), 251–278. Reprinted in *St. Petersburg Math. J.* **25**:2 (2014), 339–359. MR 3114858 Zbl 06371653

[Yafaev 2014a] D. R. Yafaev, "Diagonalizations of two classes of unbounded Hankel operators", *Bull. Math. Sci.* **4**:2 (2014), 175–198. MR 3228573

[Yafaev 2014b] D. R. Yafaev, "Quasi-diagonalization of Hankel operators", preprint, 2014. To appear in *J. Anal. Math.* arXiv 1403.3941

[Yafaev 2015] D. R. Yafaev, "On finite rank Hankel operators", *J. Funct. Anal.* **268**:7 (2015), 1808–1839. MR 3315579

Received 13 Apr 2014. Accepted 26 Nov 2014.

DIMITRI R. YAFAEV: yafaev@univ-rennes1.fr
*Institut de Recherche Mathématique de Rennes, Université de Rennes I, Campus de Beaulieu, 35042 Rennes Cedex, France*

ANALYSIS AND PDE
Vol. 8, No. 1, 2015
dx.doi.org/10.2140/apde.2015.8.223

# NODAL SETS AND GROWTH EXPONENTS OF LAPLACE EIGENFUNCTIONS ON SURFACES

GUILLAUME ROY-FORTIN

We prove a result, announced by F. Nazarov, L. Polterovich and M. Sodin, that exhibits a relation between the average local growth of a Laplace eigenfunction on a closed surface and the global size of its nodal set. More precisely, we provide a lower and an upper bound to the Hausdorff measure of the nodal set in terms of the expected value of the growth exponent of an eigenfunction on disks of wavelength-like radius. Combined with Yau's conjecture, the result implies that the average local growth of an eigenfunction on such disks is bounded by constants in the semiclassical limit. We also obtain results that link the size of the nodal set to the growth of solutions of planar Schrödinger equations with small potential.

## 1. Introduction and main results

**1.1.** ***Nodal sets of Laplace eigenfunctions.*** Let $(M, g)$ be a smooth, closed two-dimensional Riemannian manifold endowed with a $C^\infty$ metric $g$. Let $\{\phi_\lambda\}$, $\lambda \nearrow \infty$, be any sequence of eigenfunctions of the negative-definite Laplace–Beltrami operator $\Delta_g$:

$$\Delta_g \phi_\lambda + \lambda \phi_\lambda = 0. \tag{1.1.1}$$

In local coordinates, we write the Laplace–Beltrami operator as

$$\Delta_g = \frac{1}{\sqrt{g}} \sum_{i,j=1}^{2} \frac{\partial}{\partial x_i}\left(g^{ij}\sqrt{g}\,\frac{\partial}{\partial x_j}\right).$$

The nodal set of $\phi_\lambda$ is the set

$$Z_\lambda := \{p \in M : \phi_\lambda(p) = 0\}.$$

It is known [Cheng 1976] that $Z_\lambda$ is a smooth curve away from its finite singular set

$$S_\lambda := \{p \in M : \phi_\lambda(p) = \nabla\phi_\lambda(p) = 0\}.$$

Nodal sets of Laplace eigenfunctions have been of interest since the discovery of the Chladni patterns and their asymptotic properties as $\lambda \nearrow \infty$ have been intensively studied, notably in the context of quantum mechanics. In that setting, the square of a normalized eigenfunction $\phi_\lambda$ represents the probability density of a free particle in the pure state corresponding to $\phi_\lambda$ and $Z_\lambda$ can be thought of as the set where such a particle is least likely to be found. Estimating the one-dimensional Hausdorff measure $\mathcal{H}^1(Z_\lambda)$ of

Roy-Fortin has been supported by NSERC.
*MSC2010:* 58J50.
*Keywords:* spectral geometry, Laplace eigenfunctions, nodal sets, growth of eigenfunctions.

the nodal set has thus been the subject of intense studies over the last three decades, sparked by the well-known conjecture of S.-T. Yau [1982; 1993]:

**Conjecture 1.1.2.** *Let $(M, g)$ be a compact, $C^\infty$ Riemannian manifold of dimension $n$. There exist positive constants $c$ and $C$ such that*

$$c\lambda^{1/2} \le \mathcal{H}^{n-1}(Z_\lambda) \le C\lambda^{1/2}.$$

Remark that this paper is concerned with the case $n = 2$ but that the conjecture has been stated for smooth manifolds of any dimension. A common intuition in spectral geometry is that a $\lambda$-eigenfunction behaves in many ways similarly to a trigonometric polynomial of degree $\lambda^{1/2}$. As such, one can understand Yau's conjecture as a broad generalization of the fundamental theorem of algebra: counting multiplicities, a polynomial of degree $\lambda^{1/2}$ will vanish $\lambda^{1/2}$ times. The conjecture has been proved by Donnelly and Fefferman [1988] for real analytic pairs $(M, g)$ of any dimension. When $M$ is a surface with a $C^\infty$ metric, the lower bound was proved by Brüning [1978]. The current best upper bound of $\lambda^{3/4}$ obtained by [Donnelly and Fefferman 1990; Dong 1992] is still weaker than the conjectured one. Note that the current best exponent $\frac{3}{4}$ in dimension 2 gets much worse in higher dimensions. Indeed, for $n \ge 3$, the current best upper bound is $\lambda^{\sqrt{\lambda}}$ and has been obtained by Hardt and Simon [1989]. This hints that the methods used on surfaces are specific and cannot, in general, be easily extended to higher-dimensional manifolds, which is indeed the case for the results of this paper. For more details and a thorough survey of the most recent results on nodal sets of Laplace eigenfunctions, we refer to [Zelditch 2013].

**1.2.** ***An averaged measure of the local growth.*** Here and elsewhere in this article, given a ball $B(r)$ of radius $r$, $\alpha B$ will denote the concentric ball of radius $\alpha r$. In any metric space, it is possible to measure the growth of a continuous function $f$ by defining its *doubling exponent* $\beta(f, B)$ on a metric ball $B$ by

$$\beta(f, B) := \log \frac{\sup_B |f|}{\sup_{\frac{1}{2}B} |f|}.$$

The simplest example is that of the polynomial $x^n$ on the real interval $D = [-1, 1]$, for which the doubling exponent is the degree $n$, modulo a constant. Indeed, $\beta(x^n, [-1, 1]) = n \log 2$. Given two concentric balls $B$ and $\alpha B$, where $0 < \alpha < 1$, one can define the more general *$\alpha$-growth exponent* $\beta(f, B; \alpha)$ by

$$\beta(f, B; \alpha) := \log \frac{\sup_B |f|}{\sup_{\alpha B} |f|}.$$

Albeit more general, the growth exponent can still be seen as the analog of the degree of a polynomial as showcased once again by the monomial $x^n$:

$$\beta(x^n, [-1, 1]; \alpha) = \log \frac{\sup_{[-1,1]} |x|^n}{\sup_{[-\alpha,\alpha]} |x|^n} = n \log(\alpha^{-1}).$$

It is worth mentioning that the growth exponent is itself a special case of the more general Bernstein index, which measures in a similar fashion the growth of a continuous function from one compact set to a strictly larger one. For more background on the Bernstein index, we refer to [Khovanskii and Yakovenko 1996; Roytwarf and Yomdin 1997].

The metric $g$ turns $M$ into a metric space, and it is natural to define similar exponents to measure the growth of eigenfunctions on metric disks on the surface. We write $B_p(r)$ for a metric disk centered at $p \in M$ and of radius $r$. Donnelly and Fefferman [1988] show that on a smooth manifold $(M, g)$ of any dimension, the following holds for every ball $B$:

$$\beta(\phi_\lambda, B) \leq c\lambda^{1/2},$$

where $c = c(g, r, \alpha)$ is a positive constant depending only on the geometry of $M$, the radius $r$ and the scaling factor $\alpha$. From now on, we will restrict our attention to disks $B_p(r)$ of radius comparable to the wavelength: $r = k_0\lambda^{-1/2}$, where $k_0$ is a suitably small, positive constant. It turns out that, at this scale, the local study of an eigenfunction can be reduced to that of a solution of a planar Schrödinger equation (see Section 2.3), which is a central idea throughout this article. For simplicity, we write

$$\beta_p(\lambda) := \beta(\phi_\lambda, B_p(r); \alpha_0)$$

for the $\alpha_0$-growth exponent of $\phi_\lambda$ and where $\alpha_0$ is a geometric constant whose explicit value is given by (2.2.3). The quantity $\beta_p(\lambda)$ is by definition local, and motivated by Section 7.3 in [Nazarov et al. 2005], we make it global by defining the *average local growth* of a $\lambda$-eigenfunction, which is essentially the averaged $L^1$ norm of $\beta_p(\lambda)$:

$$A(\lambda) := \frac{1}{\mathrm{Vol}(M)} \int_M \beta_p(\lambda)\, \mathrm{d}V_g(p).$$

Thus, $A(\lambda)$ can be interpreted as the expected value of the $\alpha_0$-growth exponent of an eigenfunction $\phi_\lambda$ on disks of wavelength radius.

**1.3. *Results.*** We recall the basic intuition of interpreting an eigenfunction $\phi_\lambda$ as a polynomial of degree $\lambda$. In the case of a polynomial, the degree controls both the growth and the number of zeros and it is thus natural to expect a similar link for eigenfunctions. Our main result proves Conjecture 7.1 of [Nazarov et al. 2005] and provides such a link by showing that the average local growth is comparable to the size of the nodal set $Z_\lambda$ times the wavelength $\lambda^{-1/2}$.

**Theorem 1.** *Let $(M, g)$ be a smooth, closed Riemannian manifold of dimension* 2. *There exist positive constants $c_1$ and $c_2$ such that*

$$c_1\lambda^{1/2}A(\lambda) \leq \mathcal{H}^1(Z_\lambda) \leq c_2\lambda^{1/2}(A(\lambda) + 1). \tag{1.3.1}$$

The theorem provides an interesting reformulation of Yau's conjecture for surfaces with smooth metric. Recall that, in this setting, the lower bound of Conjecture 1.1.2 is proven so that, in view of Theorem 1, the conjecture holds if and only if

$$A(\lambda) = O(1).$$

Also, since the conjecture is true in the analytic case, we immediately have that $A(\lambda) = O(1)$ in such a setting. In other words, on a surface with a real analytic metric, the average local growth of an eigenfunction on balls of small radius is bounded by a constant independent of the eigenvalue.

Finally, two other main results are of interest, namely Theorems 2.1.1 and 3.1.1, each providing a link between growth exponents and the size of nodal sets of solutions to a planar Schrödinger equation. The explicit statement of these results is respectively given at the beginning of Sections 2 and 3.

**1.4. *Outline of proof and organization of the paper.*** Nazarov et al. [2005, §7.3] suggested a heuristic for the proof of Theorem 1 that essentially consisted of the following four steps:

(i) Reduce an eigenfunction $\phi_\lambda$ to a solution $F$ of a planar Schrödinger equation. This is done locally on a conformal coordinate patch by restricting $\phi_\lambda$ to a small disk of radius $\sim \lambda^{-1/2}$, which transforms the eigenvalue equation (1.1.1) into

$$\Delta F + qF = 0,$$

where $\Delta$ is the flat Laplacian and $q$ is a smooth potential with small uniform norm.

(ii) Use Lemma 3.4 from [Nazarov et al. 2005] to express $F$ as the composition $u \circ h$ of a harmonic function $u$ with a $K$-quasiconformal homeomorphism $h$ whose dilation factor $K$ is controlled.

(iii) Extend to $F$ and then to $\phi_\lambda$ some appropriate estimates linking the size of the nodal set of $u$ with its growth exponent $\beta$. Such estimates are in the spirit of Lemma 2.13 in [Nazarov et al. 2005] (see also [Gelfond 1934; Robertson 1939; Khovanskii and Yakovenko 1996]) and relate the growth exponents of a harmonic function $u$ on some disk with the number of change of signs of $u$ on the boundary of either a larger or a smaller disk.

(iv) The last step is an integral-geometric argument based on a generalized Crofton formula that allows one to recover the global statement of Theorem 1 from the local estimates obtained in the previous steps.

This approach has been successful in obtaining the lower bound for the size of the nodal set in terms of the average local growth, that is, the left inequality of Theorem 1. The details are presented in Section 3. However, as first noticed by J. Bourgain, the same approach cannot be used for the other inequality. The problem roughly resides in step (iii), where we aim to extend to $F = u \circ h$ a result of the type

$$N_u(\partial D_-) \leq \beta_u(D^+),$$

where $N_u(\partial D_-)$ is the number of zeros of $u$ on a circle $\partial D_-$ that is strictly contained in a bigger disk $D^+$ on which the doubling exponent is computed. It is impossible to do so since we have no way to ensure that the $K$-quasiconformal map $h$ will map the circle $\partial D_-$ to another circle in the domain of $F$. It might in fact map a circle to a nonrectifiable curve, which prevents one from properly counting the zeros of $F$.

Based on a private communication with Nazarov, Polterovich and Sodin, we take a different route to prove the upper bound in Theorem 1, which is inspired by [Donnelly and Fefferman 1988]. More precisely, we keep steps (i) and (iv) but replace the intermediate steps by Theorem 2.1.1, which provides a convenient estimate linking the size of the nodal set of $F$ on a small disk to its growth exponent on a bigger disk. This approach is presented in Section 2 and allows us to recover the remaining inequality of our main theorem. Theorem 2.1.1 thus plays a crucial role, and its proof is presented in Section 4. The general idea is to tile the domain of $F$ into squares of rapid and slow growth and to then notice that: (a) the nodal set in a square of slow growth is small and (b) there cannot be too many squares of rapid growth. The interested reader will also find further explanations detailing the structure of that proof in Section 4.2. Involved in the proof are notably the technical Proposition 4.2.1, which roughly proves statement (b) above, as well as the specialized Carleman estimate of Lemma 5.2.1, whose rather long derivations we respectively present in Sections 5 and 6. We conclude the article with a discussion and a few questions in Section 7.

*Notation.* Throughout the paper, we will denote positive numerical constants in the following fashion. Constants $c_1, c_2, \dots$ will be used in the statements of any result and may depend on the geometry of the manifold $M$ but nothing else. In particular, they are independent of $\lambda$. Within proofs, we will use $a_1, a_2, \dots$ for numerical constants without any dependency and $b_1, b_2, \dots$ for constants that may depend on the geometry of the surface. Often, we merge many numerical constants together to simplify the sometimes heavy notation; for example, $a_5 = a_3^{-1} a_4 (4\pi)/\operatorname{Vol}(M)$. Finally, we reset the numeration for the constants $a_i$ at each section.

We will use $D$ to denote Euclidean disks and $B$ for metric balls on the surface. Given the context, we either write $D(p, r)$ for a disk centered at $p$ of radius $r$ or just $D_p$ if the radius is known. Finally, we will keep the convention that, given a positive constant $a$ and a disk $D = D(p, r)$, $aD$ denotes the concentric disk of radius $ar$. We write $\mathbb{D}$ for the open unit disk in $\mathbb{R}^2$.

## 2. Upper bound for the length of the nodal set

In this section, we prove the right inequality of Theorem 1, which provides an upper bound to the length of the nodal set in terms of the average local growth of an eigenfunction $\phi_\lambda$. The main tool in the proof is the following, which links the size of the nodal set of a Schrödinger eigenfunction to its growth exponent:

**Theorem 2.1.1.** *Let $F : 3\mathbb{D} \to \mathbb{R}$ be a solution of*

$$\Delta F + qF = 0 \tag{2.1.2}$$

*with the potential $q \in C^\infty(3\mathbb{D})$ satisfying $\|q\|_\infty = \sup_{3\mathbb{D}} |q| < \epsilon_0$. Let also*

$$\beta := \beta\big(F, \tfrac{5}{2}\mathbb{D}; 10\big) = \log \frac{\sup_{\frac{5}{2}\mathbb{D}} |F|}{\sup_{\frac{1}{4}\mathbb{D}} |F|}.$$

*Finally, denote by $Z_F$ the nodal set $\{p \in 3\mathbb{D} : F(p) = 0\}$ of $F$. Then*

$$\mathcal{H}^1\big(Z_F \cap \tfrac{1}{60}\mathbb{D}\big) \le c_3 \beta^*,$$

*where $\beta^* := \max\{\beta, 1\}$ and $c_3$ is a positive constant.*

We remark that we do not assume here that $q$ has a constant sign. The proof of this theorem is presented in Section 5, and some information about the value of $\epsilon_0$ is given at the end of Lemma 5.4.6.

**2.2.** ***From the surface to the plane: the passage to Schrödinger eigenfunctions with small potential.*** Cover the surface $M$ with a finite number $N$ of conformal charts $(U_i, \psi_i)$, $\psi_i : U_i \subset M \to V_i \subset \mathbb{R}^2$, $i \in I = \{1, \dots, N\}$. On each of these charts, the metric is conformally flat and there exist smooth positive functions $q_i$ such that $g = q_i(x, y)(dx^2 + dy^2)$. By compactness, we can find positive constants $q_-$ and $q^+$ such that we have $0 < q_- < q_i < q^+$ for all $i = 1, \dots, N$. The metric is thus pinched between scalings of the flat metric, and we have a local equivalence of various metric notions on $M$ and in $\mathbb{R}^2$. In particular, given any subset $E \subset U_i$, the one-dimensional Hausdorff measures are equivalent:

$$b_1 \mathcal{H}^1(\psi_i(E)) \le \mathcal{H}^1(E) \le b_2 \mathcal{H}^1(\psi_i(E)). \tag{2.2.1}$$

In the same spirit, the Riemannian volume form on $M$ and the Lebesgue measure $\mathrm{d}A$ in $\mathbb{R}^2$ are equivalent in the following sense: given any integrable function $f$ on $U_i$, we have

$$b_3 \int_{V_i} f \,\mathrm{d}A \le \int_{U_i} f \,\mathrm{d}V_g \le b_4 \int_{V_i} f \,\mathrm{d}A. \tag{2.2.2}$$

Note that the explicit values of the constants $b_1, \dots, b_4$ involve only the geometric constants $q_-$ and $q^+$. We now let $B_p := B_p(k_0 \lambda^{-1/2}) \subset M$ be a metric disk and set

$$\alpha_0 := \frac{q_-}{5q^+}. \tag{2.2.3}$$

The value of the small positive constant $k_0$ will be fixed later. Recall that, at a point $p \in M$, the growth exponent $\beta_p(\lambda)$ of an eigenfunction $\phi_\lambda$ is defined by

$$\beta_p(\lambda) := \log \frac{\sup_{B_p} |\phi_\lambda|}{\sup_{\alpha_0 B_p} |\phi_\lambda|}.$$

**2.3.** ***Metric and Euclidean disks.*** In order to estimate $\beta_p(\lambda)$ from below, we define the Euclidean disks

$$D_p^+ := D_p(q_- k_0 \lambda^{-1/2}), \qquad D_p^- := \alpha_0 D_p(q^+ k_0 \lambda^{-1/2})$$

so that $D_p^-$ is a proper subset of $D_p^+$. Note that, by a Euclidean disk $D_p(r)$ centered at $p \in M$, we mean the set $\{(x, y) : x^2 + y^2 \le r^2\}$, where $(x, y)$ are local conformal coordinates around $p$. The inclusions $B_p \supset D_p^+$ and $\alpha_0 B_p \subset D_p^-$ imply

$$\log \frac{\sup_{D_p^+} |\phi_\lambda|}{\sup_{D_p^-} |\phi_\lambda|} \le \beta_p(\lambda).$$

In a conformal chart $(U_i, \psi_i)$, the eigenvalue equation $\Delta_g \phi_\lambda + \lambda \phi_\lambda = 0$ becomes

$$\Delta \phi_\lambda + \lambda q_i \phi_\lambda = 0. \tag{2.3.1}$$

With the aim of using Theorem 2.1.1, we endow the disk $3\mathbb{D}$ with the complex coordinate $z = x + iy$, fix a scaling constant $\tau = 2q^+ \alpha_0$ and define a function $F = F_{\lambda,p} : 3\mathbb{D} \to \mathbb{R}$ by $F(z) = \phi_\lambda(\tau k_0 \lambda^{-1/2} z + p)$. The scaling allows us to absorb the spectral parameter $\lambda$ in the potential. Indeed, we have

$$\Delta F = \tau^2 k_0^2 \lambda^{-1} \Delta \phi = (k_0 \tau)^2 \lambda^{-1} (-\lambda q_i) \phi_\lambda = -(k_0 \tau)^2 q_i F$$

so that $F$ satisfies (2.1.2), where $q = (k_0 \tau)^2 q_i$ is a smooth potential whose supremum norm satisfies $\|q\|_\infty < \epsilon_0$ without loss of generality. Indeed, since the family of $q_i$ is bounded, we can choose $k_0$ as

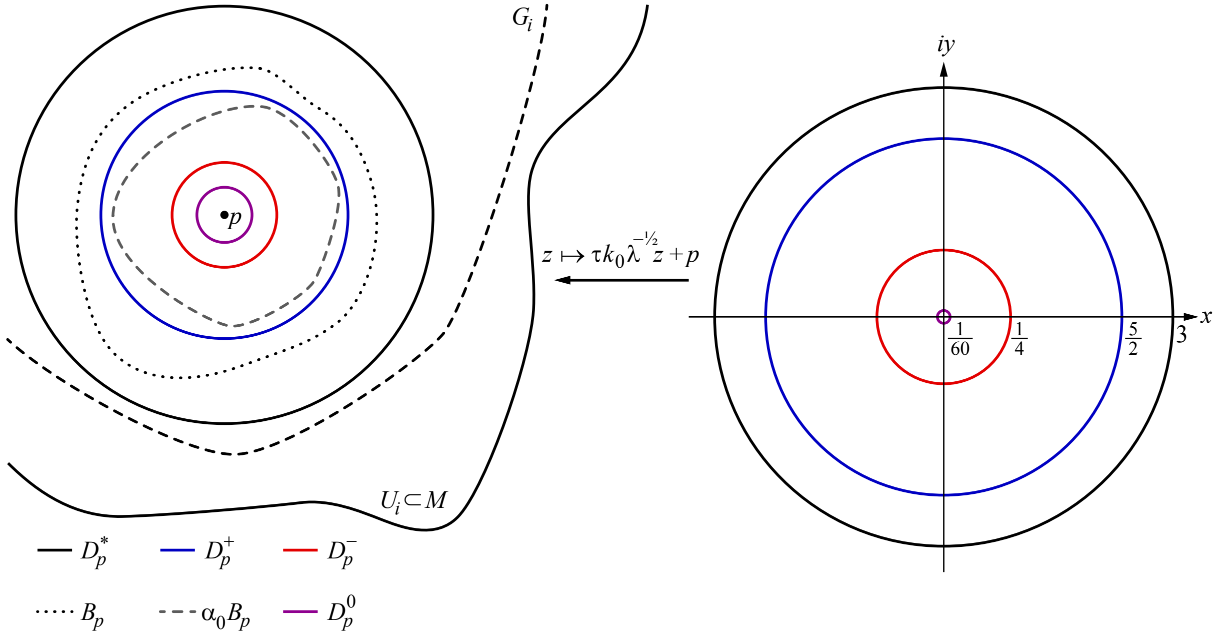


**Figure 1.** Mapping of Euclidean disks and metric balls within a conformal patch.

small as needed. The transformation $z \mapsto \tau k_0 \lambda^{-1/2} z + p$ induces the following correspondences between disks in $3\mathbb{D}$ and Euclidean disks centered at $p$:

$$\left\{|z| \leq \tfrac{1}{4}\right\} \leftrightarrow D_p^-, \qquad \left\{|z| \leq \tfrac{5}{2}\right\} \leftrightarrow D_p^+, \qquad \left\{|z| \leq \tfrac{1}{60}\right\} \leftrightarrow D_p^0,$$

where $D_p^0 = D_p\left(\frac{1}{60}\tau k_0 \lambda^{-1/2}\right)$. As a consequence, we have

$$\beta^* < \beta + 1 = \log \frac{\sup_{D_p^+} |\phi_\lambda|}{\sup_{D_p^-} |\phi_\lambda|} + 1 \leq \beta_p(\lambda) + 1. \tag{2.3.2}$$

It is important at this stage to remark that the construction of $F$ is dependent on a fixed choice of conformal chart $U_i$ both for the well-posedness of (2.3.1) as well as the very definition of the Euclidean disks. Thus, in order to allow the construction of $F = F_{\lambda,p}$ everywhere on the surface $M$, one has to choose $k_0$ small enough so that the disks $D_p^* := D_p(3k_0\tau\lambda^{-1/2})$, which are mapped onto $3\mathbb{D}$, are contained in at least one chart $U_i$ for every $p \in M$. This allows the definition of the mapping $\sigma : M \to I = \{1, \dots, N\}$, which assigns to a point $p$ a unique index $\sigma(p)$ such that $D_p^* \subset U_{\sigma(p)}$. The disjoint sets $G_i := \sigma^{-1}(i)$ form a partition of $M$. Figure 1 summarizes the setting we are in by presenting a sketch of the various correspondences between Euclidean disks in $G_i$ and those in $3\mathbb{D}$.

We now turn to the study of the nodal set $Z_\lambda$. Recall that $S_\lambda$ is the singular set of the eigenfunction $\phi_\lambda$, and consider the sets $Z_\lambda(i) := \psi_i((Z_\lambda \setminus S_\lambda) \cap G_i) \subset \mathbb{R}^2$. Since $S_\lambda$ is discrete, we have

$$\mathcal{H}^1(Z_\lambda) = \mathcal{H}^1(Z_\lambda \setminus S_\lambda) \leq b_2 \sum_{i \in I} \mathcal{H}^1(Z_\lambda(i)). \tag{2.3.3}$$

Denote by $Z_F$ the nodal set of $F$. By construction, we have

$$\mathcal{H}^1(Z_\lambda(i)\cap D_p^0)=(k_0\tau)\lambda^{-1/2}\mathcal{H}^1\big(Z_F\cap \tfrac{1}{60}\mathbb{D}\big).$$

Applying Theorem 2.1.1 and (2.3.2) now yields

$$\mathcal{H}^1(Z_\lambda(i)\cap D_p^0)\le a_2\lambda^{-1/2}(\beta_p(\lambda)+1). \tag{2.3.4}$$

We integrate the left-hand side of the last equation over the set $G_i$ and use a generalized Crofton formula (see (6) in [Hug and Schneider 2002]) to get

$$\int_{G_i}\mathcal{H}^1(Z_\lambda(i)\cap D_p^0)\,\mathrm{d}A(p)=a_3\mathcal{H}^2(D_p^0)\mathcal{H}^1(Z_\lambda(i))=a_4\lambda^{-1}\mathcal{H}^1(Z_\lambda(i)). \tag{2.3.5}$$

Recalling the equivalence (2.2.2) and combining (2.3.4) and (2.3.5) then gives

$$a_4\lambda^{-1}\mathcal{H}^1(Z_\lambda(i))\le a_2\lambda^{-1/2}\int_{G_i}(\beta_p(\lambda)+1)\,\mathrm{d}A(p)\le(a_2b_3^{-1})\lambda^{-1/2}\int_{G_i}(\beta_p(\lambda)+1)\,\mathrm{d}V.$$

Simplifying readily gives

$$\mathcal{H}^1(Z_\lambda(i))\le a_5\lambda^{1/2}\int_{G_i}(\beta_p(\lambda)+1)\,\mathrm{d}V$$

so that

$$\begin{aligned}\mathcal{H}^1(Z_\lambda)\le b_2\sum_{i\in I}\mathcal{H}^1(Z_\lambda(i))\le a_6\lambda^{1/2}\sum_{i\in I}\int_{G_i}(\beta_p(\lambda)+1)\,\mathrm{d}V=a_6\lambda^{1/2}\int_M(\beta_p(\lambda)+1)\,\mathrm{d}V\\ \le c_2\lambda^{1/2}(A(\lambda)+1).\end{aligned}$$

## 3. Lower bound for the length of the nodal set

In this section, we prove the left inequality of Theorem 1. As was the case in the previous section, the central idea is once again the use of conformal coordinates on $M$ and restriction to wavelength scales to reduce the local behavior of an eigenfunction $\phi_\lambda$ to that of $F$, a solution of a planar Schrödinger equation with small, smooth potential. The main result of this section is the following theorem, which suitably links the growth exponent of $F$ with its nodal set:

**Theorem 3.1.1.** *Let $F:\overline{\mathbb{D}}\to\mathbb{R}$ be a solution of*

$$\Delta F+qF=0 \tag{3.1.2}$$

*in $\mathbb{D}$ and with the potential $q\in C^\infty(\mathbb{D})$ satisfying $\|q\|_\infty=\sup_{\overline{\mathbb{D}}}|q|<\epsilon_1$. Denote by $|Z_F(\mathbb{S}^1)|$ the number of zeros of $F$ on the unit circle $\mathbb{S}^1$. Then*

$$\log\frac{\sup_{\rho^+\mathbb{D}}|F|}{\sup_{\rho^-\mathbb{D}}|F|}\le c_4(1+|Z_F(\mathbb{S}^1)|),$$

*where $0<\rho^-<\rho^+<\frac{1}{2}$ are fixed, small radii.*

The value of $\epsilon_1$ can be obtained in the proof of Lemma 3.3 in [Nazarov et al. 2005] while those of $\rho^-$ and $\rho^+$ are given in the proof. The constant $\rho^-$ depends on the geometry of the manifold. It is possible to get rid of this dependency if one wants Theorem 3.1.1 to be a stand-alone result. However, our aim is to prove the left inequality of Theorem 1, and as such, our choice of $p^-$ makes the rest of the argument much simpler. Also, remark that, in contrast to Theorem 3.1.1 where $F$ was defined on $\mathbb{D}$, the setting is now in $3\mathbb{D}$. This is an arbitrary choice made only in order to ease the writing of the respective proofs: confining Theorem 2.1.1 to the unit disk would have added even more complexity in the expression of the many constants needed to carry out the long proof.

**3.2.** ***Proof of Theorem 3.1.1.*** The general strategy is as follows: we first prove a similar kind of result for harmonic functions, and inspired by [Nazarov et al. 2005], we then express $F$ as the composition of a harmonic function and a $K$-quasiconformal homeomorphism. Controlling the properties of the quasiconformal homeomorphism allows one to recover the desired result. We begin with a lemma that relates the growth of harmonic functions within a disk and its nodal set on the boundary.

**Lemma 3.2.1.** *Let $v \in C^\infty(\mathbb{D}) \cap C^0(\overline{\mathbb{D}})$ be harmonic in the open unit disk, and denote by $N_v$ the number of changes of sign of $v$ on the circle $|z| = 1$. Choose $r_0$ in $0 < r_0 < \frac{1}{2}$. Then*

$$\frac{\sup_{\frac{1}{2}\mathbb{D}}|v|}{\sup_{r_0\mathbb{D}}|v|} \le \left(\frac{c_5}{r_0}\right)^{N_v}, \tag{3.2.2}$$

*where $c_5$ is a positive numerical constant.*

*Proof.* Let $u$ be the harmonic conjugate of $v$ such that $u(0) = 0$. Then the function

$$f(z) = \sum_{n=0}^{\infty} \xi_n z^n = u(z) + i v(z)$$

is holomorphic in the closed unit disk $\{|z| \le 1\}$. Suppose that

$$\sup_{r_0\mathbb{D}}|v| = \max_{|z|=r_0}|v| = 1.$$

The harmonic function $v$ changes sign $2p = N_v$ times on the circle $|z| = 1$, where $p$ is a nonnegative integer. Also, let $\mu_p := \max\{|\xi_0|, |\xi_1|, \dots, |\xi_p|\}$. By [Robertson 1939, Theorem 1, (iii)], we have

$$|f(re^{i\theta})| < c(p)\mu_p(1-r)^{-2p-1}, \quad r < 1, \tag{3.2.3}$$

where $c(p) > 0$ is a constant depending on $p$ that will be given explicitly later. (Robertson actually proves (3.2.3) in our current setting and then uses a limiting argument to obtain a slightly different statement.)

The classical Schwarz formula says that, for a function $g$ holomorphic on the open disk $r_0\mathbb{D}$ and continuous on the boundary $\{|z| = r_0\}$, we have

$$g(z) = \frac{1}{2\pi}\int_0^{2\pi} \operatorname{Re}(g(r_0e^{i\theta}))\frac{r_0e^{i\theta}+z}{r_0e^{i\theta}-z}\,\mathrm{d}\theta + i\ \operatorname{Im}(g(0)), \quad |z| < r_0.$$

Since $f = u + iv$ is holomorphic, so is $g = v - iu$ and we obviously have $|f| = |g|$ so that the following inequality holds for all $|z| \le \frac{1}{2} r_0$:

$$|f(z)| = |g(z)| = \left| \frac{1}{2\pi} \int_0^{2\pi} \operatorname{Re}(g(r_0 e^{i\theta})) \frac{r_0 e^{i\theta} + z}{r_0 e^{i\theta} - z} \, \mathrm{d}\theta + iu(0) \right|$$
$$\le \frac{1}{2\pi} \max_{|z|=r_0} |v| \int_0^{2\pi} \frac{r_0 + |z|}{|r_0 - z|} \, \mathrm{d}\theta \le 3 =: a_1.$$

Applying Cauchy's inequality for holomorphic functions to $f = \sum_{n=0}^{\infty} \xi_n z^n$ on the open disk of radius $\frac{1}{2} r_0$,

$$|\xi_n| \le \left( \frac{r_0}{2} \right)^{-n} \sup_{|z|=r_0/2} |f(z)| = a_1 \left( \frac{2}{r_0} \right)^n.$$

Hence, we have $\mu_p \le a_1 (2/r_0)^p$. Setting $r = \frac{1}{2}$ in (3.2.3) now yields

$$\left| f\left(\tfrac{1}{2} e^{i\theta}\right) \right| \le c(p) \mu_p 2^{2p+1} \le 2a_1 c(p) \left( \frac{2}{r_0} \right)^p 2^{2p} \le 2a_1 c(p) \left( \frac{4}{r_0} \right)^{2p},$$

which in turn means

$$\sup_{\frac{1}{2}\mathbb{D}} |v| = \max_{|z|=1/2} |v| \le \left| f\left(\tfrac{1}{2} e^{i\theta}\right) \right| \le 2a_1 c(p) \left( \frac{4}{r_0} \right)^{2p}.$$

Going back to [Robertson 1939], we use the explicit value of the constant $c(p)$ to get the bound

$$c(p) = 2^{2p} + \frac{(2p)!}{(p!)^2} = 2^{2p} + \binom{2p}{p} \le 2^{2p} + \left( \frac{2pe}{p} \right)^p = 2^{2p} + (2e)^p \le 2(2e)^{2p}.$$

Since we assumed that $\sup_{r_0 \mathbb{D}} |v| = 1$, we have

$$\frac{\sup_{\frac{1}{2}\mathbb{D}} |v|}{\sup_{r_0 \mathbb{D}} |v|} \le 4a_1 \left( \frac{8e}{r_0} \right)^{2p}.$$

Suppose now that $\sup_{r_0 \mathbb{D}} |v| = \tau \ne 1$, and let as before $f = u + iv$ be the holomorphic function built from $v$ and its harmonic conjugate $u$. Define $\tilde{f} = \tilde{u} + i\tilde{v}$ by $\tilde{f} = \tau^{-1} f$. Then $\sup_{r_0 \mathbb{D}} |\tilde{v}| = 1$ and

$$\frac{\sup_{\frac{1}{2}\mathbb{D}} |v|}{\sup_{r_0 \mathbb{D}} |v|} = \frac{\tau \sup_{\frac{1}{2}\mathbb{D}} |\tilde{v}|}{\tau \sup_{r_0 \mathbb{D}} |\tilde{v}|} \le 4a_1 \left( \frac{8e}{r_0} \right)^{2p} \le \left( \frac{c_5}{r_0} \right)^{2p}. \qquad \square$$

We now prove Theorem 3.1.1. By Lemmas 3.3 and 3.4 in [Nazarov et al. 2005], there exist a $K$-quasiconformal homeomorphism $h : \mathbb{D} \to \mathbb{D}$ with $h(0) = 0$, a harmonic function $v : \mathbb{D} \to \mathbb{R}$ and a solution $\varphi$ to (3.1.2) such that $F = \varphi \cdot (v \circ h)$. Moreover, the function $\varphi$ is positive and satisfies

$$1 - a_2 \epsilon_1 \le \varphi \le 1.$$

Finally, the dilation factor of the quasiconformal map $h$ satisfies

$$1 \le K \le 1 + a_3 \|q\|_\infty \le a_4.$$

We refer the reader to [Nazarov et al. 2005] for the precise values of the various constants stated above. We recall Mori's theorem (see Section III.C in [Ahlfors 1966] or [Nazarov et al. 2005]) for $K$-quasiconformal homeomorphisms:

$$\tfrac{1}{16}|z_1 - z_2|^K \le |h(z_1) - h(z_2)| \le 16|z_1 - z_2|^{1/K}.$$

Since the origin is a fixed point of $h$, we have

$$\tfrac{1}{16}|z|^K \le |h(z)| \le 16|z|^{1/K}, \quad z \in \mathbb{D}.$$

Fix a small radius $\rho^+ = \left(\frac{1}{32}\right)^{a_4}$, and consider the circle $\{|z| = \rho^+\}$. For such $z$, Mori's theorem gives $|h(z)| \le 16(\rho^+)^{1/K} \le \frac{1}{2}$ so that

$$h(\rho^+\mathbb{D}) \subset \tfrac{1}{2}\mathbb{D}.$$

Now, set $\rho^- := \frac{1}{5}\rho^+(q_-/q^+)^2$. The image by $h$ of the circle $\{|z| = \rho^-\}$ contains the circle of radius $\frac{1}{16}(\rho^-)^K \ge \frac{1}{16}(\rho^-)^{a_4} =: r_0$. As a consequence, we have

$$r_0\mathbb{D} \subset h((\rho^-)\mathbb{D}).$$

Since $F = \varphi \cdot (v \circ h)$, the bounds on $\varphi$ and the above inclusions imply

$$\frac{\sup_{\rho^+\mathbb{D}}|F|}{\sup_{\rho^-\mathbb{D}}|F|} \le a_5 \frac{\sup_{\rho^+\mathbb{D}}|v \circ h|}{\sup_{\rho^-\mathbb{D}}|v \circ h|} \le a_5 \frac{\sup_{\frac{1}{2}\mathbb{D}}|v|}{\sup_{r_0\mathbb{D}}|v|},$$

where $a_5 = (1 - a_2\epsilon_1)^{-1}$. Since $\varphi$ is positive and $h$ is a homeomorphism, the number $N_F$ of sign changes of $F$ on the unit circle is the same as that of $v$. Applying Lemma 3.2.1 now yields

$$\frac{\sup_{\rho^+\mathbb{D}}|F|}{\sup_{\rho^-\mathbb{D}}|F|} \le a_5 \left(\frac{c_5}{r_0}\right)^{N_F}.$$

Since the number $|Z_F(\mathbb{S}^1)|$ of zeros of $F$ on the unit circle is bounded below by $N_F$, taking the logarithm on both sides yields

$$\log \frac{\sup_{\rho^+\mathbb{D}}|F|}{\sup_{\rho^-\mathbb{D}}|F|} \le c_4(1 + |Z_F(\mathbb{S}^1)|),$$

where $c_4 = \max\{a_5, c_5/r_0\}$.

**3.3.** ***A lower bound for the nodal set in terms of the average local growth.*** In order to recover the right inequality of Theorem 1, we propose an argument that is very similar to the one developed in Section 2. It thus helps to refer to that section when reading the remainder of this one. The aim is to apply Theorem 3.1.1 to a function $F$ that has been built from an eigenfunction $\phi_\lambda$ and to then apply an integral-geometric argument to recover the desired result. We begin with the same setting as that of Section 2.2 and then define the Euclidean disks

$$D_p^+ := D_p(q^+ k_0 \lambda^{-1/2}), \qquad D_p^- := \alpha_0 D_p(q^- k_0 \lambda^{-1/2}).$$

The last two definitions employ the same notation as in the previous section, but the radii of the disks are different. The inclusions $B_p \subset D_p^+$ and $\alpha B_p \supset D_p^-$ imply

$$\beta_p(\lambda) \le \log \frac{\sup_{D_p^+}|\phi_\lambda|}{\sup_{D_p^-}|\phi_\lambda|}. \tag{3.3.1}$$

Let $\tau := q^+/\rho^+$ be a scaling constant, endow the unit disk with the complex coordinate $z = x + iy$ and define $F_{\lambda,p} = F : \mathbb{D} \to \mathbb{R}$ by $F(z) = F(\tau k_0 \lambda^{-1/2} z + p)$. The function $F$ solves (3.1.2), and the potential $q$ satisfies $\|q\|_\infty < \min\{\epsilon_0, \epsilon_1\}$ without loss of generality, choosing $k_0$ small enough. Recalling that $\rho^- = \frac{1}{5}\rho^+(q^-/q^+)^2$, we remark that the mapping $z \mapsto \tau k_0 \lambda^{-1/2} z + p$ induces the bijections

$$\{|z| \le \rho^+\} \leftrightarrow D_p^+, \qquad \{|z| \le \rho^-\} \leftrightarrow D_p^-.$$

An immediate consequence is

$$\log \frac{\sup_{\rho^+\mathbb{D}}|F|}{\sup_{\rho^-\mathbb{D}}|F|} = \log \frac{\sup_{D_p^+}|\phi_\lambda|}{\sup_{D_p^-}|\phi_\lambda|} \ge \beta_p(\lambda). \tag{3.3.2}$$

Notice that, for $F$ to be properly defined on $\mathbb{D}$, the Euclidean disk $D_p^0 := D_p(\tau k_0 \lambda^{-1/2})$ must lie completely within some conformal chart $U_i$. Hence, to ensure that the above construction can be carried through for any $p \in M$, we choose $k_0$ small enough that $D_p(\tau k_0 \lambda^{-1/2})$ is a proper subset of at least one conformal chart $U_i$ for every $p \in M$. This allows one to define the map $\sigma : M \to I = \{1, \dots, N\}$ that assigns to $p \in M$ a unique index $\sigma(p)$ such that $D_p(\tau k_0 \lambda^{-1/2}) \subset U_{\sigma(p)}$. Once again, the sets $G_i := \sigma^{-1}(i) \subset U_i$ form a partition of $M$. Now consider the sets $Z_\lambda(i) := \psi_i((Z_\lambda \setminus S_\lambda) \cap G_i)$, $i = 1, \dots, N$. Then

$$\mathcal{H}^1(Z_\lambda) = \mathcal{H}^1(Z_\lambda \setminus S_\lambda) \ge b_1 \sum_{i \in I} \mathcal{H}^1(Z_\lambda(i)). \tag{3.3.3}$$

Denote by $|Z_{p,\lambda}(i)|$ the number of intersection points of the circle $\partial D_p^0$ with $Z_\lambda(i)$. By construction, the following equality holds outside from the singular set, that is, almost everywhere:

$$|Z_{p,\lambda}(i)| = |Z_F(\mathbb{S}^1)|. \tag{3.3.4}$$

Applying Theorem 3.1.1 and (3.3.2) now yields

$$\beta_p(\lambda) \le c_4(1 + |Z_{p,\lambda}(i)|) \tag{3.3.5}$$

outside of $S_\lambda$. We integrate the left-hand side of the last equation over the set $G_i$ and use a generalized Crofton formula [Hug and Schneider 2002, (6)] to get

$$\int_{G_i \setminus S_\lambda} |Z_\lambda(\partial D_p^0)| \,\mathrm{d}A(p) = a_2 \mathcal{H}^1(\partial D_p^0) \mathcal{H}^1(Z_\lambda(i)) = a_3 \lambda^{-1/2} \mathcal{H}^1(Z_\lambda(i)). \tag{3.3.6}$$

Notice that, in contrast to the previous use of an analogous Crofton formula in Section 2, we have now integrated, over all planar rigid motions, the cardinality of the intersection of a one-dimensional rotation-invariant submanifold — namely the circle $\partial D_p^0$ — with the one-dimensional nodal set.

It is now straightforward to conclude

$$\begin{aligned}
A(\lambda) &= \frac{1}{\mathrm{Vol}(M)} \sum_{i\in I} \int_{G_i\setminus S_\lambda} \beta_p(\lambda)\,\mathrm{d}V_g \\
&\le (b_4 c_4)\left(1+\frac{1}{\mathrm{Vol}(M)} \sum_{i\in I} \int_{G_i\setminus S_\lambda} |Z_{p,\lambda}(i)|\,\mathrm{d}A(p)\right) \\
&= a_5\left(1+\frac{a_3}{\mathrm{Vol}(M)}\lambda^{-1/2} \sum_{i\in I} \mathcal{H}^1(Z_\lambda(i))\right) \\
&\le a_6(1+\lambda^{-1/2}\mathcal{H}^1(Z_\lambda)) \\
&\le c_1 \mathcal{H}^1(Z_\lambda)\lambda^{-1/2},
\end{aligned}$$

where the last inequality uses the fact that the lower bound in Yau's conjecture holds for surfaces, preventing $\lambda^{-1/2}\mathcal{H}^1(Z_\lambda)$ to be too small.

## 4. Nodal set and growth of planar Schrödinger eigenfunctions with small potential

This section is dedicated to the proof of Theorem 2.1.1. We start with a function $F : 3\mathbb{D} \to \mathbb{R}$ that satisfies the equation $\Delta F + qF = 0$ on $3\mathbb{D}$. The potential $q$ is smooth and has a small uniform norm: $\|q\|_\infty < \epsilon_0$. Recall that

$$\beta = \log \frac{\sup_{\frac{5}{2}\mathbb{D}} |F|}{\sup_{\frac{1}{4}\mathbb{D}} |F|}$$

and that $\beta^* = \max\{\beta, 1\}$.

**4.1.** ***A configuration of disks and annuli.*** We start with some notation for disks and annuli within our main setting, which takes place in the disk $3\mathbb{D}$. We denote a finite set of *small disks* by

$$D_\nu = D(z_\nu, \delta) \subset \tfrac{1}{60}\mathbb{D}, \quad 1 \le \nu \le N,$$

where the radius $\delta > 0$ is suitably small. We will say that such a set of small disks is $\gamma$-*separated* if it satisfies $|z_\mu - z_\nu| \ge 2\gamma\delta$ for all $\mu \ne \nu$, where $\gamma$ is some positive constant. One has to understand the $\gamma$-separation condition as disjointness after a scaling of factor $\gamma$. For example, in Figure 2, the disks $D_1$ and $D_2$ are $\gamma$-separated while the pair $D_\nu$ and $D_N$ is not.

For a small $0 < a \ll 1$, we now let $D_\nu(a) := (1-2a)D_\nu$ and define the following annuli:

- $A_\nu = \{(1-2a)\delta < |z - z_\nu| < (1-a)\delta\}$,
- $A_{\nu'} = \{(1-3a)\delta < |z - z_\nu| < (1-4/(3a))\delta\}$,
- $A_{\nu''} = \left\{\left(1-\tfrac{3}{2}a\right)\delta < |z - z_\nu| < (1-a)\delta\right\}$.

We regroup the collection of annuli $A_\nu$ under $A = \bigcup_\nu A_\nu$. Figure 3 provides a close-up of the various annuli defined above.

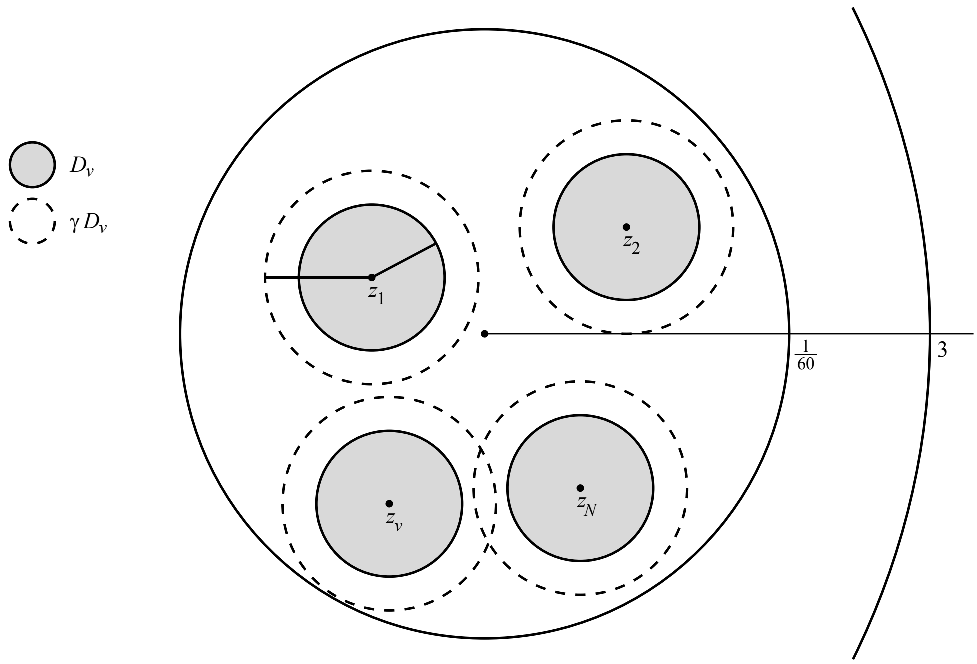


**Figure 2.** A finite set of disks $D_\nu$ and scaled disks within $\frac{1}{60}\mathbb{D}$.

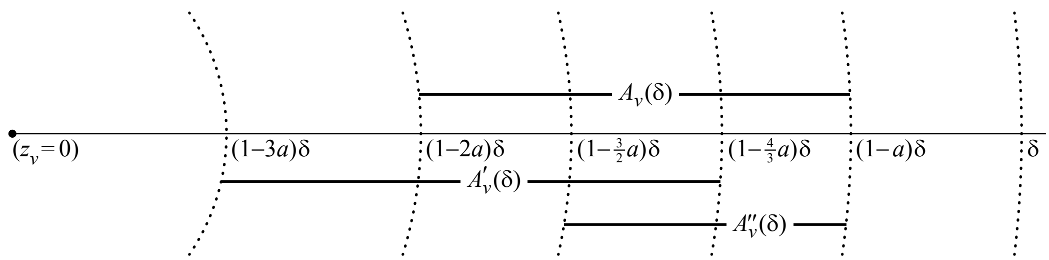


**Figure 3.** Various annuli within a disk $D_\nu$ of radius $\delta$ centered in $z_\nu$.

Given $M > 0$, we say that a disk $D_\nu$ is a *disk of M-rapid growth* or simply a *rapid disk* if

$$M \int_{A_{\nu'}} F^2 \leq \int_{A_{\nu''}} F^2. \tag{4.1.1}$$

We say the radius $\delta$ is $\beta^*$-related if it satisfies

$$\delta < \frac{1}{60}, \qquad \delta\beta^* < \frac{1}{2}. \tag{4.1.2}$$

Finally, we fix the separation constant to $\gamma := \delta^{-1/2}$.

**4.2.** ***Intermediate results.*** We first state a result that shows that, if the potential is small enough and if we fix the growth threshold $M$ sufficiently high, there cannot be too many disks of rapid growth. In fact, it turns out that the number of such disks is bounded above by a constant times the growth exponent $\beta^*$.

**Proposition 4.2.1.** *Suppose that the radius of a collection of $\gamma$-separated small disks in $\frac{1}{60}\mathbb{D}$ satisfies the constraints* (4.1.2), *and let $\mathcal{N} = \mathcal{N}(M)$ denote the number of such disks that are of $M$-rapid growth. Then*

$$\mathcal{N} \leq c_5 \beta^*,$$

*provided that $\|q\|_\infty < \epsilon_0$ and $M > M_0$, where $c_5$, $\epsilon_0$ and $M_0$ are positive constants.*

The rather long proof, inspired by that of Proposition 4.7 in [Donnelly and Fefferman 1990], is presented in Section 5. The next result is Proposition 5.14 in the same reference and links the growth condition and the local length of the nodal set.

**Proposition 4.2.2.** *Suppose that the disk of radius $\epsilon$ centered at $z_\mu$ is not $M_0$-rapid, that is,*

$$\int_{(1-\frac{3}{2}a)\epsilon < |z - z_\mu| < (1-a)\epsilon} F^2 \leq M^{-1} \int_{(1-3a)\epsilon < |z - z_\mu| < (1-\frac{4}{3}a)\epsilon} F^2$$

*holds. Then*

$$\mathcal{H}^1(Z_F \cap D(z_\mu, c_6\epsilon)) \leq c_7 \epsilon,$$

*where $c_6, c_7 > 0$ are positive constants.*

The last two propositions allow us to lay out a general strategy to prove Theorem 2.1.1. Indeed, we now know that there cannot be too many disks of rapid growth and the nodal set of a slow disk cannot be too big. Conjugating those two ideas in the right way will allow us to bound the global length of the nodal set by the growth exponent of $F$.

The proof is based on an iterative process that will be indexed by $k = 0, 1, 2, \ldots$. We begin the first step $k = 0$ by fixing some $\delta(0)$ satisfying the constraints (4.1.2) and then divide the square $P = \{(x, y) : |x|, |y| < \frac{1}{60}\}$ into a grid of squares whose sides have length $\delta(0)$. We distribute those smaller squares into two categories. The *rapid squares* $R_i(0)$, $i = 1, 2, \ldots, r(0)$, are those that contain at least one point $z_i(0) \in R_i(0)$ such that $D_i = D(z_i, \delta)$ is a disk of $M$-rapid growth of the function $F$. Here we have fixed $M = M_0$ to allow the use of Proposition 4.2.1. If that condition is not satisfied, we consider the square to be a *slow square* and label it $S_j(0)$, $j = 1, 2, \ldots, s(0)$.

We now proceed to the next step $k = 1$ and set $\delta(1) = \frac{1}{2}\delta(0)$. We bisect the rapid squares $R_i(0)$ of the previous step into four smaller squares and split those newly obtained squares into rapid squares $R_i(1)$, $i = 1, 2, \ldots, r(1)$, and slow squares $S_j(1)$, $j = 1, 2, \ldots, s(1)$, depending on whether they include a point that is the center of an $M$-rapid disk of radius $\delta(1)$. Note that the slow squares of the previous step are left untouched. Figure 4 gives a representation of the tiling process.

We repeat the process so that, at step $k$, we have $\delta(k) = 2^{-k}\delta(0)$ as well as some rapid squares $R_i(k)$ and slow squares $S_j(k)$. Let $I(k) = \{1, 2, \ldots, r(k)\}$ be the indexing set of the rapid squares obtained at step $k$. To simplify notation, we will sometimes write $\delta$ instead of $\delta(0)$ in what follows and until the end of the section.

**Lemma 4.2.3.** *Denote by $|I(k)|$ the cardinality of the finite set $I(k)$, i.e., the number of rapid squares at step $k$. There exists a constant $c_8 > 0$ such that, for each step $k = 0, 1, 2, \ldots$, we have*

$$|I(k)| \leq c_8 \delta^{-1} \beta^*.$$

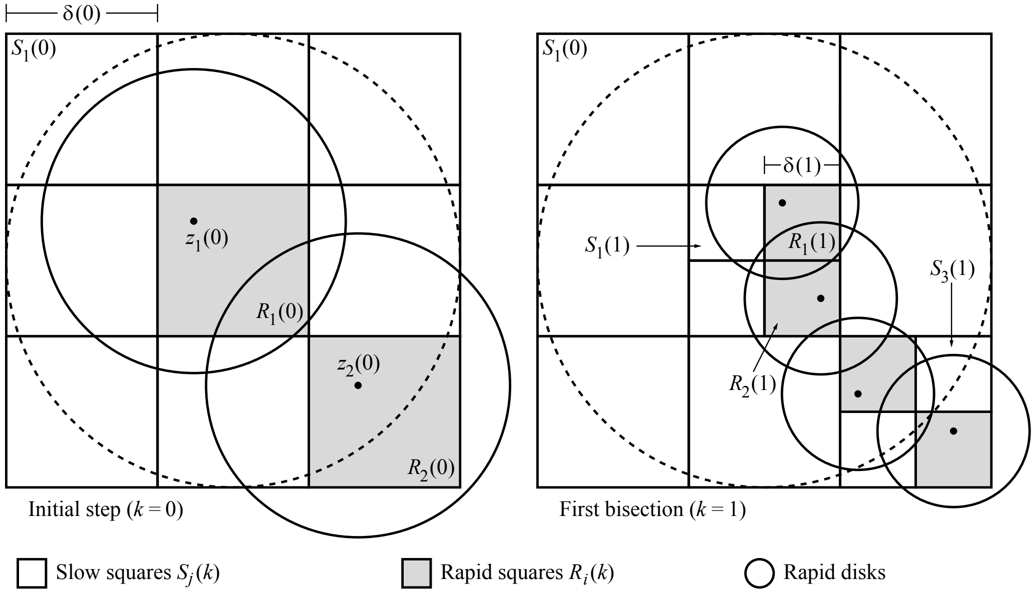


**Figure 4.** Iterative tiling of P in rapid and slow squares.

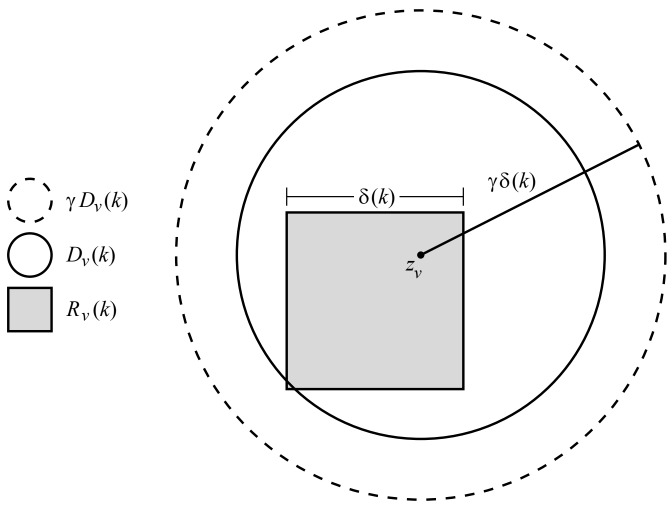


**Figure 5.** A close-up of a rapid square.

*Proof.* Recall that $\delta(k) := 2^{-k}\delta(0)$. Since $\delta(0)$ satisfies the constraints (4.1.2), it follows that $\delta(k)$ is $\beta^*$-related for all $k \in \mathbb{N} \cup \{0\}$.

We choose some $\nu \in I(k)$ and recall that there is one rapid-growth disk $D_\nu(k)$ whose center $z_\nu$ lies in $R_\nu(k)$. Notice that, since $\gamma\delta(k) > \sqrt{2}\delta(k)$, we have $R_\nu(k) \subset \gamma D_\nu(k)$ as shown in Figure 5.

Thus,

$$\bigcup_{\nu\in I(k)} R_\nu(k) \subset \bigcup_{\nu\in I(k)} \gamma D_\nu(k).$$

We now choose a *maximal* subcollection of disjoint disks $\gamma D_\nu$ and denote by $I^*(k) \subset I(k)$ the corresponding set of indices. Notice that disjointness of two scaled disks $\gamma D_\nu$ and $\gamma D_\mu$ is equivalent to $\gamma$-separation of $D_\nu$ and $D_\mu$. By maximality, for $\mu \in I(k) \setminus I^*(k)$, there exists $\nu \in I^*(k)$ such that $|z_\mu - z_\nu| \le 2\gamma\delta(k)$. In this case and for all $z \in \gamma D_\mu(k)$, we thus have

$$|z - z_\nu| \le |z - z_\mu| + |z_\mu - z_\nu| \le \gamma\delta(k) + 2\gamma\delta(k) < 4\gamma\delta(k).$$

As a consequence, we get the inclusion $\gamma D_\mu(k) \subset 4\gamma D_\nu(k)$, where $\mu$ represents a disk excluded from the maximal subset. This in turn means

$$\bigcup_{\nu\in I(k)} \gamma D_\nu(k) \subset \bigcup_{\nu\in I^*(k)} 4\gamma D_\nu(k).$$

Hence,

$$\bigcup_{\nu\in I(k)} R_\nu(k) \subset \bigcup_{\nu\in I^*(k)} 4\gamma D_\nu(k).$$

We compare the respective areas of the regions covered by the last inclusion and get $|I(k)|\delta^2(k) \le 16\pi\gamma^2\delta^2(k)|I^*(k)|$. By Proposition 4.2.1, $|I^*(k)| \le c_5\beta^*$ and we finally get

$$|I(k)| \le 16\pi\gamma^2|I^*(k)| \le 16\pi c_5\gamma^2\beta^* = c_8\delta^{-1}(0)\beta^*,$$

which concludes the proof since $I$ is precisely the set indexing the rapid squares. □

**Lemma 4.2.4.** *Denote by $|J(k)|$ the number of slow squares $S_j(k)$ obtained at step $k$. Then, for any $k = 0, 1, 2, \ldots,$ we have*

$$|J(k)| \le 4c_8\delta^{-1}\beta^*.$$

*Proof.* By construction, we have $|J(k)| \le 4|I(k-1)| \le 4c_8\delta^{-1}\beta^*$. □

**Lemma 4.2.5.** *There exists a constant $c_9$ such that, for each slow square $S_j(k)$ and each $k = 0, 1, 2, \ldots,$*

$$\mathcal{H}^1(Z_F \cap S_j(k)) \le c_9 2^{-k}\delta.$$

*Proof.* If $z_\mu$ lies in some slow square $S_i(k)$, then the disk $D(z_\mu, \delta(k))$ is slow, which means it satisfies

$$\int_{(1-3a)\delta<|z-z_\mu|<(1-\frac{4}{3}a)\delta} F^2 > M^{-1} \int_{(1-\frac{3}{2}a)\delta<|z-z_\mu|<(1-a)\delta} F^2.$$

By Proposition 4.2.2, we thus have

$$\mathcal{H}^1(Z_F \cap \mathfrak{D}(z_\mu, c_6 2^{-k}\delta)) \le c_7 2^{-k}\delta,$$

which holds for all $z_\mu \in S_j(k)$. We can now pick a finite collection of $N_0 = N_0(c_6)$ points $z_l \in S_j(k)$ such that the reunion of the associated disks $D(z_l, c_6 2^{-k}\delta)$ cover $S_j(k)$. The collection being finite, we have

$$\mathcal{H}^1(Z_F \cap S_j(k)) \le \sum_{l=1}^{N_0} \mathcal{H}^1(Z_F \cap D(z_l, c_6 2^{-k}\delta)) \le (N_0 c_7) 2^{-k}\delta = c_9 2^{-k}\delta. \qquad \square$$

The next result is exactly Lemma 6.3 in [Donnelly and Fefferman 1990].

**Lemma 4.2.6.** *The union* $\bigcup_{j\in J(k),\ k\in\mathbb{N}\cup\{0\}} S_j(k)$ *covers the whole square*

$$P = \left\{|x|, |y| \le \tfrac{1}{60}\right\}$$

*except for the singular set* $\mathcal{S}_F := \{z \in P : F(z) = \nabla F(z) = 0\}$.

The last lemma allows us to discard the singular set when studying the length of the nodal set of $F$.

**Lemma 4.2.7.** *Let* $\mathcal{S}_F$ *be the singular set of* $F$ *in* $P$. *Then*

$$\mathcal{H}^1(\mathcal{S}) = 0.$$

*Proof.* It is well-known (see for instance [Bers 1955; Han and Lin 2007]) that the singular set $\mathcal{S}$ of a $F$ is a submanifold of codimension 2, which means here that it is a finite set of points, whence $\mathcal{H}^1(\mathcal{S}) = 0$. □

We are now ready to complete the proof of Theorem 2.1.1.

*Proof of Theorem 2.1.1.* Using all of the above lemmas, we have

$$\begin{aligned}
\mathcal{H}^1\big(Z_F \cap \tfrac{1}{60}\mathbb{D}\big) \le \mathcal{H}^1(Z_F \cap P) &\underset{6,7}{=} \sum_{k=0}^{\infty} \sum_{j\in J(k)} \mathcal{H}^1(Z_F \cap S_j(k)) \\
&\underset{5}{\le} c_9\delta \sum_{k=0}^{\infty} \sum_{j\in J(k)} 2^{-k} \underset{4}{\le} (c_9\delta) 4c_8\delta^{-1}\beta^* \sum_{k=0}^{\infty} 2^{-k} \\
&= 4c_8c_9\beta^* \sum_{k=0}^{\infty} 2^{-k} \le c_3\beta^*.
\end{aligned}$$

□

## 5. Proof of Proposition 4.2.1

We divide the rather long proof into six subsections. The treatment is based on the proof of Proposition 4.7 in [Donnelly and Fefferman 1990].

**5.1.** ***Setting.*** Using the same hypotheses, we will actually prove a slightly different statement. We let $t := \beta + 1$. It follows from the fact that $\delta\beta^* < \frac{1}{2}$ that

$$\delta t < 1. \tag{5.1.1}$$

We normalize $F$ by the condition $\sup_{3\mathbb{D}}|F| = 1$, which has no effect whatsoever on the growth exponent. Finally, we can choose the uniform norm of the potential to be conveniently small: $\|q\|_\infty < \epsilon_0 < 1$. We will show that there exists a constant $c_5 > 0$ such that, for a large enough $M = M_0$, the number $\mathcal{N} = \mathcal{N}(M)$ of $\gamma$-separated, $M$-rapid disks satisfies

$$\mathcal{N} < c_5 t,$$

which implies the result since $t \le 2\beta^* = 2\max\{\beta, 1\}$. We recall that we are still in the setting of disks and annuli described in Section 4.1, that is, we have an arbitrary, finite collection of open disks $D_\nu \subset \frac{1}{60}\mathbb{D}$,

$1 \le \nu \le N$, each of radius $\delta$. Moreover, the collection of disks is $\gamma$-separated: the disks are mutually disjoint after a scaling of factor $\gamma$,

$$|z_\mu - z_\nu| \ge 2\gamma\delta \quad \text{for all } \mu \ne \nu,$$

where $\gamma = \delta^{-1/2}$.

**5.2. *A Carleman-type estimate.*** The starting point of the proof is (2.4) of [Donnelly and Fefferman 1990], which is an estimate in the spirit of Carleman, relating the weighted $L^2$ norm of a function to that of some of its derivatives.

**Lemma 5.2.1.** *Let $t > 0$, and define*

$$P(z) := \prod_\nu (z - z_\nu).$$

*There exists a constant $c_{10} > 0$ such that, for any $f \in C_0^\infty\big(3\mathbb{D} \setminus \bigcup_\nu D_\nu(a)\big)$, we have*

$$\int_{3\mathbb{D}} |\Delta f|^2 |P|^{-2} e^{t|z|^2} \ge c_{10}\bigg(t^2 \int_{3\mathbb{D}} |f|^2 |P|^{-2} e^{t|z|^2} + \delta^{-2} \int_A |\nabla f|^2 |P|^{-2} e^{t|z|^2}\bigg). \tag{C1}$$

The rather long development of that inequality is postponed to Section 6. Our first goal is to replace $|\nabla f|^2$ by $|f|^2$ in the right-hand side of the Carleman estimate. To do so, we will need the next two lemmas.

**Lemma 5.2.2.** *There exist positive constants $c_i$, $i = 11, \dots, 14$, such that, for any $w_1, w_2 \in A_\nu$,*

$$\text{(i)} \quad c_{11} \le \frac{e^{t|w_1|^2}}{e^{t|w_2|^2}} \le c_{12}, \qquad \text{(ii)} \quad c_{13} \le \frac{|P(w_1)|}{|P(w_2)|} \le c_{14}.$$

*Proof.* Since $w_1, w_2 \in \frac{1}{60}\mathbb{D}$, we have

$$\big|t|w_1|^2 - t|w_2|^2\big| = t\big|(|w_1| - |w_2|)(|w_1 + w_2|)\big| \le t\big||w_1| - |w_2|\big| \le t|w_1 - w_2| \le 2t\delta.$$

Since $t\delta \le 1$, the result (i) now follows from exponentiation.

We now prove (ii). We have

$$\begin{aligned}\big|\log|P(w_1)| - \log|P(w_2)|\big| &= \bigg|\sum_\mu \log|w_1 - z_\mu| - \sum_\mu \log|w_2 - z_\mu|\bigg| \\ &\le \big|\log|w_1 - z_\nu| - \log|w_2 - z_\nu|\big| + \sum_{\mu \ne \nu}\big|\log|w_1 - z_\mu| - \log|w_2 - z_\mu|\big|.\end{aligned}$$

We first consider the first term of the right-hand side of the above inequality. Suppose without loss of generality that $w_1$ is farther from $z_\nu$ than $w_2$, that is, $|w_1 - z_\nu| = \max\{|w_1 - z_\nu|, |w_2 - z_\nu|\}$. Then since both $w_1$ and $w_2$ belong to the annulus $A_\nu$, we have

$$\begin{aligned}\big|\log|w_1 - z_\nu| - \log|w_2 - z_\nu|\big| &= \log|w_1 - z_\nu| - \log|w_2 - z_\nu| \\ &\le \log(1-a)\delta - \log(1-2a)\delta \\ &= \log\frac{1-a}{1-2a} = a_2,\end{aligned}$$

where $a_2 > 0$. It now remains to estimate $\sum_{\mu\neq\nu}\big|\log|w_1 - z_\mu| - \log|w_2 - z_\mu|\big|$. By the mean value theorem applied to $w \mapsto |w - z_\mu|$, there exists some point $w \in \{(1-\tau)w_1 + \tau w_2 : 0 \leq \tau \leq 1\}$ such that

$$\big|\log|w_1 - z_\mu| - \log|w_2 - z_\mu|\big| = |w - z_\mu|^{-1}|w_1 - w_2|.$$

The triangle inequality also implies $|z_\mu - z_\nu| \leq |w - z_\mu| + |w - z_\nu| \leq 2|w - z_\mu|$, whence $|w - z_\mu|^{-1} \leq 2|z_\mu - z_\nu|$ and

$$\big|\log|w_1 - z_\mu| - \log|w_2 - z_\mu|\big| \leq 2\frac{|w_1 - w_2|}{|z_\mu - z_\nu|} \leq \frac{4\delta}{|z_\mu - z_\nu|}.$$

We now have

$$\sum_{\mu\neq\nu}\big|\log|w_1 - z_\mu| - \log|w_2 - z_\mu|\big| \leq 4\delta\sum_{\mu\neq\nu}|z_\mu - z_\nu|^{-1}. \tag{5.2.3}$$

For $z \in \gamma D_\mu$, $\mu \neq \nu$, we have $|z - z_\nu| + |z_\mu - z_\nu| \leq 2|z_\mu - z_\nu|$, from which we easily get

$$\int_{\gamma D_\mu}|z - z_\nu|^{-1} \geq \frac{1}{2}\int_{\gamma D_\mu}\frac{1}{|z_\mu - z_\nu|} = \frac{\pi(\gamma\delta)^2}{2|z_\mu - z_\nu|}.$$

We define $E_\nu := \bigcup_{\mu\neq\nu}\gamma D_\mu$, and we now have

$$4\delta\sum_{\mu\neq\nu}|z_\mu - z_\nu|^{-1} \leq \frac{8\delta}{\pi(\gamma\delta)^2}\sum_{\mu\neq\nu}\int_{\gamma D_\mu}|z - z_\nu|^{-1} = \frac{8}{\pi\gamma^2\delta}\int_{E_\nu}|z - z_\nu|^{-1}. \tag{5.2.4}$$

Let $B_\nu$ be the disk centered at $z_\nu$ whose total area is the same as $E_\nu$; that is, $\text{Area}(B_\nu) = \text{Area}(E_\nu) = (N-1)\pi(\gamma\delta)^2$. Remark that the maximum number of $\gamma$-separated disks of radius $\delta$ in $3\mathbb{D}$ is of the order $(\gamma\delta)^{-2}$; that is, there exists a positive constant $c$, independent of $\gamma$ and $\delta$, such that the cardinality $N$ of our collection of disks satisfies $N < c(\gamma\delta)^{-2}$. We consequently have

$$\int_{E_\nu}|z - z_\nu|^{-1} \leq \int_{B_\nu}|z - z_\nu|^{-1} \leq 4\sqrt{\text{Area}(E_\nu)} \leq 4\sqrt{\pi N}\gamma\delta \leq 4\sqrt{c\pi}. \tag{5.2.5}$$

Combining (5.2.3), (5.2.4) and (5.2.5) now gives

$$\sum_{\mu\neq\nu}\big|\log|w_1 - z_\mu| - \log|w_2 - z_\mu|\big| \leq \frac{32\sqrt{c\pi}}{\pi\gamma^2\delta} = \frac{a_3}{\gamma^2\delta} = a_3$$

since $\gamma = \delta^{-1/2}$. Finally,

$$\big|\log|P(w_1)| - \log|P(w_2)|\big| \leq a_2 + a_3,$$

from which the result follows via exponentiation. □

The second lemma is a Poincaré-like inequality.

**Lemma 5.2.6.** *Suppose $f \in C^\infty(A_\nu)$ and vanishes on the inner boundary $|z| = (1-2a)\delta$ of $A_\nu$. Then*

$$\int_{A_\nu}|\nabla f|^2 \geq \frac{c_{15}}{\delta^2}\int_{A_\nu}|f|^2, \tag{5.2.7}$$

*where $c_{15}$ is a positive constant.*

*Proof.* We introduce polar coordinates $(r,\theta)$ on $A_\nu$. Since $f((1-2a)\delta,\theta)\equiv 0$, the fundamental theorem of calculus yields

$$f(r,\theta)=\int_{(1-2a)\delta}^{r}\frac{\partial f}{\partial s}(s,\theta)\,ds.$$

Hence,

$$\int_{A_\nu}|f|^2\,\mathrm{d}A=\int_0^{2\pi}\int_{(1-2a)\delta}^{(1-a)\delta}\biggl(\int_{(1-2a)\delta}^{r}\frac{\partial f}{\partial s}(s,\theta)\,ds\biggr)^2 r\,dr\,d\theta.$$

By Cauchy–Schwarz, we have

$$\biggl(\int_{(1-2a)\delta}^{r}\frac{\partial f}{\partial s}(s,\theta)\,ds\biggr)^2\le\int_{(1-2a)\delta}^{r}\Bigl(\frac{\partial f}{\partial s}\Bigr)^2 ds\int_{(1-2a)\delta}^{r}1^2\,ds\le a\delta\int_{(1-2a)\delta}^{r}\Bigl(\frac{\partial f}{\partial s}\Bigr)^2 ds.$$

Consequently,

$$\begin{aligned}\int_{A_\nu}|f|^2&\le a\delta\int_0^{2\pi}\int_{(1-2a)\delta}^{(1-a)\delta}\int_{(1-2a)\delta}^{(1-a)\delta}\Bigl(\frac{\partial f}{\partial s}\Bigr)^2 ds\,r\,dr\,d\theta\\&\le a\delta\int_0^{2\pi}\frac{r^2}{2}\bigg|_{(1-2a)\delta}^{(1-a)\delta}\int_{(1-2a)\delta}^{(1-a)\delta}\Bigl(\frac{\partial f}{\partial s}\Bigr)^2\frac{s}{(1-2a)\delta}\,ds\,d\theta\\&\le c_{15}\delta^2\int_0^{2\pi}\int_{(1-2a)\delta}^{(1-a)\delta}|\nabla f|^2 s\,ds\,d\theta=c_{15}\delta^2\int_{A_\nu}|\nabla f|^2.\end{aligned}$$

□

Fix one $w_\nu\in A_\nu$ for all $1\le\nu\le N$. Then, for each $\nu$, we have

$$\begin{aligned}\int_{A_\nu}|\nabla f|^2|P|^{-2}e^{t|z|^2}&\ge(c_{11}c_{13}^2)\frac{e^{t|w_\nu|^2}}{|P(w_\nu)|^2}\int_{A_\nu}|\nabla f|^2\ge(c_{11}c_{13}^2c_{15})\frac{e^{t|w_\nu|^2}}{\delta^2|P(w_\nu)|^2}\int_{A_\nu}f^2\\&\ge(c_{11}^2c_{13}^4c_{15})\delta^{-2}\int_{A_\nu}f^2|P|^{-2}e^{t|z|^2},\end{aligned}$$

where we have used, respectively, Lemmas 5.2.2, 5.2.6 and then 5.2.2 again. The Carleman estimate (C1) thus becomes

$$\int_{3\mathbb{D}}|\Delta f|^2|P|^{-2}e^{t|z|^2}\ge a_4\biggl(t^2\int_{3\mathbb{D}}f^2|P|^{-2}e^{t|z|^2}+\delta^{-4}\int_A f^2|P|^{-2}e^{t|z|^2}\biggr),\tag{C2}$$

where $a_4:=\min\{c_{11}^2c_{13}^4c_{15},c_{10}\}$.

**5.3.** ***A suitable cut-off for F.*** We now apply the previous estimate to $f=\theta F$, where $\theta$ is a suitable cut-off. More precisely, the cut-off $\theta$ satisfies the following properties:

(i) $0\le\theta\le 1$, $\theta\in C_0^\infty\bigl(2\mathbb{D}\setminus\bigcup_\nu D_\nu\bigr)$,

(ii) $\theta(z)\equiv 1$ on $\bigl\{z:|z|<1,\ |z-z_\nu|>\bigl(1-\frac{3}{2}a\bigr)\delta\bigr\}$,

(iii) $|\nabla\theta|+|\Delta\theta|\le a_5$ on $\{|z|>1\}$,

(iv) $|\nabla\theta|\le a_6\delta^{-1}$ and $|\Delta\theta|\le a_7\delta^{-2}$ for $|z-z_\nu|\le\bigl(1-\frac{3}{2}a\bigr)\delta$.

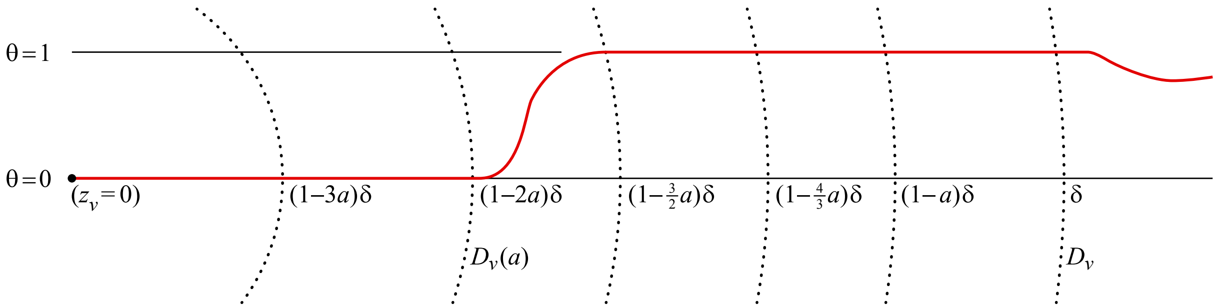


**Figure 6.** A smooth cut-off $\theta$ defined on $2\mathbb{D}$.

The property (iv) allows us to control the growth properties of the cut-off in terms of the radius $\delta$ of the disks. Figure 6 summarizes the property of the cut-off.

Using the properties of $\theta$, we have the following:

**Lemma 5.3.1.** *Let $F$ and $\theta$ be as defined in our current setting. Then*

$$|\Delta(\theta F)| \le 5(q^2F^2 + |\nabla\theta|^2|\nabla F|^2 + F^2|\Delta\theta|^2).$$

*Proof.* The proof is a simple computation:

$$\begin{aligned} |\Delta(\theta F)|^2 &= |\theta\Delta F + 2(\nabla\theta\cdot\nabla F) + F\Delta\theta|^2 \\ &\le (|\theta\Delta F| + 2|\nabla\theta||\nabla F| + |F\Delta\theta|)^2 \\ &\le 5(\theta^2|-qF|^2 + |\nabla\theta|^2|\nabla F|^2 + F^2|\Delta\theta|^2) \\ &\le 5(q^2F^2 + |\nabla\theta|^2|\nabla F|^2 + F^2|\Delta\theta|^2). \end{aligned}$$ □

Applying (C2) to $\theta F$ now yields

$$\int_{2\mathbb{D}} |\Delta\theta F|^2|P|^{-2}e^{t|z|^2} \ge a_4\left(t^2\int_{2\mathbb{D}}|\theta F|^2|P|^{-2}e^{t|z|^2} + \delta^{-4}\int_A|\theta F|^2|P|^{-2}e^{t|z|^2}\right).$$

Using Lemma 5.3.1 to estimate the left-hand side of the above equation, we now get

$$\int_{2\mathbb{D}}\big(q^2F^2+|\nabla\theta|^2|\nabla F|^2+F^2|\Delta\theta|^2\big)|P|^{-2}e^{t|z|^2} \ge \frac{a_4}{5}\left(t^2\int_{2\mathbb{D}}|\theta F|^2|P|^{-2}e^{t|z|^2}+\delta^{-4}\int_A|\theta F|^2|P|^{-2}e^{t|z|^2}\right).$$

Now, since our potential is small, $\|q\|_\infty < \epsilon_0$, the first term of the left-hand side can without loss of generality (by picking a smaller constant if needed) be absorbed by the right-hand side, yielding

$$\int_{2\mathbb{D}}\big(|\nabla\theta|^2|\nabla F|^2 + F^2|\Delta\theta|^2\big)|P|^{-2}e^{t|z|^2} \ge a_8\left(t^2\int_{2\mathbb{D}}|\theta F|^2|P|^{-2}e^{t|z|^2} + \delta^{-4}\int_A|\theta F|^2|P|^{-2}e^{t|z|^2}\right). \quad \text{(C3)}$$

The remainder of the proof consists mostly of improvements of the left- and right-hand sides of this last estimate.

**5.4.** ***Using elliptic theory to improve the left-hand side of* (C3).** We now work on the left-hand side of the last Carleman estimate. By the definition of the cut-off $\theta$, we have $|\nabla\theta| = |\Delta\theta| \equiv 0$ on $2\mathbb{D} \setminus \big(A = \bigcup_\nu A_\nu \cup \{1 \le |z| \le 2\}\big)$ so that it makes sense to write (LHS) $= I + \sum_\nu I_\nu$, where

$$I = \int_{1<|z|<2} \zeta(z), \qquad I_\nu = \int_{A_\nu} \zeta(z)$$

and

$$\zeta(z) = \big(|\nabla\theta|^2|\nabla F|^2 + F^2|\Delta\theta|^2\big)|P|^{-2}e^{t|z|^2}.$$

The following lemma uses elliptic theory to improve estimates of both $I$ and $I_\nu$:

**Lemma 5.4.1.** *There exist positive constants $c_{11}$ and $c_{12}$ such that*

$$\text{(i)} \quad I \le c_{16}e^{4t}\max_{|z|\ge 1}|P|^{-2}\int_{3/4<|z|<9/4} F^2, \qquad \text{(ii)} \quad I_\nu \le c_{17}\delta^{-4}\max_{A_\nu}(|P|^{-2}e^{t|z|^2})\int_{A'_\nu} F^2.$$

*Proof.* Recalling the various assumptions on the cutoff $\theta$, we immediately have

$$\begin{aligned} I &\le a_5 e^{4t}\max_{1\le|z|\le 2}|P|^{-2}\int_{1<|z|<2}(F^2+|\nabla F|^2) \\ &= a_5 e^{4t}\max_{1\le|z|\le 2}|P|^{-2}\|F\|^2_{H^1(\Omega')}, \end{aligned} \tag{5.4.2}$$

where $H^1 = W^{1,2}$ is the habitual Sobolev space and $\Omega' = \{1 < |z| < 2\}$. We now apply Theorem 8.8 in [Gilbarg and Trudinger 1998] with $L = \Delta$, $u = F$ and $f = -qF$ to get

$$\begin{aligned} \|F\|_{W^{2,2}(\Omega')} &\le a_9\big(\|F\|_{L^2(\Omega)} + \|qF\|_{L^2(\Omega)}\big) \\ &\le a_9\max\{1, \operatorname{Area}(\Omega)\epsilon_0\}\|F\|_{L^2(\Omega)} \\ &= a_{10}\|F\|_{L^2(\Omega)}, \end{aligned}$$

which holds for any subdomain $\Omega$ such that $\Omega' \Subset \Omega$, that is,

$$\sup_{x\in\partial\Omega,\ y\in\Omega'}|x-y| > 0.$$

We set $\Omega := \left\{\frac{3}{4} < |z| < \frac{9}{4}\right\}$ so that the above condition is satisfied. Since $\|\cdot\|_{W^{1,2}} \le \|\cdot\|_{W^{2,2}}$, we have

$$\|F\|^2_{H^1(\Omega')} \le a_{10}^2\|F\|^2_{L^2(\Omega)}$$

so that estimate (5.4.2) becomes

$$I \le (a_5 a_{10}^2)e^{4t}\max_{1\le|z|\le 2}|P|^{-2}\|F\|^2_{L^2(\Omega)} = c_{11}e^{4t}\max_{|z|\ge 1}|P|^{-2}\int_{3/4<|z|<9/3} F^2.$$

We now prove the second part of the lemma. We define $\bar{A}_\nu := (1-2a)\delta < |z-z_\nu| < (1-3a/2)\delta \subset A_\nu$. Since $\theta(z) \equiv 1$ for $(1-3a/2)\delta < |z| < (1-a)\delta$, we have

$$\begin{aligned} I_\nu &\le \max_{A_\nu}\big(|P|^{-2}e^{t|z|^2}\big)\int_{\bar{A}_\nu}(|\nabla\theta|^2|\nabla F|^2 + F^2|\Delta\theta|^2) \\ &\le \max\{a_6^2, a_7^2\}\max_{A_\nu}\big(|P|^{-2}e^{t|z|^2}\big)\left[\int_{\bar{A}_\nu}\delta^{-2}|\nabla F|^2 + \int_{\bar{A}_\nu}\delta^{-4}F^2\right]. \end{aligned} \tag{5.4.3}$$

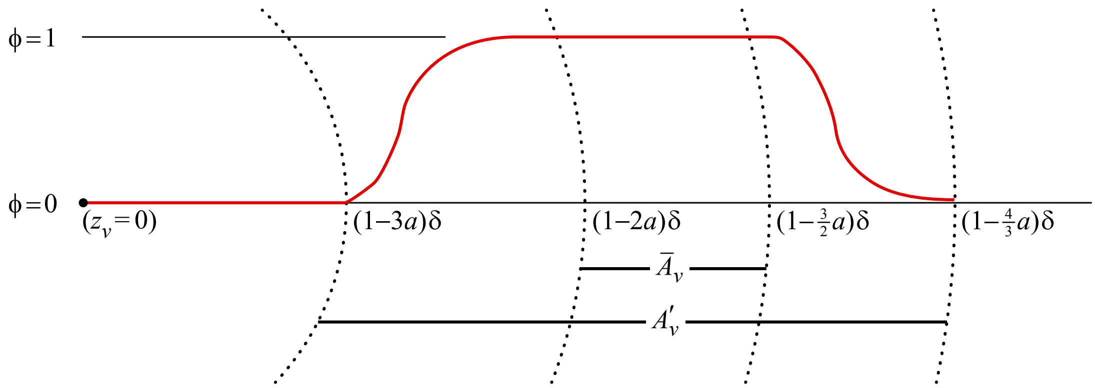


**Figure 7.** A second cutoff $\phi$ on the annuli.

Our goal is now to get rid of the gradient in the first integral of the last equation above. To do so, we set $\bar{I}_\nu := \int_{\bar{A}_\nu} |\nabla F|^2$ and introduce another cutoff $\phi \in C_0^\infty(A'_\nu)$ that satisfies

(i) $0 \le \phi \le 1$,

(ii) $\phi(z) \equiv 1$ on $\bar{A}_\nu$,

(iii) $|\nabla\phi| \le a_{11}(\phi\delta^{-1})$.

Figure 7 summarizes the properties of this cutoff function.

Using Green's identity and since $\phi$ vanishes on the boundary of $A'_\nu$, we notice that

$$\int_{A'_\nu} q\phi F^2 = -\int_{A'_\nu} \phi F \Delta F = \int_{A'_\nu} \nabla(\phi F)\cdot\nabla F = \int_{A'_\nu} F(\nabla F\cdot\nabla\phi) + \int_{A'_\nu} \phi|\nabla F|^2.$$

Thus, since $\|q\|_\infty < 1$, we get

$$\int_{A'_\nu} \phi|\nabla F^2| \le \int_{A'_\nu} \phi F^2 + \frac{a_{11}}{\delta}\int_{A'_\nu} \phi|F|\|\nabla F\|. \tag{5.4.4}$$

Now, for any nonnegative numbers $a$, $b$ and $c$ and $k > 0$, we have the elementary inequality $abc \le \frac{1}{2}(ab^2/k - kac^2)$, which we apply to our setting to get

$$\phi\left(\frac{|F|}{\delta}\right)\|\nabla F\| \le \frac{1}{2}\left(\frac{\phi F^2}{k\delta^2} + k\phi|\nabla F|^2\right).$$

We integrate over $A'_\nu$ and then choose $k$ small enough to absorb $\frac{1}{2}(k\phi|\nabla F|^2)$ in the left-hand side of (5.4.4) so that it becomes

$$\int_{A'_\nu} \phi|\nabla F|^2 \le \max\left\{1, \frac{a_{11}}{2k}\right\}\frac{1}{\delta^2}\int_{A'_\nu} \phi F^2.$$

Going back to the definition of $\bar{I}_\nu$, we now have

$$\bar{I}_\nu = \int_{\bar{A}_\nu} |\nabla F|^2 \le \int_{A'_\nu} \phi|\nabla F|^2 \le a_{12}\frac{1}{\delta^2}\int_{A'_\nu} \phi F^2 \le a_{12}\delta^{-2}\int_{A'_\nu} F^2.$$

Plugging this into (5.4.3) yields

$$\begin{aligned} I_\nu &\le \max\{a_6^2, a_7^2\} \max_{A_\nu}\big(|P|^{-2}e^{t|z|^2}\big)\left[\delta^{-2}\bar{I}_\nu + \int_{\bar{A}_\nu} \delta^{-4}F^2\right] \\ &\le \big(\max\{a_6^2, a_7^2\}\max\{1, a_{12}\}\big) \max_{A_\nu}\big(|P|^{-2}e^{t|z|^2}\big)\delta^{-4}\int_{\bar{A}_\nu} F^2 \\ &\le c_{17}\max_{A_\nu}\big(|P|^{-2}e^{t|z|^2}\big)\delta^{-4}\int_{A'_\nu} F^2. \end{aligned}$$

□

By Lemma 5.2.2, we have

$$\max_{A_\nu}\big(|P|^{-2}e^{t|z|^2}\big) \le a_{13}\min_{A_\nu}\big(|P|^{-2}e^{t|z|^2}\big).$$

Applying the estimates of Lemma 5.4.1 to the left-hand side of (C3) then gives

$$(\text{LHS}) = I + \sum_\nu I_\nu \le a_{14}\bigg(e^{4t}\max_{|z|\ge 1}|P|^{-2}\int_{3/4<|z|<9/4} F^2 + \delta^{-4}\sum_\nu \min_{A_\nu}\big(|P|^{-2}e^{t|z|^2}\big)\int_{A'_\nu} F^2\bigg), \quad (5.4.5)$$

where $a_{14} = \max\{c_{16}, c_{16}a_{13}\}$.

The next lemma introduces the growth exponent $\beta$ of F in an expression that links the $L^2$ norms of $F$ on two annuli of different sizes.

**Lemma 5.4.6.** *There exists a positive constant $c_{18}$ such that*

$$\int_{3/4<|z|<9/4} F^2 \le c_{18}e^{2\beta}\int_{1/4<|z|<1/2} F^2.$$

*Proof.* First, recall that the potential $q$ satisfies $\|q\|_\infty < \epsilon_0$. On the one hand,

$$\iint_{3/4<|z|<9/4} F^2\,\mathrm{d}A < \iint_{\frac{5}{2}\mathbb{D}} F^2\,\mathrm{d}A \le \left(\frac{25\pi}{4}\right)\sup_{\frac{5}{2}\mathbb{D}} F^2. \quad (5.4.7)$$

On the other hand, the definition of the growth exponent yields

$$\sup_{\frac{5}{2}\mathbb{D}} F^2 = e^{2\beta}\sup_{\frac{1}{4}\mathbb{D}} F^2. \quad (5.4.8)$$

Following an approach similar to Lemma 4.9 in [Nazarov et al. 2005], we now represent $F$ as the sum of its Green potential and Poisson integral. More precisely, for $|z| < \frac{1}{4}$ and given any fixed radius $\rho \in \big(\frac{1}{4}, \frac{1}{2}\big]$,

$$F(z) = \iint_{\rho\mathbb{D}} p(\zeta)F(\zeta)G_\rho(z,\zeta)\,\mathrm{d}A(\zeta) + \int_{\rho\mathbb{S}^1} F(\zeta)P_\rho(z,\zeta)\,ds(\zeta), \quad (5.4.9)$$

where $G_\rho(z,\zeta) = \log|(\rho^2 - z\bar{\zeta})/\rho(z-\zeta)|$ and $P_\rho(z,\zeta) = (\rho^2 - |z|^2)/|\zeta - z|^2$. We write $I_1$ and $I_2$, respectively, for the double integral and the (line) integral above and notice that

$$F^2 = I_1^2 + 2I_1I_2 + I_2^2 \le 4(I_1^2 + I_2^2). \quad (5.4.10)$$

Using Cauchy–Schwartz, we get the upper bound

$$\begin{aligned} I_1^2 &\le \iint_{\rho\mathbb{D}} p^2(\zeta)F^2(\zeta)\,\mathrm{d}A(\zeta)\iint_{\rho\mathbb{D}} G_\rho^2(z,\zeta)\,\mathrm{d}A(\zeta) \\ &\le a_{15}\iint_{\rho\mathbb{D}} p^2(\zeta)F^2(\zeta)\,\mathrm{d}A(\zeta) \le a_{15}\|p\|_\infty^2 \iint_{\rho\mathbb{D}} F^2(\zeta)\,\mathrm{d}A(\zeta) \\ &\le a_{15}\epsilon_0^2 \iint_{\frac{1}{2}\mathbb{D}} F^2(\zeta)\,\mathrm{d}A(\zeta). \end{aligned} \tag{5.4.11}$$

In the above, we have $a_{15} = \sup_{\rho\in(1/4,1/2]}\sup_{z\in\frac{1}{4}\mathbb{D}} \iint_{\rho\mathbb{D}} G_\rho^2(z,\zeta)\,\mathrm{d}A(\zeta)$. Similarly,

$$I_2^2 \le \int_{\rho\,\mathbb{S}^1} F^2(\zeta)\,ds(\zeta)\int_{\rho\mathbb{S}^1} P_\rho^2(z,\zeta)\,ds(\zeta) \le a_{16}\int_{\rho\mathbb{S}^1} F^2(\zeta)\,ds(\zeta) \tag{5.4.12}$$

with $a_{16} = \sup_{\rho\in(1/4,1/2]}\sup_{z\in\frac{1}{4}\mathbb{D}}\int_{\rho\mathbb{S}^1} P_\rho^2(z,\zeta)\,ds(\zeta)$. Now, recalling that the representation of $F$ in (5.4.9) holds for any $|z| \le \frac{1}{4}$ and substituting (5.4.11) and (5.4.12) into (5.4.10), we get

$$\sup_{z\in\frac{1}{4}\mathbb{D}} F^2 \le a_{17}\biggl(\epsilon_0^2\iint_{\frac{1}{2}\mathbb{D}} F^2\,\mathrm{d}A + \int_{\rho\mathbb{S}^1} F^2\,ds\biggr) \quad \text{for all } \rho\in\bigl(\tfrac{1}{4},\tfrac{1}{2}\bigr]$$

with $a_{17} = 4\max\{a_{15}, a_{16}\}$. Averaging over all $\rho$ yields

$$\begin{aligned} \sup_{z\in\frac{1}{4}\mathbb{D}} F^2 &\le a_{17}\frac{16}{3\pi}\biggl(\epsilon_0^2\iint_{\frac{1}{2}\mathbb{D}} F^2\,\mathrm{d}A + \iint_{1/4<|z|<1/2} F^2\,\mathrm{d}A\biggr) \\ &= a_{18}\epsilon_0^2\iint_{\frac{1}{4}\mathbb{D}} F^2\,\mathrm{d}A + a_{18}(1+\epsilon_0^2)\iint_{1/4<|z|<1/2} F^2\,\mathrm{d}A \\ &\le \Bigl(\frac{a_{18}\pi}{16}\Bigr)\epsilon_0^2\sup_{z\in\frac{1}{4}\mathbb{D}} F^2 + a_{18}(1+\epsilon_0^2)\iint_{1/4<|z|<1/2} F^2\,\mathrm{d}A \\ &= a_{19}\epsilon_0^2\sup_{z\in\frac{1}{4}\mathbb{D}} F^2 + a_{18}(1+\epsilon_0^2)\iint_{1/4<|z|<1/2} F^2\,\mathrm{d}A. \end{aligned} \tag{5.4.13}$$

Hence,

$$(1-a_{19}\epsilon_0^2)\sup_{z\in\frac{1}{4}\mathbb{D}} F^2 \le a_{18}(1+\epsilon_0^2)\iint_{1/4<|z|<1/2} F^2\,\mathrm{d}A.$$

It suffices to choose $\epsilon_0$ small enough so that $(1-a_{19}\epsilon_0^2)$ is positive to finally obtain

$$\sup_{z\in\frac{1}{4}\mathbb{D}} F^2 \le \frac{a_{18}(1+\epsilon_0^2)}{1-a_{19}\epsilon_0^2}\iint_{1/4<|z|<1/2} F^2\,\mathrm{d}A. \tag{5.4.14}$$

Linking (5.4.7), (5.4.8) and (5.4.14) together concludes the proof. □

To finalize our estimate of the left-hand side of (C3), we need one last lemma.

**Lemma 5.4.15.** *Let $N$ be the number of disks $D_\nu$ in our collection, that is, $N = \deg P$. Then there exists a positive constant $c_{19}$ such that*

$$\max_{z\ge 1}|P|^{-2} \le e^{-c_{19}N}\min_{|z|\le\frac{1}{2}}|P|^{-2}.$$

*Proof.* For $|z| \geq 1$, we have

$$\frac{1}{|z-z_\nu|} \leq \frac{1}{|z|-|z_\nu|} \leq \frac{1}{1-1/60} = \frac{60}{59}$$

while, for $|z| \leq 1/2$, we have

$$\frac{1}{|z-z_\nu|} \geq \frac{1}{|z|+|z_\nu|} \geq \frac{1}{1/2+1/60} = \frac{60}{31}.$$

As a consequence,

$$\max_{|z|\geq 1}|P|^{-2} \leq \left(\tfrac{60}{59}\right)^{2\deg P} = \left(\tfrac{31}{59}\right)^{2\deg P}\left(\tfrac{60}{31}\right)^{2\deg P} \leq \left(\tfrac{31}{59}\right)^{2\deg P} \min_{|z|\leq 1/2}|P|^{-2}.$$

We set $c_{19} = -2\log\left(\tfrac{31}{59}\right)$ to conclude the proof. □

Applying the results of the last two lemmas to (5.4.5), we obtain a final estimate for the left-hand side of (C3):

$$\begin{aligned}
\text{(LHS)} &\leq a_{14}\left(e^{4t}e^{-c_{19}N}\min_{|z|\leq 1/2}|P|^{-2}c_{18}e^{2\beta}\int_{1/4<|z|<1/2}F^2 + \delta^{-4}\sum_\nu \min_{A_\nu}\left(|P|^{-2}e^{t|z|^2}\right)\int_{A'_\nu}F^2\right) \\
&\leq a_{20}\left(e^{6t-c_{19}N}\min_{|z|\leq 1/2}|P|^{-2}\int_{1/4<|z|<1/2}F^2 + \delta^{-4}\sum_\nu \min_{A_\nu}\left(|P|^{-2}e^{t|z|^2}\right)\int_{A'_\nu}F^2\right),
\end{aligned} \tag{5.4.16}$$

where $a_{20} = a_{14}\max\{c_{18}, 1\}$, since $\beta < t$.

**5.5.** ***Improving the right-hand side of* (C3).** Recalling that $t > 1$ as well as the various properties of the cut-off, we estimate the right-hand side of (C3):

$$\begin{aligned}
a_8^{-1}\text{(RHS)} &= t^2\int_{2\mathbb{D}}|\theta F|^2|P|^{-2}e^{t|z|^2} + \delta^{-4}\sum_\nu\int_{A_\nu}|\theta F|^2|P|^{-2}e^{t|z|^2} \\
&\geq \left(\frac{\pi}{4}\right)\min_{|z|\leq 1/2}|P|^{-2}\int_{|z|<1/2}F^2 + \delta^{-4}\sum_\nu \min_{A_\nu}\left(|P|^{-2}e^{t|z|^2}\right)\int_{A_\nu}F^2 \\
&\geq a_{21}\left(\min_{|z|<1/2}|P|^{-2}\int_{1/4<|z|<1/2}F^2 + \delta^{-4}\sum_\nu \min_{A_\nu}\left(|P|^{-2}e^{t|z|^2}\right)\int_{A''_\nu}F^2\right),
\end{aligned} \tag{5.5.1}$$

where $a_{21} = \min\{\pi/4, 1\}$.

**5.6.** ***Conclusion.*** At last, putting together the estimates (C3), (5.4.16) and (5.5.1) yields

$$\begin{aligned}
&e^{6t-c_{19}N}\min_{|z|\leq 1/2}|P|^{-2}\int_{1/4<|z|<1/2}F^2 + \delta^{-4}\sum_\nu \min_{A_\nu}\left(|P|^{-2}e^{t|z|^2}\right)\int_{A'_\nu}F^2 \\
&\qquad\geq a_{22}\left(\min_{|z|<1/2}|P|^{-2}\int_{1/4<|z|<1/2}F^2 + \delta^{-4}\sum_\nu \min_{A_\nu}\left(|P|^{-2}e^{t|z|^2}\right)\int_{A''_\nu}F^2\right),
\end{aligned}$$

where $a_{22} = a_{21}a_8/a_{20}$. Recall that a disk $D_\nu$ is said to be $M$-rapid if

$$M\int_{A'_\nu}F^2 \leq \int_{A''_\nu}F^2.$$

Suppose now that all the disks of our collection are $M$-rapid, i.e., that $\mathcal{N} = N$, and assume without loss of generality that $a_{22} > 1$ (otherwise, the argument still works: it suffices to pick a larger $M$). We get

$$e^{6t-c_{19}N} \min_{|z|\le 1/2} |P|^{-2} \int_{1/4<|z|<1/2} F^2 + \delta^{-4} \sum_\nu \min_{A_\nu}\bigl(|P|^{-2} e^{t|z|^2}\bigr) \int_{A'_\nu} F^2$$
$$\ge \min_{|z|<1/2} |P|^{-2} \int_{1/4<|z|<1/2} F^2 + M\delta^{-4} \sum_\nu \min_{A_\nu}\bigl(|P|^{-2} e^{t|z|^2}\bigr) \int_{A'_\nu} F^2. \quad (5.6.1)$$

We get a contradiction if $N > \frac{6}{c_{19}} t \iff c_{19} N > c_6 t$, and the proof is completed.

## 6. An inequality in the spirit of Carleman

Carleman estimates are known to be useful in obtaining unique continuation results as well as growth estimates (see for instance [Koch and Tataru 2001]). It is thus not surprising that the estimate (C1) has played a crucial role in the proof of the growth estimate presented in the previous section. For completeness, we present here one way to obtain such an inequality, which follows very closely the approach taken in Section 2 of [Donnelly and Fefferman 1990].

**6.1.** ***An elementary inequality in a weighted Hilbert space.*** We let $\mathcal{D} \subset \mathbb{C}$ be open and bounded and $\varphi : \mathcal{D} \to \mathbb{R}$ be a smooth real-valued function. Let also $\mathcal{H} = L^2(\mathcal{D}, e^{-\varphi}\, dx\, dy)$ be the Hilbert space of complex-valued square-integrable functions on $\mathcal{D}$ with respect to the weight $e^{-\varphi}$. Finally, let $u \in C_0^\infty(\mathcal{D}) \subset H$. We introduce the differential operators

$$\partial := \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right), \qquad \bar\partial := \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right), \qquad \bar\partial^* := e^{-\varphi}\partial(e^{-\varphi}\,\cdot\,).$$

Easy computations allow one to verify the following facts:

(i) For any real-valued function $\psi$, $\bar\partial\partial\psi = \frac{1}{4}\Delta\psi$.

(ii) By the Cauchy–Riemann equations, $u$ is holomorphic if and only if $\bar\partial u = 0$.

(iii) $\bar\partial^*$ is the adjoint operator of $\bar\partial$.

(iv) $[\bar\partial, \bar\partial^*]u = \bigl(\frac{1}{4}\Delta\varphi\bigr)u$, where the interior of the parentheses acts on $u$ by multiplication.

**Lemma 6.1.1.** *Let $\Phi : \mathcal{D} \to \mathbb{R}$ be a smooth, positive function. Then*

$$\int_{\mathcal{D}} |\bar\partial u|^2 \Phi \ge \int_{\mathcal{D}} \tfrac{1}{4}(\Delta \log \Phi)|u|^2 \Phi,$$

*where the integrals are taken with respect to the usual Lebesgue measure, that is, not in the weighted Hilbert space $\mathcal{H}$.*

*Proof.* Put $\varphi := -\log\Phi$, i.e., $e^{-\varphi} = \Phi$. In the following, the norms and inner products are taken in the Hilbert space $\mathcal{H}$:

$$\begin{aligned} 0 \le \|\bar\partial^* u\|^2 = (\bar\partial^* u, \bar\partial^* u) &= (\bar\partial\bar\partial^* u, u) \\ &= (\bar\partial^*\bar\partial u, u) + ([\bar\partial, \bar\partial^*]u, u) \\ &= (\bar\partial u, \bar\partial u) + ([\bar\partial, \bar\partial^*]u, u) \\ &= \|\bar\partial u\|^2 + \int_{\mathcal{D}} \big(\tfrac{1}{4}\Delta\varphi\big)|u|^2 e^{-\varphi}. \end{aligned}$$

Thus, $\|\bar\partial u\|^2 \ge -\frac{1}{4}\int_{\mathcal{D}}(\Delta\varphi)|u|^2 e^{-\varphi} = \frac{1}{4}\int_{\mathcal{D}}(\Delta\log\Phi)|u|^2\Phi$. □

**6.2. *A specialized choice of weight function.*** The remainder of the section aims to specialize the choice of $\Phi$ in order to obtain a more refined inequality. In particular, we will build a weight function that has singularities on a crucial set of points. In the following, $a$ is a small, positive constant: $0 < a \ll 1$.

**Lemma 6.2.1.** *There exists a function $\Psi_0(z)$, defined for $|z| > (1-2a)$, such that*

(i) $a_1 \le \Psi_0(z) \le a_2$, *where* $a_1, a_2 > 0$,

(ii) $\Psi_0(z) \equiv 1$ *on* $\{|z| > 1\}$,

(iii) $\Delta\log\Psi_0 \ge 0$ *on* $\{|z| > (1-2a)\}$,

(iv) $\Delta\log\Psi_0 \ge a_3 > 0$ *on* $\{1-2a < |z| < 1-a\}$.

*Proof.* First, choose $\psi_0(z)$ to be a radial function, i.e., depending only on $r = |z|$. Let $h(r) \ge 0$ be smooth and such that $h(r) \ge a_3$ for $1-2a < r < 1-a$ and $h(r) = 0$ for $|z| > 1-a/2$. Now consider the radial Laplacian

$$\Delta\log\psi_0(r) = \left(\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr}\right)\log\psi_0(r),$$

which has smooth coefficients on $r > 1-2a$. By the fundamental theorem for ordinary differential equations, we let $\log\psi_0(r)$ be the solution of the second-order ODE

$$\begin{cases} \Delta\log\psi_0(r) = h(r), \\ \log\psi_0(1) = 0, \\ \log{\psi_0}'(1) = 0. \end{cases}$$

The function $\psi_0$ satisfies all the requirements. □

We now let $D_\nu := \{z : |z - z_\nu| < \delta\}$, $1 \le \nu \le N$, denote a finite collection of disks in the open unit disk $\mathbb{D}$ and let $D_\nu(a)$ be the closure of $(1-2a)D_\nu$. Define $\Phi_0 : \mathbb{C} \setminus \bigcup_\nu D_\nu(a)$ by

$$\Phi_0(z) = \begin{cases} 1 & \text{if } z \notin \bigcup_\nu D_\nu, \\ \psi_0\Big(\dfrac{z - z_\nu}{\delta}\Big) & \text{if } z \in D_\nu. \end{cases}$$

We have that $\log\Phi_0(z) = \log\Psi_0(w(z))$, where $w(z) = (z - z_\nu)/\delta$ and $w'(z) = 1/\delta$. Thus,

$$\Delta\log\Phi_0(z) = \frac{1}{\delta^2}\Delta\log\psi_0(w(z)) \ge a_3$$

for $z \in A_\nu = \{(1-2a)\delta < |z| < (1-a)\delta\}$. By Lemma 6.2.1, we have

(i) $a_1 \le \Phi_0(z) \le a_2$,

(ii) $\Delta \log \Phi_0 \ge 0$ for all $z \in \mathbb{C} \setminus \bigcup_\nu D_\nu(a)$,

(iii) $\Delta \log \Phi_0 \ge a_3/\delta^2$ for all $z \in A_\nu(\delta)$.

Let $t > 0$ be a constant, and denote by $A$ the union $\bigcup_\nu A_\nu(\delta)$. We want to apply Lemma 6.1.1 to $\Phi(z) := \Phi_0(z)e^{t|z|^2}$. For $u \in C_0^\infty\big(\mathbb{C} \setminus \bigcup_\nu D_\nu(a)\big)$, we assume that $\mathscr{D}$ is a bounded domain such that $\operatorname{supp} u \subset \mathscr{D}$ and $A \subset \mathscr{D} \subset \mathbb{C} \setminus \bigcup_\nu D_\nu(a)$. Applying the lemma gives

$$\int_{\mathscr{D}} |\bar\partial u|^2 \Phi_0(z) e^{t|z|^2} \ge \int_{\mathscr{D}} \tfrac{1}{4}\big(\Delta \log \Phi_0 e^{t|z|^2}\big)|u|^2 \Phi_0 e^{t|z|^2}. \tag{6.2.2}$$

But $\log \Phi_0 e^{t|z|^2} = \log \Phi_0 + t|z|^2$, and the right-hand side of the above inequality satisfies

$$\begin{aligned}\text{(RHS)} &= \left[\int_A + \int_{\mathscr{D}\setminus A}\right] \tfrac{1}{4}(\Delta \log \Phi_0)|u|^2 \Phi_0(z) e^{t|z|^2} + t\int_{\mathscr{D}} |u|^2 \Phi_0(z) e^{t|z|^2} \\ &\ge \frac{a_3}{\delta^2}\int_A |u|^2 \Phi_0(z) e^{t|z|^2} + t\int_{\mathscr{D}} |u|^2 \Phi_0(z) e^{t|z|^2}.\end{aligned}$$

Since $\Phi_0$ is bounded, we get

$$\int_{\mathscr{D}} |\bar\partial u|^2 e^{t|z|^2} \ge \frac{a_4}{\delta^2}\int_A |u|^2 e^{t|z|^2} + a_5 t\int_{\mathscr{D}} |u|^2 e^{t|z|^2}. \tag{6.2.3}$$

Define the holomorphic function $P(z) := \prod_\nu (z - z_\nu)$, and replace $u \mapsto u/P$. Then

$$\bar\partial\Big(\frac{u}{P}\Big) = \frac{\bar\partial u P - u\bar\partial P}{P^2} = \frac{\bar\partial u}{P}.$$

Since $u/P \in C_0^\infty(\mathscr{D})$, (6.2.3) becomes

$$\int_{\mathscr{D}} |\bar\partial u|^2 |P|^{-2} e^{t|z|^2} \ge \frac{a_4}{\delta^2}\int_A |u|^2 |P|^{-2} e^{t|z|^2} + a_5 t\int_{\mathscr{D}} |u|^2 |P|^{-2} e^{t|z|^2}. \tag{6.2.4}$$

All of the above discussion is valid for $u : \mathscr{D} \to \mathbb{C}$. We now choose $f : \mathscr{D} \to \mathbb{R}$. We have $|\bar\partial f| = |\partial f| = |\nabla f|$. We choose $u = \partial f$, whence $\bar\partial u = \bar\partial\partial f = \frac{1}{4}\Delta f$, which yields

$$\int_{\mathscr{D}} |\Delta f|^2 |P|^{-2} e^{t|z|^2} \ge \frac{a_6}{\delta^2}\int_A |\nabla f|^2 |P|^{-2} e^{t|z|^2} + a_7 t\int_{\mathscr{D}} |\bar\partial f|^2 |P|^{-2} e^{t|z|^2}.$$

We work on the last integral. Applying Lemma 6.1.1 to $\Phi = |P|^{-2} e^{t|z|^2}$, we get

$$\int_{\mathscr{D}} |\bar\partial f|^2 |P|^{-2} e^{t|z|^2} \ge \frac{1}{4}\int_{\mathscr{D}} \big(\Delta \log\big(|P|^{-2} e^{t|z|^2}\big)\big)|f|^2 |P|^{-2} e^{t|z|^2}.$$

Also, $\log |P|^{-2} = -\log \prod_\nu |z - z_\nu|^2 = -\sum_\nu \log |z - z_\nu|^2$, whence

$$\Delta \log\big(|P|^{-2} e^{t|z|^2}\big) = -\sum_\nu \delta(z - z_\nu) + 4t,$$

where $\delta$ is the Dirac delta, meaning that the sum above vanishes on $\mathcal{D}$. Thus,

$$\int_{\mathcal{D}} |\bar{\partial} f|^2 |P|^{-2} e^{t|z|^2} \geq t \int_{\mathcal{D}} |f|^2 |P|^{-2} e^{t|z|^2}.$$

Finally, (6.2.4) becomes the desired Carleman estimate

$$\int_{\mathcal{D}} |\Delta f|^2 |P|^{-2} e^{t|z|^2} \geq \frac{a_6}{\delta^2} \int_A |\nabla f|^2 |P|^{-2} e^{t|z|^2} + a_7 t^2 \int_{\mathcal{D}} |f|^2 |P|^{-2} e^{t|z|^2}, \tag{C1}$$

which holds for any $f \in C_0^\infty\big(\mathbb{R}^2 \setminus \bigcup_\nu D_\nu(a)\big)$, with $\mathcal{D}$ a bounded open set such that $A \subset \mathcal{D} \subset \bigcup_\nu D_\nu(a)$.

## 7. Discussion

**7.1.** ***Higher dimensions.*** In this paper, we have studied eigenfunctions of the Laplace–Beltrami operator on closed $C^\infty$ surfaces and have underlined a natural interpretation of Yau's conjecture in light of Theorem 1. Since the conjecture is expected to hold in any dimension, it is natural to ask:

**Question 7.1.1.** Does Theorem 1 hold for a compact, smooth manifold of dimension $n \geq 3$?

It seems reasonable to expect that the result holds in higher dimensions. On the one hand, as previously stated, Yau's conjecture on the size of nodal sets is formulated for manifolds of any dimension. On the other hand, some fundamental results for the growth exponents of eigenfunctions are known to hold in any dimension, most notably the Donnelly–Fefferman growth bound

$$\beta(\phi_\lambda, B) = \log \frac{\sup_B |\phi_\lambda|}{\sup_{\frac{1}{2}B} |\phi_\lambda|} \leq c\sqrt{\lambda}, \tag{7.1.2}$$

where $B$ is any metric ball (see for instance [Donnelly and Fefferman 1988; Mangoubi 2013; Nazarov et al. 2005]). However, the approach we have used relies crucially on the reduction of an eigenfunction $\phi_\lambda$ to a planar solution $F$ to a Schrödinger equation, a transformation made possible by the existence of local conformal coordinates, a fact that does not generalize in dimension $n \geq 3$. One would therefore need to follow a fundamentally different approach to prove a result in the spirit of Theorem 1 in that setting. Nazarov et al. [2005] give a simpler proof of the growth bound (7.1.2) in the setting of closed surfaces. A generalization of that proof in higher dimensions has been given by Mangoubi [2013], notably using a clever extension of eigenfunctions on an $n$-dimensional manifold $M$ to harmonic functions on the $(n+1)$-dimensional manifold $M \times \mathbb{R}$ (see also [Lin 1991; Jerison and Lebeau 1999; Nazarov et al. 2005]). We believe that a similar treatment could be useful in attempting to generalize Theorem 1.

**7.2.** ***How to measure the growth: generalization to $L^q$ norms.*** Our measure of the growth of eigenfunctions has been made through growth exponents defined on small metric disks on which we have taken the $L^\infty$ norm. Indeed, we recall that

$$\beta_p(\lambda) = \log \frac{\sup_B |\phi_\lambda|}{\sup_{\alpha_0} B |\phi_\lambda|},$$

where $B$ is a metric ball of small radius centered at $p \in M$. For $1 \leq q \leq \infty$, define the more general $q$-growth exponent $\beta_p^q(\lambda)$ of an eigenfunction $\phi_\lambda$:

$$\beta_p^q(\lambda) := \log \frac{\|\phi_\lambda\|_{L^q(B)}}{\|\phi_\lambda\|_{L^q(\alpha_0 B)}},$$

where $B$ is once again a suitably small metric ball centered at $p$. Notice that $\beta_p(\lambda) = \beta_p^\infty(\lambda)$. Consider the average of such quantities on the surface; that is, define

$$B^q(\lambda) := \frac{1}{\mathrm{Vol}(M)} \int_M \beta_p^q(\lambda)\, \mathrm{d}V_g,$$

and then ask:

**Question 7.2.1.** For which $q \in [1, \infty)$, if any, do we have the following analogue of Theorem 1:

$$c B^q(\lambda) \lambda^{1/2} \leq \mathcal{H}^1(Z_\lambda) \leq C(B^q(\lambda)\lambda^{1/2} + 1)?$$

Keeping our setting of closed surfaces, it would suffice to prove analogues of Theorems 2.1.1 and 3.1.1 for $q$-growth exponents of planar Schrödinger eigenfunctions to answer positively the last question, but there does not seem to be an obvious way to tackle this problem.

## Acknowledgements

This research is part of my PhD thesis at the Université de Montréal under the supervision of Iosif Polterovich. I am very grateful to him for suggesting the problem and for his constant support and many discussions that have been both very helpful and enjoyable. I also want to thank Dan Mangoubi for his support and useful explanations as well as Leonid Polterovich for helpful remarks. I am also grateful to Steve Zelditch for his suggestions on the exposition as well as some interesting questions. Thanks to Agathe Bray-Bourret for her help with some figures. Finally, I want to specially underline the precious help of Misha Sodin, whose contribution has been more than instrumental in the completion of this article. The main ideas used in the proof of Theorem 2.1.1 are based on the notes provided by him, and I am extremely grateful to have benefited from his support and help.

## References

[Ahlfors 1966] L. V. Ahlfors, *Lectures on quasiconformal mappings*, Math. Studies **10**, Van Nostrand, Princeton, NJ, 1966. MR 34 #336 Zbl 0138.06002

[Bers 1955] L. Bers, "Local behavior of solutions of general linear elliptic equations", *Comm. Pure Appl. Math.* **8**:4 (1955), 473–496. MR 17,743a Zbl 0066.08101

[Brüning 1978] J. Brüning, "Über Knoten von Eigenfunktionen des Laplace–Beltrami Operators", *Math. Z.* **158**:1 (1978), 15–21. MR 57 #17732 Zbl 0349.58012

[Cheng 1976] S.-Y. Cheng, "Eigenfunctions and nodal sets", *Comment. Math. Helv.* **51**:1 (1976), 43–55. MR 53 #1661 Zbl 0334.35022

[Dong 1992] R.-T. Dong, "Nodal sets of eigenfunctions on Riemann surfaces", *J. Differential Geom.* **36**:2 (1992), 493–506. MR 93h:58159 Zbl 0776.53024

[Donnelly and Fefferman 1988] H. Donnelly and C. Fefferman, "Nodal sets of eigenfunctions on Riemannian manifolds", *Invent. Math.* **93**:1 (1988), 161–183. MR 89m:58207 Zbl 0659.58047

[Donnelly and Fefferman 1990] H. Donnelly and C. Fefferman, "Nodal sets for eigenfunctions of the Laplacian on surfaces", *J. Amer. Math. Soc.* **3**:2 (1990), 333–353. MR 92d:58209 Zbl 0702.58077

[Gelfond 1934] A. Gelfond, "Über die harmonischen Funktionen", *Trav. Inst. Phys.-Math. Stekloff* **5** (1934), 149–158. Zbl 0009.17204

[Gilbarg and Trudinger 1998] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, revised 2nd ed., Grundlehren der Math. Wissenschaften **224**, Springer, Berlin, 1998. MR 2001k:35004 Zbl 1042.35002

[Han and Lin 2007] Q. Han and F.-H. Lin, "Nodal sets of solutions of elliptic differential equations", preprint, 2007, http://www3.nd.edu/~qhan/nodal.pdf.

[Hardt and Simon 1989] R. Hardt and L. Simon, "Nodal sets for solutions of elliptic equations", *J. Differential Geom.* **30**:2 (1989), 505–522. MR 90m:58031 Zbl 0692.35005

[Hug and Schneider 2002] D. Hug and R. Schneider, "Kinematic and Crofton formulae of integral geometry: recent variants and extensions", pp. 51–80 in *Homenatge al professor Lluís Santaló i Sors*, edited by C. Barceló i Vidal, Diversitas **34**, Universitat de Girona, 2002.

[Jerison and Lebeau 1999] D. Jerison and G. Lebeau, "Nodal sets of sums of eigenfunctions", pp. 223–239 in *Harmonic analysis and partial differential equations* (Chicago, 1996), edited by M. Christ et al., Univ. Chicago, 1999. MR 2001b:58035 Zbl 0946.35055

[Khovanskii and Yakovenko 1996] A. Khovanskii and S. Yakovenko, "Generalized Rolle theorem in $\mathbb{R}^n$ and $\mathbb{C}$", *J. Dynam. Control Systems* **2**:1 (1996), 103–123. MR 97f:26016 Zbl 0941.26009

[Koch and Tataru 2001] H. Koch and D. Tataru, "Carleman estimates and unique continuation for second-order elliptic equations with nonsmooth coefficients", *Comm. Pure Appl. Math.* **54**:3 (2001), 339–360. MR 2001m:35075 Zbl 1033.35025

[Lin 1991] F.-H. Lin, "Nodal sets of solutions of elliptic and parabolic equations", *Comm. Pure Appl. Math.* **44**:3 (1991), 287–308. MR 92b:58224 Zbl 0734.58045

[Mangoubi 2013] D. Mangoubi, "The effect of curvature on convexity properties of harmonic functions and eigenfunctions", *J. Lond. Math. Soc.* (2) **87**:3 (2013), 645–662. MR 3073669 Zbl 06188917

[Nazarov et al. 2005] F. Nazarov, L. Polterovich, and M. Sodin, "Sign and area in nodal geometry of Laplace eigenfunctions", *Amer. J. Math.* **127**:4 (2005), 879–910. MR 2006j:58049 Zbl 1079.58026

[Robertson 1939] M. S. Robertson, "The variation of the sign of $V$ for an analytic function $U + iV$", *Duke Math. J.* **5** (1939), 512–519. MR 1,9b Zbl 0021.41503

[Roytwarf and Yomdin 1997] N. Roytwarf and Y. Yomdin, "Bernstein classes", *Ann. Inst. Fourier* (*Grenoble*) **47**:3 (1997), 825–858. MR 98h:34009a Zbl 0974.30524

[Yau 1982] S.-T. Yau, "Survey on partial differential equations in differential geometry", pp. 3–71 in *Seminar on differential geometry*, Ann. of Math. Stud. **102**, Princeton Univ., 1982. MR 83i:53003 Zbl 0478.53001

[Yau 1993] S.-T. Yau, "Open problems in geometry", pp. 1–28 in *Differential geometry, 1: Partial differential equations on manifolds* (Los Angeles, 1990), edited by R. Greene and S.-T. Yau, Proc. Sympos. Pure Math. **54**, Amer. Math. Soc., Providence, RI, 1993. MR 94k:53001 Zbl 0801.53001

[Zelditch 2013] S. Zelditch, "Eigenfunctions and nodal sets", pp. 237–308 in *Surveys in differential geometry: geometry and topology*, edited by H.-D. Cao and S.-T. Yau, Surv. Differ. Geom. **18**, International Press, 2013. MR 3087922 Zbl 06296849

Received 6 Sep 2014. Accepted 26 Nov 2014.

GUILLAUME ROY-FORTIN: groyfortin@dms.umontreal.ca
*Département de mathématiques et de statistique, Université de Montréal, CP 6128 succ., Centre-Ville, Montréal, QB H3C 3J7, Canada*

## Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at msp.org/apde.

**Originality**. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

**Language**. Articles in APDE are usually in English, but articles written in other languages are welcome.

**Required items**. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

**Format**. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

**References**. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

**Figures**. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

**White space**. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

**Proofs**. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

ANALYSIS & PDE

Volume 8 No. 1 2015



2157-5045(2015)8:1;1-D