Non-Euclidean Singular Integrals

and

the Porous Medium Equation

Contents

Chapter 1

Introduction

This work is concerned with non-Euclidean singular integrals and the large time regularity of solutions to the porous medium equation.

It is hard to overrate the importance of singular integral operators. This has two reasons: there are many such operators and there is an elaborate theory of them with connections to topics ranging from complex analysis over probability to numerical analysis. It is no surprise that the theory of singular integrals has been developed into various directions, among them pseudo-differential operators, sharp criteria for boundedness (T(1) theorems), the theory of Muckenhoupt weights, application to regularity in Lipschitz domains and, closer to our interest, singular integrals on non-Euclidean spaces.

The second part of the title is concerned with the porous medium equation. The motion of an ideal gas in a porous medium can be modeled by the *porous medium equation* (with normalized constants)

$$\rho_t - \Delta \rho^m = 0 \quad \text{ in } \mathbb{R}^n \times \mathbb{R}^+$$

where ρ is the density of the gas and $m > 1$ is a real number. In contrast to parabolic problems the porous medium equation admits compactly supported solutions. This is its most striking feature.

We show under weak assumptions that the boundary of the support of ρ is a smooth hyper-surface in space time for large t. The problem is transformed to a quasi-linear sub-elliptic problem, for which we establish an analogue of the regularity theory for quasi-linear parabolic equations. This requires the theory of singular integral operators in a non-Euclidean setting and Gaussian estimates of the fundamental solution of a linear degenerate parabolic problem. The study of both, Gaussian estimates and singular integrals, strongly relates geometric and analytic properties.

On a formal level the pressure $p = \frac{m}{m-1}\rho^{m-1}$ satisfies $p_t = (m-1)p\Delta p + |\nabla p|^2$ which suggests that the velocity in normal direction of the boundary of the support is $v = -\frac{p_t}{\partial_\nu p} = -\partial_\nu p$ where ν is the exterior normal.

The behavior near a point of the boundary of its support depends crucially on the concentration of the gas near that point. This, together with some control of level surfaces, is an important local property termed non-degeneracy.

Definition 1 (Non-degeneracy). The density ρ is *non-degenerate* in the set $Q(x_0, t_0, R) := \{(x,t) : |x - x_0| < R, t_0 - R < t < t_0\}$ if there exists $\delta > 0$ such that

$$\rho(x,t)^{m-1} \geq \delta d(x, (\text{ supp } \rho(t))^0)$$

for $(x,t) \in Q(x_0, t_0, R)$. Here ^{0} denotes the complement of the set.

Non-degeneracy can be ensured by a condition on the initial data ρ_0: if the support Ω of ρ_0 is compact, $\rho_0 \in C^{1,\alpha}(\Omega)$ and

$$|\nabla \rho_0^{m-1}| + \rho_0^{m-1} \geq \kappa > 0$$

in Ω then ρ_0 is *α-admissible*, a notion which will be defined in Chapter 5. Suppose that ρ_0 is α-admissible. There exists t_0 such that for all (x,t) in the boundary of the support of ρ with $t > t_0$ the following is true: there exists a relatively open neighborhood $U \subset$ supp ρ of (x,t) (which contains part of the boundary of the support), a unit vector ν and $\delta > 0$ such that the pressure p is Lipschitz continuous in U, C^2 in the interior of U and $\nu \cdot \nabla p \leq -\delta$ in the interior of U. In particular the solution is non-degenerate there. In that situation one may make a change of dependent and independent variables to transform the problem to one with fixed boundary

$$v_t - \sum_{i=1}^{n-1} x_n \partial_{x_i}^2 v + x_n^{-\sigma} \partial_{x_n} \left(x_n^{1+\sigma} \frac{1 + \sum_{i=1}^{n-1} |\partial_i v|^2}{\partial_{x_n} v} \right) = 0$$

in a relatively open subset of the half-space $\{(x,t) : x_n \geq 0\}$ where again v is Lipschitz continuous and differentiable away from $x_n = 0$.

The linearization in the simplest case leads to the operator

$$v \to Lv = v_t - \sum_{i=1}^{n} \partial_{x_i}(x_n \partial_{x_i} v) - \sigma \partial_{x_n} v.$$

The spatial part of the operator defines a Riemannian metric on $H = \{x \in \mathbb{R}^n : x_n > 0\}$ by $g_x(v,w) = x_n^{-1} v \cdot w$. The geodesics can be computed. The Riemannian metric induces a metric d on the closed half space. The measure $\mu_\sigma = x_n^\sigma dx$ is naturally related to L. This is the geometric data occuring in the estimates of the fundamental solution. Derivatives of the fundamental solution define singular integral operators in a non-Euclidean context. To be more precise: The theory of singular integrals and the theory of Muckenhoupt weights are both available in the geometric setting given by the triple (H, d, μ_σ). The Aronson-type estimates established in Chapter 4 allow to apply the theory of singular integrals to give

Theorem 2. *Suppose that $\sigma > -1$ and $-1 < \rho < p - 1 + p\sigma$. Then*

$$\|v_t\|_{L^p(\mu_\rho)} + \|\nabla v\|_{L^p(\mu_\rho)} + \|x_n D_x^2 v\|_{L^p(\mu_\rho)} \leq c\|Lv\|_{L^p(\mu_\rho)}.$$

These estimates do not hold for $\sigma \leq -1$ or $\rho \leq -1$ or $\rho \geq p - 1 + p\sigma$.

Remark 3. Even the case $\sigma = \rho = 0$ and $p = 2$ is far from being obvious. These estimates should be considered as weighted estimates with the 'unweighted' case $\rho = \sigma$.

Among the consequences of the Aronson-type estimates is part of the following:

Theorem 4. *Let U be a relatively open subset of the support of ρ. Suppose that $\rho^{m-1} \in C^{0,1}(U) \cap C^1(\overset{o}{U})$ satisfies $\partial_{x_n} \rho^{m-1} < -\delta < 0$ in the interior $\overset{o}{U} = U \cap \{\rho > 0\}$. Then the following is true.*
1. The interface in U, $\Sigma \cap U$, is locally given by $x_n = h(x_1, \ldots, x_{n-1}, t)$ where $h \in C^\infty_{loc}$. Let K' be a compact subset of the domain of h. Then there exists R and C such that

$$\sup_{K', \alpha} \alpha!^{-1} \alpha_0!^{-1} R^{|\alpha|} |\partial^\alpha_{x,t} h| \leq C$$

where the supremum is taken over all multiindices $\alpha = (\alpha_1, \ldots, \alpha_{n-1}, \alpha_0)$.
2. The pressure p is smooth in U up to the boundary of the support. Let K be a compact subset of U. Then there exists R and C such that

$$\sup_{(x,t) \in K, \alpha} \alpha!^{-1} \alpha_0!^{-1} R^{|\alpha|} |\partial^\alpha_{x,t} p| \leq C.$$

This implies that these solutions are analytic in x up to the boundary and they are Gevrey regular in t. Moreover the boundary of the support determines the solution. The assumptions are satisfied for large t if $\rho(.,0)$ is compactly supported and α-admissible. There is a partial converse:

Theorem 5. *Suppose that Σ_0 is a smooth hyper-surface which is locally given as the graph of a function h by $x_n = h(x_1, \ldots, x_{n-1}, t)$. Suppose that h satisfies the estimates of Theorem 4 uniformly on compact subsets of its domain of definition and $h_t > 0$. Then there exists a neighborhood and a unique local solution to the porous medium equation which has Σ_0 locally as boundary of its support.*

The main contributions of this work are large-time regularity for the porous medium equation, the Aronson type estimates and the connection to singular integrals.

Chapter 2 contains material on singular integrals and weighted norm estimates. This topic is well-developed in the literature, even if it is not a simple task to collect all we need from the literature. More or less all results and proofs here are contained in the excellent treatise of Stein [125]. In Chapter 3 this theory is applied to different operators in partial differential equations: elliptic operators with drift, the Oseen operator, elliptic problems including homogenization and certain sub-elliptic operators.

In Chapter 4 the Aronson type estimates are established and their consequences explored. We adopt the argument of Nash [108] in the version of Fabes and Stroock [53] to the problem at hand since the lifting procedure (Rothschild and Stein [118]) requires more regularity and probabilistic arguments (Kusuoka and Stroock [86], [87], [88] and [89]) seem not to give what we need. In the second part we use arguments of Grüter and Widman [67] to obtain Schauder estimates.

In Chapter 5 we develop the theory of the porous medium equation. We collect certain results from the literature. The first part relies, among other things, heavily on the use of a comparison principle to obtain Lipschitz continuity of various quantities and non-degeneracy. Here we follow Caffarelli, Vazquez and Wolanski [30]. We need delicate geometric constructions before the problem is in form accessible by the tools and estimates in Chapter 2 and Chapter 4. The Aronson type estimates (and their consequences) are crucial to bridge the gap between weak and smooth solutions to the porous medium equation.

Finally we control the growth of derivatives in Chapter 6, which gives analyticity and a Cauchy-Kowalevski theorem, Theorem 5. The result is surprising as it shows that even at the boundary transversal derivatives behave much better than time derivatives. There is a delicate interplay between degeneracy of the equation and the commutators of the operator with derivatives. The regularity result is likely to be a nonlinear version of similar results for operators on strongly pseudo-convex domains in $\mathbb{C}^n$ (see the notes of Chapter 5).

It is a pleasure to thank colleagues, coworkers, friends and my family, without whom this work would not have been possible and in particular Prof. Dr. W. Jäger for his strong support.

Chapter 2

Singular Integrals

2.1 Introduction

The Schwarz kernel theorem states that rather general linear operators have distributional kernels.

The kernels of many operators in analysis have a very particular structure: they are approximately translation invariant, and they are smooth off the diagonal. This corresponds to the fact that for a large class of differential operators L, called hypo-elliptic operators, u is smooth if Lu is smooth.

The theory of singular integrals provides tools for estimating these operators in various ways. After a study of the basic (and important) Euclidean case this theory developed considerably in different directions.

One of these is the theory of weights. It studies the basic problem: Given an operator T (linear or sub-linear). Characterize the class of functions ω such that

$$\int |Tf|^p \omega \, dx \leq c \int |f|^p \omega \, dx.$$

For many operators there is the same class (called A_p) for which this estimate holds if and only if $\omega \in A_p$.

There is an elaborate theory of operators which satisfy the Euclidean kernel estimates, but where the L^p bound is not obvious for any p. The basic result here is the T(1) theorem of David and Journe [49] with applications to the study of elliptic equations in Lipschitz domains (see Kenig [79] and Dahlberg and Kenig [74] for related results and references) and several other problems.

The theory of pseudo-differential operators elaborates on another aspect of singular integral operators: it provides a calculus for computing composition operators and connects the theory closer to the Fourier transform.

The theory of singular integral operators does not rely on the Euclidean structure. An examination of the proofs shows that (Whitney) covering lemmata and doubling properties of the measures are essential ingredients.

The same is true for other basic results, like the John-Nirenberg inequality, Moser's generalization of this inequality to parabolic equations, the construction of Hardy spaces and BMO, and, closer related to our needs, the theory of weights.

The motivation for considering more general metrics comes from the occurrence of such operators in complex analysis, nilpotent Lie-groups, probability, and squares of vector fields.

We follow the standard outline: estimate of the Hardy Littlewood maximal function using the Vitali covering lemma, estimate of singular integrals using the Whitney covering lemma through the Calderón-Zygmund decomposition. Then we deviate somehow and provide estimates for a maximal operator built on truncated singular integrals. This is needed for the weighted estimates.

We introduce the Muckenhoupt condition, prove weak estimates for the Hardy-Littlewood maximal function, then the reverse Hölder inequality which implies some openness of the A_p condition and hence strong weighted estimates for the Hardy-Littlewood maximal function. Weighted estimates for singular integrals are obtained via a good λ inequality or relative distributional inequality and the estimate for the maximal operator.

This chapter is basic for the sequel. The following results are used in the sequel: the boundedness of singular integral operators, Theorem 2.2.8, boundedness of truncated operators, Corollary 2.2.12 and the weighted estimates of Corollary 2.3.9. The criteria for Muckenhoupt weights Proposition 2.3.11 and Proposition 2.3.12 are not proven. They are applied in Chapter 3, but they are not needed for the porous medium equation.

2.2 Estimates for singular integral operators

Let (Ω, d) be a *complete metric space* with metric d, equipped with the topology generated by all balls $B(x,\rho) = \{y : d(x,y) < \rho\}$ (in other words: a subset A is open if for every element x there is a ball $B(x,\rho) \subset A$). There may be sets besides the empty set and the full set which are simultaneously open and closed: Ω is not required to be connected, not even locally.

Let μ be a nontrivial *Borel regular doubling measure*. Borel regularity means that $d(.,x)$ is measurable and for every open set A we have $\mu(A) = \sup \mu(K)$ where the supremum is taken over all compact subsets K of A. We call a measure *doubling* if there exists a constant $b > 0$ with

$$\mu(B(x,3\rho)) \leq b\mu(B(x,\rho)).$$

The constant b will be used in the whole chapter. It is a consequence that there is a dense countable subset in Ω. The spaces and metrics we have in mind include the following:

1. $\Omega = \mathbb{R}^n$, d the Euclidean distance, where the Lebesgue measure is a natural measure to consider.
2. $\Omega \subset \mathbb{R}^n$ is a bounded open domain, d is the Euclidean distance.
3. $\Omega = \mathbb{Z}^n$, d is the Euclidean distance. Here the counting measure occurs naturally.
4. $\Omega \subset \mathbb{Z}^n$, d is the Euclidean distance.
5. $\Omega = \mathbb{R}^{n-1} \times \mathbb{R}^+$, $d(x,y) = \frac{|x-y|}{\sqrt{|x-y|+x_n+y_n}}$. Interesting measures are $\mu = x_n^\sigma dx$ with $\sigma > -1$.

Here 1. is the standard setting for singular integrals, 2. occurs in the connection of elliptic boundary value problems, 3. and 4. occur in discrete problems and in homogenization, 5. occurs in the study of the porous medium equation. This list is by no means complete and one should at least mention additionally nilpotent groups and the boundary of strongly pseudo-convex domains with their natural structures.

There are two associated Hardy-Littlewood maximal functions.

$$Mf(x) := \sup_\rho \mu(B(x,\rho))^{-1} \int_{B(x,\rho)} |f| dx$$

and the larger one

$$\tilde{M}f(x) := \sup_{B(y,\rho)\ni x} \mu(B(y,\rho))^{-1} \int_{B(y,\rho)} |f| dx.$$

It is clear from the definition that $\{x \in \Omega : \tilde{M}f(x) > \lambda\}$ is open for all λ. Hence $\tilde{M}f$ is measurable. In our general context $\tilde{M}f$ is the more useful maximal function of the two. The basic result about maximal functions is the following

Theorem 2.2.1. *Let f be a function defined in Ω.*

1. *If $f \in L^p(\mu)$ for one $1 \leq p \leq \infty$, then $\tilde{M}f$ is finite μ-almost everywhere.*

2. *If $f \in L^1(\mu)$ then, for $\lambda > 0$,*

$$\mu(\{x : \tilde{M}f(x) > \lambda\}) \leq \frac{b}{\lambda} \int |f| d\mu. \tag{2.1}$$

3. *If $1 < p \leq \infty$ and $f \in L^p(\mu)$ then*

$$\|\tilde{M}f\|_p \leq 2\left(\frac{bp}{p-1}\right)^{1/p} \|f\|_p.$$

Proof. It is clear that $\tilde{M}f(x) \leq \|f\|_\infty$ and $\tilde{M}(f+g)(x) \leq \tilde{M}f(x) + \tilde{M}g(x)$. Suppose that (2.1) holds. Then

$$\mu(\{\tilde{M}(f) > \lambda\}) \leq \frac{2b}{\lambda} \int (|f| - \lambda/2)_+ d\mu(x)$$

and

$$\begin{aligned}
\|\tilde{M}f\|_p^p &= p \int_0^\infty \mu(\{\tilde{M}f(x) > \lambda\}) \lambda^{p-1}\, d\lambda \\
&\leq 2bp \int_0^\infty \lambda^{p-2} \int_{\{x:|f|>\lambda/2\}} |f| d\mu(x)\, d\lambda \\
&\leq 2bp \int_\Omega \int_0^{2|f|} \lambda^{p-2} d\lambda |f|\, d\mu(x) \\
&\leq \frac{2^p bp}{p-1} \|f\|_p^p
\end{aligned}$$

by a multiple application of Fubini's Theorem to the characteristic function of the set $A_f = \{(x,t) \in \Omega \times \mathbb{R}^+ : 0 \leq t < f(x)\}$ resp. to $A_{\tilde{M}f}$. The proof of the weak inequality (2.1) requires

Lemma 2.2.2 (Vitali). *Let E be a finite union of balls B_i. Then there exists a disjoint sub-collection B_{i_j} with*

$$\sum \mu(B_{i_j}) \geq b^{-1} \mu(E). \tag{2.2}$$

Proof. We order the maximal disjoint sub-collections by the lexigraphical order of their radii. We take a maximal sub-collection B_{i_j} with respect to this ordering. We claim that $E \subset \bigcup B(x_{i_j}, 3r_{i_j})$, which implies the estimate. Let $x \in E$. It is contained in a ball B_i. This ball is either contained in the sub-collection, or it has nonempty intersection with a ball B_{i_j} with radius $r_{i_j} \geq r_i$ of the sub-collection by the maximality of the sub-collection. By the triangle inequality $B_i \subset B_{3r_{i_j}}(x_{i_j})$. This completes the proof. $\square$

We define the open set $E_\lambda = \{x : \tilde{M}f(x) > \lambda\}$. Let $E \subset E_\lambda$ be a compact subset. For $x \in E$ there exists a ball B_x containing x such that, by the definition of E_λ,

$$\mu(B_x) \leq \frac{1}{\lambda} \int_{B_x} |f(y)| d\mu.$$

This ball is contained in E_λ by the definition of $\tilde{M}$. There exists a finite sub-collection covering E. Let $\tilde{E}$ be their union. Clearly $E \subset \tilde{E} \subset E_\lambda$. There exists a disjoint sub-collection with property (2.2). Then

$$\mu(\tilde{E}) \leq b \sum_j \mu(B_{i_j}) \leq \frac{b}{\lambda} \int_{\tilde{E}} |f| d\mu(x).$$

We arrive at the conclusion if we take the supremum over all compact subsets E of E_λ. □

Corollary 2.2.3. *Suppose that $f \in L^1_{loc}$. Then*

$$\begin{aligned}\lim_{r\to 0} \sup_{B(y,r)\ni x} \mu(B(y,r))^{-1} &\int_{B(y,r)} f(z)\, d\mu(z) \\ &= \lim_{r\to 0} \inf_{B(y,r)\ni x} \mu(B(y,r))^{-1} \int_{B(y,r)} f(z)\, d\mu(z) \\ &= f(x)\end{aligned}$$

μ–almost everywhere.

Proof. We may assume that f is supported in a fixed ball. Continuous functions with compact support are dense in L^1. For continuous functions with compact support the statement is clear. Let

$$\begin{aligned}Tf(x) = &\lim_{r\to 0} \sup_{B(y,r)\ni x} \mu(B(y,r))^{-1} \int_{B(y,r)} f(z)\, d\mu(z) \\ &- \lim_{r\to 0} \inf_{B(y,r)\ni x} \mu(B(y,r))^{-1} \int_{B(y,r)} f(z)\, d\mu(z).\end{aligned}$$

Then $Tf = 0$ if f is continuous. Otherwise, given ε there exists a continuous function g with $\int |f-g| d\mu \leq \delta$. Then

$$Tf(x) \leq Tg(x) + T(f-g)(x) = T(f-g)(x)$$

hence

$$\mu(\{Tf(x) > \lambda\}) \leq \mu(\{\tilde{M}(f-g) > \lambda/2\}) \leq \frac{2b\delta}{\lambda} \|f\|_{L^1}.$$

We let δ tend to zero and obtain $Tf(x) \leq \lambda$ almost everywhere. The positive number λ is arbitrary, hence the result. □

We need more refined covering results for the treatment of singular integral operators.

Lemma 2.2.4 (Whitney). *Given an open set $A \subset \Omega$, $A \neq \Omega$, there exists a collection of balls $B(x_i, r_i)$ such that:*

1. *The balls are pairwise disjoint.*
2. $A = \cup_i B(x_i, 3r_i)$
3. *Let $F = A^0$ be the complement of A. Then $d(x_i, F) = 3r_i$.*

Remark 2.2.5. In $\mathbb{R}^n$ there is a stronger version of the Whitney covering lemma: open sets may be decomposed into closed cubes with disjoint interiors with distance to the complement less than a fixed number times their diameter. There is an immediate substitute:

We define inductively, using the sequence of balls of Lemma 2.2.4,

$$Q_k = B(x_k, d(x_k, F)) \cap \Big(\bigcup_{j<k} Q_j\Big)^0 \cap \Big(\bigcup_{k>j} B(x_k, d(x_k, F)/3))\Big)^0.$$

These sets provide a substitute for cubes, i.e. they provide a disjoint covering of the open set A with diameter comparable to the distance to the complement since

$$B(x_j, d(x_j, F)/3) \subset Q_j \subset B(x_j, d(x_j, F)).$$

Remark 2.2.6. The Whitney covering lemma is only related to the metric of Ω, not to the measure.

Proof. The balls $B(x, d(x, F)/3)$ with $x \in A$ cover A. We choose a maximal pairwise disjoint sub-collection of these balls centered at the points x_j. It remains to prove that

$$\bigcup B(x_j, d(x_j, F)) = A.$$

Let $x \in A$. There exists x_j such that $B(x, d(x, F)/3) \cap B(x_j, d(x_j, F)/3) \neq \{\}$ by construction. If $d(x, F) < 2d(x_j, F)$ then $d(x, x_j) < \frac{1}{3}d(x, F) + \frac{1}{3}d(x_j, F) < d(x_j, F)$ and $x \in B(x_j, d(x_j, F))$. Otherwise, if $d(x, F) \geq 2d(x_j, F)$ then

$$\overline{B(x_j, d(x_j, F))} \subset B(x, d(x, F)/3 + 4d(x_j, F)/3) \subset B(x, d(x, F))$$

which is impossible since the set on the lefthand side contains at least one point of F. $\square$

The major step towards estimates of singular integral operators is the Calderón-Zygmund decomposition, which decomposes L^1 functions into simpler parts: into a sum of so-called atoms, with mean zero and controlled size and support, plus a bounded function.

Proposition 2.2.7. *Let $f \in L^1(\mu)$ be given and $\lambda > \frac{1}{\mu(\Omega)} \int_\Omega |f| d\mu$. Then there exists a decomposition $f = g + b$, with $b = \sum_k b_k$ and a sequence of balls B_k so that:*

1. *$|g| \le b^2 \lambda$ almost everywhere.*
2. *Each b_k is supported in B_k and*
$$\int |b_k| d\mu \le 2b^2 \lambda \mu(B_k) \text{ and } \int b_k \, d\mu = 0.$$
3. *The measure of the balls is controlled by*
$$\sum_k \mu(B_k) \le \frac{b^2}{\lambda} \int |f| \, d\mu.$$

Proof. Let $E_\lambda = \{x : \tilde{M} f(x) > \lambda\}$ and let $F = E_\lambda^0$ be its complement. We choose a sequence of points x_j as in Lemma 2.2.4. Let Q_j be the disjoint covering of E_λ as in Remark 2.2.5. We define

$$g(x) = \begin{cases} f(x) & \text{if } x \notin E_\lambda \\ \frac{1}{\mu(Q_k)} \int_{Q_k} f d\mu & \text{if } x \in Q_k \end{cases}$$

and

$$b_k = \chi_{Q_k} \left(f(x) - \frac{1}{\mu(Q_k)} \int_{Q_k} f(y) d\mu(y) \right).$$

Then $|g| \le \lambda$ almost everywhere in F because of Corollary 2.2.3. By construction $B(x_j, 3d(x_j, F))$ contains a point of F hence

$$\int_{B(x_j, 3d(x_j, F))} |f| \, d\mu \le \lambda \mu(B(x_j, 3d(x_j, F)))$$

thus

$$\frac{1}{\mu(Q_j)} \int_{Q_j} |f| \le b^2 \lambda$$

and hence $|g| \leq b^2\lambda$ in E_λ. Now with $B_j = B(x_j, d(x_j, F))$,

$$\mu(B_j)^{-1} \int |b_j| d\mu \leq 2\mu(Q_j)^{-1} \int_{Q_j} |f(x)| d\mu \leq 2b^2\lambda.$$

This is the second part of the theorem. For the third part we observe that

$$\sum \mu(B(x_k, d(x_k, F))) \leq b \sum \mu(B(x_k, d(x_k, F)/3)) \leq b\mu(E_\lambda).$$

For every $x \in E_\lambda$ there exists a ball $B \subset E_\lambda$ with $x \in B$ and $\int_B |f| > \frac{\lambda}{b^2}|B|$. We apply the Vitali covering lemma, more precisely (2.2), in the same way as for the proof of Theorem 2.2.1, to get

$$\mu(E_\lambda) \leq \frac{b}{\lambda} \int_{E_\lambda} |f| \, d\mu. \tag{2.3}$$

This completes the proof. $\square$

We study singular integral operators in the following setting. Let $1 < q \leq \infty$ and $T : L^q \to L^q$ be a linear operator with

$$\|Tf\|_{L^q(\mu)} \leq a_q \|f\|_{L^q(\mu)}, \tag{2.4}$$

for all $f \in L^q(\mu)$ and a measurable kernel function K such that $x \to K(x, y)$ is locally integrable in $\Omega \backslash \{y\}$ for all x and $y \to K(x, y)$ is locally integrable in $\Omega \backslash \{x\}$ for all y. Thus

$$Tf(x) = \int K(x, y) f(y) d\mu(y)$$

for $f \in L^\infty(\mu)$ with bounded support and all x in the complement of the support of f. Moreover we assume

$$\sup_{y, \bar{y} \in \Omega, y \neq \bar{y}} \int_{\{y : d(x,y)| \geq 3d(y, \bar{y})\}} |K(x, y) - K(x, \bar{y})| \, d\mu(x) \leq a. \tag{2.5}$$

Theorem 2.2.8. *1. Suppose that $f \in L^1(\mu)$. Let $d = (4a_q b^2)^q b^2 + 4ab^4 + 2b^3$. Then, for all $\lambda > 0$,*

$$\mu(|Tf(x)| > \lambda) \leq \frac{d}{\lambda} \|f\|_{L^1(\mu)}. \tag{2.6}$$

2. If $1 < p \leq q$ then T extends to a bounded operator on $L^p(\mu)$ and

$$\|Tf\|_{L^p(\mu)} \leq \left(\frac{2dp}{p-1} + \frac{2^q a_q^q q}{q-p} \right)^{1/p} \|f\|_{L^p(\mu)}.$$

Proof. Suppose that (2.6) holds and that $1 < p < q$. We define

$$f^\lambda(x) := \begin{cases} \frac{\lambda f(x)}{|f(x)|} & \text{if } f(x) \geq \lambda \\ f(x) & \text{if } f(x) \leq \lambda \end{cases}$$

and $f_\lambda = f - f^\lambda$. Then

$$\begin{aligned} \|Tf\|^p_{L^p(\mu)} &= p \int_0^\infty \mu(\{x : Tf(x) > \lambda\})\lambda^{p-1}\, d\lambda \\ &\leq p \int_0^\infty \left(\mu(\{Tf_\lambda > \lambda/2\}) + \mu(\{Tf^\lambda > \lambda/2\})\right) \lambda^{p-1}\, d\lambda \\ &\leq 2dp \int_0^\infty \int_{|f|\geq\lambda} |f| d\mu\, \lambda^{p-2}\, d\lambda + 2^q p \int_0^\infty \|Tf^\lambda\|^q_{L^q(\mu)} \lambda^{p-1-q}\, d\lambda \\ &\leq \frac{2dp}{p-1}\|f\|^p_{L^p(\mu)} + 2^q a_q^q p \int_0^\infty \|f^\lambda\|^q_{L^q(\mu)}\, \lambda^{p-1-q}\, d\lambda \\ &= \frac{2dp}{p-1}\|f\|^p_{L^p(\mu)} + 2^q a_q^q pq \int_0^\infty \int_0^\lambda \mu(\{|f| > \rho\})\rho^{q-1}\, d\rho \lambda^{p-1-q}\, d\lambda \\ &= \left(\frac{2dp}{p-1} + \frac{2^q a_q^q q}{q-p}\right) \|f\|^p_{L^p(\mu)} \end{aligned}$$

where we used subadditivity in the first inequality, (2.6) and Tschebyscheff's inequality in the second, Fubini and (2.4) in the third inequality and again Fubini's theorem in the last two inequalities.

It remains to prove (2.6). Let $\lambda > 0$ and let $f = g + b$ be the Calderon-Zygmund decomposition with respect to $\lambda/2$. Then it suffices to show

$$\mu(\{x : |Tg| > \lambda/2\}) \leq \frac{(4a_q b^2)^q b^2)}{\lambda}\|f\|_{L^1(\mu)} \tag{2.7}$$

and

$$\mu(\{x : |Tb| > \lambda/2\}) \leq \frac{4ab^4 + 2b^3}{\lambda}\|f\|_{L^1(\mu)}. \tag{2.8}$$

Now, using estimate (2.3),

$$\begin{aligned} \int |g|^q d\mu &= \int_{E_{\lambda/2}} |g|^q d\mu + \int_{E^0_{\lambda/2}} |g|^q d\mu \\ &\leq (b^2\lambda/2)^q \mu(E_{\lambda/2}) + \lambda^{q-1}\|f\|_{L^1(E^0_{\lambda/2})} \\ &\leq (2^{q-1}b^{2q+2} + 1)\lambda^{q-1}\|f\|_{L^1(\mu)} \end{aligned}$$

hence, by Tschebyscheff's inequality

$$\mu\{x : |Tg|(x) \geq \lambda/2\} \leq \frac{2^q}{\lambda^q}\|Tg\|^q_{L^q(\mu)} \leq \frac{2^q a_q^q}{\lambda^q}\|g\|^q_{L^q(\mu)} \leq \frac{(4a_q b^2)^q b^2}{\lambda}\|f\|_{L^1(\mu)}.$$

This is estimate (2.7). It remains to estimate Tb. This again consists of two estimates: one in $E_{\lambda/2}$ and one in $E^0_{\lambda/2}$. Let B^*_k the ball with the same center but three times the radius as B_k. As a consequence of (2.5) we have $\int_{(B^*_k)^0} |Tb_k|\, d\mu \leq a\|b_k\|_{L^1}$. This implies by Proposition 2.2.7

$$\int_{(\cup B^*_k)^0} |Tb|\, d\mu \leq a \sum \|b_k\|_{L^1(\mu)} \leq 2ab^4\|f\|_{L^1(\mu)}$$

and $\mu\left(\cup B^*_k\right) \leq \frac{2b^3}{\lambda}\|f\|_{L^1(\mu)}$, which implies by Tschebyscheffs inequality (2.8). The proof is complete. $\square$

Theorem 2.2.8 is the typical estimate for singular integral operators. Its assumptions are reasonably weak. Many questions remain however open. For example: what happens when we regularize the singular integral? Can one get more local estimates?

To answer these questions we strengthen the assumptions on the kernel K. Let

$$V(x, y) = \mu(B(x, d(x, y))) + \mu(B(y, d(x, y)).$$

To estimate truncated operators we shall use the assumption

$$K(x, y) \leq \frac{a}{V(x, y)}. \tag{2.9}$$

We define the truncated operator (first for compactly supported integrable functions)

$$T_\varepsilon f(x) = \int_{d(x,y)>\varepsilon} K(x, y) f(y) d\mu(y).$$

Theorem 2.2.9. *Suppose* $T : L^q(\mu) \to L^q(\mu)$ *with*

$$\|Tf\|_{L^q(\mu)} \leq a_q\|f\|_{L^q(\mu)}$$

has a kernel which satisfies (2.9). *Then*

$$\|T_\varepsilon f\|_q \leq (a_q + ab^2)b^3\|f\|_q.$$

Proof. We fix $\bar{x}$. Let $E_\lambda(\bar{x}) = \{y : V(\bar{x}, y) < \lambda\}$ and $R = 2\inf\{r : \mu(B(\bar{x}, r)) \geq b\lambda\}$. Then $\mu(B(\bar{x}, R/3)) \geq \lambda$, hence $E_\lambda \subset B(\bar{x}, R)$ and $\mu(E_\lambda) \leq \mu(B(\bar{x}, R)) \leq b^2\lambda$. Thus

$$\|K_\varepsilon(\bar{x}, .)\|^p_{L^p(\mu)} \leq p \int_0^{\sup|K_\varepsilon|} \mu(\{K(\bar{x}, .) > \lambda\})\lambda^{p-1} d\lambda \leq \frac{apb^2}{p-1} \sup|K_\varepsilon|^p$$

and T_ε is well-defined on $L^q(\mu)$ for $1 \leq q < \infty$.

We define $\tilde{T}_\varepsilon = T - T_\varepsilon$ on $L^p(\mu)$. Let χ_ε be the characteristic function of the ball around $\bar{x}$ with radius ε. We claim that

$$\|\chi_{\varepsilon/3}\tilde{T}_\varepsilon f\|_{L^q(\mu)} \leq (a_q + ab^2)\|\chi_{\frac{4\varepsilon}{3}} f\|_{L^q(\mu)}. \tag{2.10}$$

The result follows then by a covering argument as follows. Let K be a compact subset. Every x is contained in a ball of radius $\varepsilon/9$. By the proof of the Vitali covering lemma, Lemma 2.2.2 there exists a disjoint sub-collection $B(x_j, \varepsilon/9)$ such that the balls $B(x_j, \varepsilon/3)$ cover K. Let x be a point which lies in N balls of the type $B(x_j, \frac{4\varepsilon}{3})$. All these balls are contained in $B(x, 3\varepsilon)$. In particular the N balls with same radius $\varepsilon/9$ are contained in $B(x, 3\varepsilon)$. There exists one whose measure is minimal, let say with index j_0. Then

$$N\mu(B(x_{j_0}, \varepsilon/9)) \leq \mu(B(x, 3\varepsilon)) \leq b^3\mu(B(x_{j_0}, \varepsilon/9))$$

thus $N \leq b^3$ and

$$\begin{aligned} \|T_\varepsilon f\|^q_{L^q(K)} &\leq \sum \|T_\varepsilon f\|^q_{L^q(B(x_j, \varepsilon/3))} \\ &\leq (a_q + ab^2) \sum \|f\|^q_{L^q(B(x_j, 4\varepsilon/3))} \\ &\leq (a_q + ab^2) b^3 \|f\|^q_{L^q(\mu)}. \end{aligned}$$

It remains to prove estimate (2.10). Let χ_r be the characteristic function of the ball $B(\bar{x}, r)$. By construction $\chi_{\varepsilon/3}\tilde{T}_\varepsilon f = \chi_{\varepsilon/3}\tilde{T}_\varepsilon \chi_{4\varepsilon/3} f$. Because of the support of K_ε we have $\chi_{\varepsilon/3}\tilde{T}_\varepsilon\chi_{2\varepsilon/3} = \chi_{\varepsilon/3} T \chi_{2\varepsilon/3}$ and thus

$$\|\chi_{\varepsilon/3}\tilde{T}_\varepsilon\chi_{2\varepsilon/3} f\|_{L^q(\mu)} = \|\chi_{\varepsilon/3} T \chi_{2\varepsilon/3} f\|_{L^q(\mu)} \leq a_q \|\chi_{2\varepsilon/3} f\|_{L^q(\mu)}.$$

It remains to estimate $\tilde{T}_\varepsilon(\chi_{4\varepsilon/3} - \chi_{2\varepsilon/3})f$ which is given by

$$\int_{B(x,\varepsilon)\cap B(\bar{x}, 4\varepsilon/3)\setminus B(\bar{x}, 2\varepsilon/3)} K(x, y) f(y) d\mu(y).$$

Now, for $|x - \bar{x}| \leq \varepsilon/3$, using (2.9),

$$\int_{2\varepsilon/3 \leq |\bar{x} - y| \leq 4\varepsilon/3} |K(x, y)| d\mu(y) \leq ab^2$$

and, for $2\varepsilon/3 \le |\bar{x} - y| \le 4\varepsilon/3$,

$$\int_{|x-\bar{x}|\le\varepsilon/3} |K(x,y)|dy \le ab^2$$

which gives altogether (2.10). □

We are mainly interested in the following consequence.

Corollary 2.2.10. *Suppose the assumptions of Theorem 2.2.9 are satisfied. Then there exists a sequence $\varepsilon_i \to 0$ and $v \in L^\infty$ such that:*

$$T_{\varepsilon_j} f \to Tf - vf \quad \text{weakly in } L^q. \tag{2.11}$$

Moreover $\|v\|_\infty \le (a_q + ab^2)b^3 + a_q$.

Proof. Theorem (2.2.9) states that the sequence T_ε is a uniformly bounded family of operators on L^p. For any $v \in L^p$ there is a weakly converging subsequence. Choosing a dense countable subset and a suitable subsequence we obtain the convergence to an operator T_0. Weak limits do not increase the norm. Thus $\|T_0 v\|_{L(L^q)} \le (a_q + ab^2)b^3$. The operator $T - T_0$ is bounded on L^p. Its kernel is by construction zero away from the diagonal. Hence, for any open set U with characteristic function χ_U, $(T-T_0)\chi_U f = \chi_U(T-T_0)f$. This implies for $g \in L^\infty$: $(T-T_0)gf = g(T-T_0)f$. We define v on the open bounded set U by $v|_U = (T-T_0)\chi_U$. It is not hard to check that v is uniquely defined. The identity (2.11) holds at least for characteristic functions f of bounded open sets. The linear combinations of these functions however are dense in L^p. Finally, the norm of the multiplication operator $g = T - T_0$ is $\|g\|_\infty$, which is bounded by its operator norm, $(a_q + ab^2)b^3 + a_q$. □

We want to strengthen the statement about the convergence $T_\varepsilon \to T_0$ to pointwise convergence almost everywhere. This is done via a maximal operator. Let $\zeta(r)$ be the smallest monotone nondecreasing function such that

$$|K(x,y) - K(x,\bar{y})| \le \zeta\left(\frac{d(y,\bar{y})}{\max\{d(x,y), d(x,\bar{y})\}}\right) V(x,y)^{-1} \tag{2.12}$$

if $d(y,\bar{y}) \le \max\{d(x,y), d(x,\bar{y})\}/2$. We assume that

$$\zeta(1/2) + \int_0^{1/2} \zeta(s)\frac{ds}{s} = a_D < \infty. \tag{2.13}$$

Let, for suitable operators T,

$$T_* f(x) := \sup_\varepsilon |T_\varepsilon f(x)|.$$

For later use we are interested in the estimate

$$|\{x : T_* f(x) > \lambda\}| \leq \frac{c}{\lambda} \|f\|_{L^1}. \tag{2.14}$$

It is clear from the proof of Theorem 2.2.9 that $T_\varepsilon f$ of Theorem 2.2.9 is continuous if $f \in L^p$, $1 \leq p < \infty$ and if T satisfies (2.12) and (2.13). Hence $T_* f$ is measurable and $\{x : T_* f > \lambda\}$ is open.

Theorem 2.2.11. *Suppose that the assumptions of Theorem 2.2.8,* (2.9) *and inequality* (2.13) *are satisfied. Then there exists a constant c depending only on b, a_D, a_q, q such that*

1. $\|T_* f\|_{L^q(\mu)} \leq \frac{c}{q-1} \|f\|_{L^q(\mu)}$.
2. *The weak-type estimate* (2.14) *holds.*

Proof. We claim that for $r > 0$

$$T_* f \leq c(r)\{\tilde{M}(|Tf|^r)^{1/r} + \tilde{M}(f)\} \tag{2.15}$$

where $c(r)$ is determined in the proof of the inequality. This implies (with $r = 1$) immediately assertion 1. Assertion 2 is more difficult. Estimate (2.14) is equivalent to

$$\||T_* f|^r\|_{1/r,\infty}^{1/r} \leq c\|f\|_1$$

where the $\|.\|_{1/r,\infty}$ are the Lorentz 'norms' $\|f\|_{p,\infty} = \sup_\lambda \lambda\mu(\{x : |f(x)| > \lambda\})^{1/p}$. We remark that $\|\tilde{M}f\|_{p,\infty} \leq c\|f\|_{p,\infty}$. Indeed,

$$\begin{aligned} \lambda\mu(\{\tilde{M}f > \lambda\})^{1/p} \leq \lambda\mu(\{\tilde{M}(f_{\lambda/2}) > \lambda/2\})^{1/p} &\leq \lambda^{\frac{p-1}{p}} \left(2\int |f_{\lambda/2}| d\mu\right)^{1/p} \\ &\leq 2 \sup_\alpha \alpha\mu(\{|f| > \alpha\})^{1/p}. \end{aligned}$$

Thus, for $r < 1$ and generic constants depending only on the constants mentioned in the theorem,

$$\begin{aligned} \||T_* f|^r\|_{1/r,\infty}^{1/r} &\leq c\|\tilde{M}(|Tf|^r)\|_{1/r,\infty}^{1/r} + c\||\tilde{M}f|^r\|_{1/r,\infty}^{1/r} \\ &\leq c\||Tf|^r\|_{1/r,\infty} + c\||\tilde{M}f|^r\|_{1/r,\infty}^{1/r} \\ &\leq c\|f\|_{L^1(\mu)} \end{aligned}$$

where the last inequality follows from the weak type inequality of Theorem 2.2.8. Thus it suffices to prove the pointwise estimate (2.15) for T_ε for fixed

ε. This we do now. Cutting the singularity off has a regularizing effect. More precisely we claim

$$|T(1-\chi_\varepsilon)f(x) - T(1-\chi_\varepsilon)f(\bar{x})| \le a_D b\tilde{M}f(\bar{x}) \text{ if } x \in B(\bar{x}, \varepsilon/2). \tag{2.16}$$

It is the proof of this inequality where the Dini continuity comes into play. This difference is bounded by

$$\begin{aligned}
&\int_{\{y:d(\bar{x},y)\ge\varepsilon\}} |K(x,y) - K(\bar{x},y)||f(y)|\, d\mu(y) \\
&\qquad = \sum_{k=0}^{\infty} \int_{\{y:2^k\varepsilon<d(\bar{x},y)<2^{k+1}\varepsilon\}} |K(x,y)-K(\bar{x},y)||f(y)|d\mu(y) \\
&\qquad \le \sum_{k=1}^{\infty} \eta\left(\frac{1}{2^k}\right)\mu(B(\bar{x},2^k\varepsilon))^{-1}\int_{B(\bar{x},2^{k+1}\varepsilon)} |f(y)|d\mu(y) \\
&\qquad \le a_D b\tilde{M}f(\bar{x}).
\end{aligned}$$

We use inequality (2.16) in the form

$$|T_\varepsilon f(\bar{x})| \le |Tf(x)| + |T\chi_\varepsilon f(x)| + a_D b\tilde{M}f(\bar{x}) \text{ for all } x \in B(\bar{x},\varepsilon/2) \tag{2.17}$$

which says that we can bound T_ε pointwise in terms of the maximal function of f and the evaluation of Tf and $T\chi_\varepsilon f$ in a neighborhood of $\bar{x}$. We will use this by providing a quantified version of the obvious statement that both can not be large on a large set. Now

$$\begin{aligned}
\mu(\{x\in B(\bar{x},\varepsilon/2) : |Tf(x)| > \lambda)\}) &\le \lambda^{-r}\int_{B(\bar{x},\varepsilon/2)} |Tf(x)|^r d\mu \\
&\le \lambda^{-r}\mu(B(\bar{x},\varepsilon/2)\tilde{M}(|Tf|^r)(\bar{x})
\end{aligned}$$

for any $r > 0$. Thus, if $\lambda > 4^{1/r}[\tilde{M}(|Tf|^r)(\bar{x})]^{1/r}$ then

$$\mu(\{x\in B(\bar{x},\varepsilon/2) : |Tf(x)| > \lambda\}) \le \frac{1}{4}\mu(B(\bar{x},\varepsilon/2)).$$

The weak-type estimate of Theorem 2.2.8 gives

$$\mu(\{x\in B(\bar{x},\varepsilon/c) : |T\chi_\varepsilon f(x)|>\lambda)\})\le\frac{d}{\lambda}\int_{B(\bar{x},\varepsilon)}|f|d\mu\le\frac{db}{\lambda}\mu(B(\bar{x},\varepsilon/2))\tilde{M}f(\bar{x})$$

and if $\lambda \ge 4db\tilde{M}f(\bar{x})$ then

$$\mu(x\in B(\bar{x},\varepsilon/2) : |T\chi_\varepsilon f(x)| > \lambda) \le \frac{1}{4}\mu(x \in B(\bar{x},\varepsilon/2)).$$

Now, if $\lambda \ge \max\{4db\tilde{M}f(\bar{x}), 4^{1/r}[\tilde{M}(|Tf|^r)(\bar{x})]^{1/r}\}$ then there exists x with $|Tf(x)| \le \lambda$ and $|Tf_1(x)| \le \lambda$. This yields estimate (2.15):

$$T_\varepsilon f(\bar{x}) \le 4^{1/r}\tilde{M}(|Tf|^r)(\bar{x})^{1/r} + (4d + a_D)b\tilde{M}f(\bar{x}).$$

□

The maximal function $T_* f$ cannot be too large when $\tilde{M}f$ is small. For later use we state this corollary to the proof of Theorem 2.2.11.

Corollary 2.2.12. *There exists d, depending only on the data of Theorem 2.2.11 such that the following is true. Let Q be a measurable set with $B(z,r) \subset Q \subset B(z,3r)$, $c > 0$, $0 < a < 1$ and $\lambda > 0$ and suppose that*

$$T_* f(x) > a\lambda \quad \textit{for all} \quad x \in Q$$

and there exists $\bar{x} \in B(z,3r)$ with

$$T_* f(\bar{x}) \le a\lambda. \tag{2.18}$$

Then

$$\mu(\{x \in Q : T_* f(x) > \lambda; \tilde{M}f(x) \le c\lambda\}) \le \frac{dc}{1-a}\mu(Q).$$

Proof. Let $B = B(\bar{x}, 12r)$ and $f_1 = f - \chi_B f$. Then, by Theorem 2.2.11,

$$\mu(\{x : T_* \chi_B f(x) > \alpha\}) \le \frac{\tilde{d}}{\alpha}\int_B |f|\, d\mu \le \frac{\tilde{d}b^3 c\lambda}{\alpha}\mu(Q) \tag{2.19}$$

if there is one point $y \in Q$ where $\tilde{M}f(y) \le c\lambda$. Otherwise there is nothing to show.

We apply inequality (2.16) to f_1:

$$|T_\varepsilon f_1(\bar{x}) - T_\varepsilon f_1(x)| \le a_D b^2 \tilde{M}f_1(y) \le a_D b^2 c\lambda$$

for $d(x,\bar{x}) \le 6r$. Hence

$$|T_\varepsilon f_1(x)| \le |T_\varepsilon f(\bar{x})| + |T_\varepsilon \chi_B f(\bar{x})| + a_D b^2 c\lambda.$$

If $\varepsilon < 6r$ then $T_\varepsilon f_1(x) = T_{6r} f_1(x)$. If $\varepsilon > 24r$ then $T_\varepsilon \chi_B f(\bar{x}) = 0$. For $6r \le \varepsilon \le 24r$ we have $|T_\varepsilon \chi_B f(\bar{x})| \le e\tilde{M}(\chi f)(y)$ because the kernel is bounded by $e\mu(B)^{-1}$. We take the supremum with respect to ε. Then

$$T_* f_1(x) \le T_* f(\bar{x}) + a_D b^2 c\lambda + e\tilde{M}f(\bar{y}) \le (a + \bar{d}c)\lambda \tag{2.20}$$

for $x \in Q$. Certainly

$$\{x \in Q : T_* f > \lambda\} \subset \{x \in Q : T_* \chi_B f > \alpha\} \cup \{x \in Q : T_* f_1 > \lambda - \alpha\}.$$

We define $d = 2\max\{\bar{d}, \tilde{d}b^3\}$. We suppose that $a < 1/2 - dc$. Otherwise there is nothing to show. We choose $\alpha = \frac{1-a}{2}\lambda$. The second set is empty by estimate (2.20), and the definition of d. We obtain the assertion from estimate (2.19). □

2.3 Weighted norm estimates

We continue to work in the setting of the last section.

Definition 2.3.1. *Let $1 < p < \infty$. The nonnegative, μ measurable function ω is in the Muckenhoupt class A_p if*

$$\sup_B \mu(B)^{-1} \int_B \omega \, d\mu \left[\mu(B)^{-1} \int_B \omega^{-q/p} \, d\mu \right]^{p/q} \leq a < \infty \tag{2.21}$$

with $1/p + 1/q = 1$. Here the sup is taken with respect to all balls B. The best constant a is called the A_p constant of ω.

In the sequel we identify ω with the measure $\omega\mu$. We are interested in conditions insuring

$$\|\tilde{M} f\|_{L^p(\omega)} \leq c\|f\|_{L^p(\omega)} \tag{2.22}$$

for all $f \in L^p(\mu)$ and finally in conditions insuring

$$\|T f\|_{L^p(\omega)} \leq c\|f\|_{L^p(\omega)}$$

where T is a singular integral operator. It is an important non-obvious fact that all these estimates are closely related.

We define the mean over the ball B by $f_B = \mu(B)^{-1} \int_B f \, d\mu$. There is still another related inequality:

$$f_B^p \leq \frac{a^p}{\omega(B)} \int_B |f|^p d\omega \tag{2.23}$$

for all balls. Here a is a positive constant. If inequality (2.23) holds for all balls B then $\tilde{M}|f|^p \leq c\tilde{M}_\omega |f|^p$ where

$$\tilde{M}_\omega f(x) = \sup_{B(y,r) \ni x} \omega(B(y,r))^{-1} \int_{B(y,r)} |f(z)| \, d\omega(z)$$

is the maximal function with respect to the measure ω. Inequality (2.23) is equivalent to the Muckenhoupt condition (2.21) and it is a consequence of (2.22): $\tilde{M} f(x) \geq f_{B(x,R)}$ and (2.23) follows from (2.22) applied to $\chi_B f$ with $a = c$. Suppose that (2.21) is true. Then we have for all nonnegative f , by Hölder's inequality,

$$\begin{aligned} f_B &= \mu(B)^{-1} \int_B f d\mu = \mu(B)^{-1} \int_B f \omega^{1/p} \omega^{-1/p} \, d\mu \\ &\leq \mu(B)^{-1} \left(\int_B f^p \omega \, d\mu \right)^{1/p} \left(\int_B \omega^{q/p} d\mu \right)^{1/q} \\ &\leq \omega(B)^{-1} a^p \int f_B d\omega \end{aligned}$$

which implies estimate (2.23) with a from (2.21). To prove the converse we suppose that (2.23) holds. We would like to set $f = \omega^{-p/q}$, which we cannot do unless we know more about ω. Thus we set $f_\varepsilon = (\omega+\varepsilon)^{-p/q}$ and let $\varepsilon \to 0$. Hence (2.23) implies (2.21) and the constants are the same.

There is a simple but important consequence of characterization (2.23):

$$\omega(B(x,3r)) \leq ba\omega(B(x,r)) \tag{2.24}$$

which we obtain from choosing $f = \chi_{B(x,R)}$. Then $f_{B(x,9r)} \geq 1/b$ and thus ω is doubling.

Proposition 2.3.2. *Let ω be a nonnegative nontrivial Borel measure on Ω and $1 < p < \infty$. Then*

$$\sup_\lambda \lambda\omega(\{\tilde{M}f > \lambda\})^{1/p} \leq c(A,\omega)\|f\|_{L^p(\omega)} \tag{2.25}$$

for all $f \in L^p(\omega)$ if and only if ω is absolutely continuous with respect to μ and if $\omega \in A_p$.

Proof. Suppose that the operator satisfies (2.25). The derivation of the L^p boundedness of the maximal function from the weak-type estimate in Theorem 2.2.1 (which is a special case of the Marcinkiewicz interpolation theorem) implies here that $f \to \tilde{M}f$ is bounded in $L^q(\omega)$ for all $q > p$. Suppose there exists a set K with $\mu(K) = 0$ and $\omega(K) > 0$. Let χ_n be the characteristic function of the set $\{x : 0 < d(x,K) < 1/n\}$. We have necessarily that $\omega(B_{1/n}(x)) > 0$ for all x because of the doubling condition. Clearly $\chi_n \to 0$ in $L^p(\omega)$ by the dominated convergence theorem because the integrand converges to 0 μ–almost everywhere. But $\tilde{M}_\omega\chi_n(x) = 1$ for all $x \in K$ which gives a contradiction with

$$\|\tilde{M}\chi_n\|_{L^q(\omega)} \leq c\|\chi_n\|_{L^q(\omega)}$$

since the right hand side tends to zero but not the left hand side. Therefore $\omega = h\mu$, where we again set $\omega = h$ by an abuse of notation.

Now let $f \geq 0$ be supported in a ball $B(\bar{x},R) =: B$. If $x \in B$ then $B(x,R) \subset B(\bar{x},2R)$. Hence $\tilde{M}f(x) \geq b^{-1}f_B$ for $x \in B$ and, by inequality (2.25)

$$\omega(B) \leq \omega(\{x : \tilde{M}f(x) > b^{-1}f_B\}) \leq \frac{c^p b^p}{f_B^p}\int_B |f(x)|^p d\omega$$

and $f_B^p\omega(B) \leq c^p b^p \int_B |f|^p d\omega$ hence $\omega \in A_p$.

To prove the converse we remark that $\tilde{M}|f|^p(x) \leq c\tilde{M}_\omega(|f|^p)(x)$ follows from (2.23). Since ω is doubling we infer from Theorem 2.2.1,

$$\omega(\{x : \tilde{M}_\omega f^p(x) > \lambda\}) \leq \frac{d}{\lambda}\int |f|^p d\omega$$

which implies the weak-type inequality (2.25). $\square$

The next step is a crucial and deep one. It is a reverse Hölder inequality and it will imply that $\omega \in A_{p-\varepsilon}$ for some $\varepsilon > 0$ if $\omega \in A_p$. This implies immediately (together with Proposition 2.3.2 and the Marcinkiewicz interpolation theorem) that (2.22) is equivalent to (2.21).

We begin with a property shared by all Muckenhoupt weights. Later we shall see that it actually characterizes the union of all Muckenhoupt weights.

If $\omega \in A_p$ and $0 < \lambda < 1$ then there exists $0 < \beta < 1$ such that

$$F \subset B, \mu(F) \geq \lambda\mu(B) \quad \Longrightarrow \quad \omega(F) \geq \beta\omega(B). \tag{2.26}$$

This is seen by taking $f = \chi_F$. Then $f_B = \frac{\mu(E)}{\mu(B)}$, $\int f^p d\omega^p = \omega(E)$ and hence it follows from (2.23) that

$$\lambda \leq \frac{\mu(E)}{\mu(B)} \leq a\left(\frac{\omega(E)}{\omega(B)}\right)^{1/p}.$$

Definition 2.3.3. *We say that $\omega \in A_\infty$ if there exist $0 < \delta, \varepsilon < 1$ such that for all measurable sets Q, for which there are x and R with $B(x, R) \subset Q \subset B(x, 3R)$, and all $E \subset Q$ with $\mu(E) \leq \varepsilon\mu(Q)$ also $\omega(E) \leq \delta\omega(Q)$.*

Remark 2.3.4. Weights in A_∞ are doubling. If r/R is sufficiently large then $\mu(B(x, R)) \leq \varepsilon\mu(B(x, r))$, hence

$$\omega(B(x, 3R)) \leq \omega(B(x, r)) \leq \delta^{-1}\omega(B(x, R)).$$

Inequality (2.26) implies $\omega \in A_\infty$ because of the following: We choose $\varepsilon < 1/b$. Let Q and F be as in Definition 2.3.3. Let F be the complement of E in $B(x, 3R)$. Then $\mu(F) \geq (1 - b\varepsilon)\mu(B)$ hence $\omega(F) \geq \beta\omega(B(x, 3R))$ and

$$\omega(F \cap Q) \geq \beta\omega(B(x, 3R)) - \omega(B(x, 3R)\backslash Q) \geq \beta\omega(Q).$$

Taking complements we obtain $\omega \in A_\infty$. Much more is true: $\omega \in A_\infty$ implies $\omega \in A_p$ for some p. This will be a consequence of the following proposition.

Proposition 2.3.5. *Suppose that $\omega \in A_\infty$. Then there exists $r > 1$ and $c > 0$ with*

$$\left(\frac{1}{\mu(B)}\int_B \omega^r d\mu\right)^{1/r} \leq \frac{c}{\mu(B)}\int \omega\, d\mu$$

for all balls B.

Proof. Let ε and δ be the constants of Definition 2.3.3. We fix one ball $B = B(x_0, R_0)$. The proof makes heavy use of covering lemmas. We have to make sure first that all balls and cubes occuring in the relevant coverings can be controlled by data in B. Let $\Lambda := \omega(B(x_0, 3R_0))/\mu(B)$ and

$$\tilde{E}_k := \left\{x : \text{ there exists a ball } B_0 \ni x \text{ of radius } r < R_0 \text{ with } \frac{\omega(B_0)}{\mu(B_0)} > d^k\Lambda\right\}$$

where $d = b^5/\varepsilon$. Let $B_1 \subset B(x_0, 3R_0)$ be a ball with radius $R \leq R_0$ and $\omega(B_1)/\mu(B_1) > d^k\Lambda$. We claim that then $R \leq R_0/27$. Suppose this is wrong. Then

$$\omega(B(x_0, 3R_0)) \geq \omega(B_1) \geq d^k\Lambda\mu(B_1) \geq b^{5k}/\varepsilon^k\omega(B(x_0, 3R_0))\mu(B_1)/\mu(B)$$

which is a contradiction because of the doubling property if $k \geq 1$. Hence $R \leq R_0/27$. Let $\bar{B}$ be a ball in $\tilde{E}_k \cap B(x_0, 3R_0)$, $k \geq 1$. We want to show that its radius is smaller than $R_0/9$. It can be covered by balls for which $\omega(\tilde{B})/\mu(\tilde{B}) \geq d\Lambda$, with radius $< R_0/27$. This implies by the Vitali covering lemma that $\omega(\bar{B})/\mu(\bar{B}) \geq b^3\Lambda$, hence the assertion follows by a repetition of the argment above.

We apply the Whitney covering Lemma 2.2.4 to $\tilde{E}_k$ and define E_k as the union of those sets of Remark 2.2.5 which intersect with $B(x_0, R_0)$. Then E_k is contained in $B(x_0, 2R_0)$. This completes the preparations.

Let Q be a set of the Whitney decomposition of $\tilde{E}_{k-1}$ which intersects with B. Let x_j be the sequence of points and Q_j the sets from the Whitney decomposition of E_k (obtained from the Whitney decomposition of $\tilde{E}_k$ above) with nonempty intersection with Q. Let $R = d(x_j, \tilde{E}^0_{k-1})$. There exists a point $z \in \overline{B(x_j, R)} \cap \tilde{E}^0_{k-1}$, and by construction

$$\frac{\omega(Q_j)}{\mu(Q_j)} \leq b^2\frac{\omega(B(z, 3R))}{\mu(B(z, 3R))} \leq b^2 d^{k-1}\Lambda. \tag{2.27}$$

Every point of the interior of Q_j is contained in a ball $B_0 \subset E_k$ of radius $\leq R_0/27$ with $\frac{\omega(B_0)}{\mu(B_0)} \geq d^k\Lambda$ or

$$\mu(B_0) \leq d^{-k}\Lambda^{-1}\omega(B_0).$$

Hence, by the Vitali covering Lemma 2.2.2 and the doubling property of μ, $\mu(Q_j) \le d^{-k} b \Lambda^{-1} \omega(Q_j)$. We sum over all Q_j and get, using again the doubling property,

$$\mu(Q \cap E_k) \le d^{-k} \Lambda^{-1} b \omega(Q) \le d^{-1} b^3 \mu(Q).$$

Thus by the A_∞ property $\omega(E_k) \le \delta \omega(E_{k-1}) \le \delta^k \omega(E_1) \le \omega(B(x_0, 3R_0))$ and

$$\omega(\{x \in B : \tilde{M} \chi_B \omega(x) > \lambda\}) \le \delta^{-1} \lambda^{\frac{\ln \delta}{5 \ln b - \ln \varepsilon}} \omega(B(x_0, 3R_0))$$

and, with $r = 1 + \frac{\ln \delta}{2(5 \ln b - \ln \varepsilon)}$,

$$\begin{aligned} \|\omega \chi_B\|^r_{L^r(\mu)} = \|\omega^{r-1} \chi_B\|_{L^1(\omega)} &= \int_0^\infty \omega(\{x \in B : \omega(x) > \lambda^{1/(r-1)}\}) d\lambda \\ &\le \delta^{-1} \int_0^\infty \min\{1, \lambda^{-2}\} d\lambda \omega(B(x_0, 3R_0)) \\ &\le 2\tilde{b} \delta^{-1} \omega(B). \end{aligned}$$

where $\tilde{b}$ is the doubling constant of ω. □

Corollary 2.3.6. *If $\omega \in A_p$ then there exists $q < p$ with $\omega \in A_q$.*

Proof. If $1/p + 1/q = 1$, $1 < p, q < \infty$ and $\omega \in A_p$ then $\sigma = \omega^{-q/p} \in A_q \subset A_\infty$. By Proposition 2.3.5 σ satisfies a reverse Hölder inequality for all balls B:

$$\left(\int_B \sigma^r \, d\mu \right)^{1/r} \le c \int_B \sigma d\mu.$$

This is the A_{p_1} condition for $p_1 = 1 + \frac{p-1}{r} < p$. □

Corollary 2.3.7. *Let μ be a nonnegative Borel measure on Ω and $1 < p < \infty$. Then $f \to \tilde{M} f$ is a bounded operator from $L^p(\omega)$ to itself if and only if $\omega \in A_p$.*

Proof. Suppose that $\omega \in A_p$. Then it is also in A_q for some $q < p$. Hence it satisfies estimate (2.25) of Proposition 2.3.2 for this q. We have seen in the proof of Proposition 2.3.2 that then $\tilde{M} f : L^p(\omega) \to L^p(\omega)$ is bounded. Conversely $L^p(\omega)$ boundedness implies $\omega \in A_p$. This has been discussed at the beginning of this section. □

Theorem 2.3.8. *Suppose that $\omega \in A_\infty$ and that T satisfies the assumptions of Theorem 2.2.11. If $1 < p < \infty$ then*

$$\|T_* f\|_{L^p(\omega)} \le c \|\tilde{M} f\|_{L^p(\omega)}$$

provided the right hand side is finite.

Proof. Let ε and δ be the constants of ω in the characterization of A_∞. Let $a < 1$ be given. We claim that there exists c such that

$$\omega(\{x : T_*f(x) > \lambda; \tilde{M}f(x) \leq c\lambda\}) \leq \delta\omega(\{x : T_*f(x) > a\lambda\}). \tag{2.28}$$

We suppose that this estimate holds. Then

$$\begin{aligned} \|T_*f\|^p_{L^p(\omega)} &= p\int_0^\infty \lambda^{p-1}\omega(\{x : |T_*f(x)| > \lambda\})\,d\lambda \\ \leq p\int_0^\infty &\delta\lambda^{p-1}\omega(\{|T_*f(x)| > a\lambda\}) + c\lambda^{p-1}\omega(\{x : \tilde{M}f(x) > c\lambda\})\,d\lambda \\ \leq a^{-p}&\delta\|T_*f\|^p_{L^p(\omega)} + c^{-p}\|\tilde{M}f\|^p_{L^p(\omega)}. \end{aligned} \tag{2.29}$$

We choose a so that $a^{-p}\delta < 1$. We obtain the desired conclusion provided we know that $T_*f \in L^p(\omega)$.

We prove (2.28). We recall that $\{T_*f > a\lambda\}$ is open. We use a Whitney covering. Let $B(x_0, r) \subset Q \subset B(x_0, 3r)$ be one of the cubes. It suffices to prove the claim in the form

$$\omega(\{x \in Q : T_*f(x) > \lambda; \tilde{M}f(x) \leq c\lambda\}) \leq \delta\omega(Q) \tag{2.30}$$

assuming $T_*f(x) > a\lambda$ in Q and $T_*f(x) \leq a\lambda$ for one x with $d(x, x_0) = 3r$. We know by Corollary 2.2.12 that

$$\mu(\{x \in Q : T_*f(x) > \lambda; \tilde{M}f(x) \leq c\lambda\}) \leq \frac{Ac}{1-a}\mu(Q).$$

We choose c small so that $\frac{Ac}{1-b} \leq \varepsilon$. Then estimate (2.30) follows from the definition of (2.3.3).

The assertion $Mf \in L^p(\omega)$ implies that $\int_\Omega \omega/(1 + V(x, y))^p d\mu(y) < \infty$. From the kernel estimate it follows that for all $f \in L^q(\Omega)$

$$T_*(f) \leq cV(x, y)^{-1}$$

which shows that the estimate holds for Hölder continuous (with respect to d) compactly supported functions f. These are dense in $L^p(\omega)$ and an approximation yields the result. □

Corollary 2.3.9. *Suppose that $1 < p < \infty$, $\omega \in A_p$ and T satisfies the assumptions of Theorem 2.2.11. Then*

$$\|Tf\|_{L^p(\omega)} \leq c\|f\|_{L^p(\omega)}.$$

Proof. By Corollary 2.2.10 $|Tf(x)| \leq |T_* f(x)| + c|f(x)|$ hence

$$\|Tf\|_{L^p(\omega)} \leq c\|f\|_{L^p(\omega)} + \|T_* f\|_{L^p(\omega)} \leq c\|f\|_{L^p(\omega)} + \tilde{c}\|\tilde{M}f\|_{L^p(\omega)} \leq \hat{c}\|f\|_{L^p(\omega)}.$$

□

We conclude this section by providing tools for constructing interesting functions in A_p. It turns out to be useful to introduce the class A_1.

Definition 2.3.10. *A non-vanishing non-negative function ω is in A_1 if $\tilde{M}\omega \leq c\omega$ almost everywhere.*

There is an alternative characterization.

Proposition 2.3.11. *Suppose that $\omega \in A_1$. Then there exist a nonnegative function f and $0 < q < 1$ such that $(\tilde{M}f)^q/\omega$ is bounded from above and below almost everywhere. Conversely, if f is a nonnegative measure with $\tilde{M}f < \infty$ almost everywhere then $\tilde{M}f^q \in A_1$ for all $0 < q < 1$.*

Proposition 2.3.12. *Let $1 < p < \infty$ and let $\omega \in A_p$. Then there exist ω_1 and $\omega_2 \in A_1$ with $\omega = \omega_1\omega_2^{1-p}$. Conversely, if $\omega_1, \omega_2 \in A_1$ then $\omega_1\omega_2^{1-p} \in A_p$.*

See Stein [125] for a proof of Proposition 2.3.11 and Proposition 2.3.12.

2.4 Notes

The theory of singular integral operators began with complex methods in the one-dimensional setting (Zygmund [139], [140]) and with the study of the Hilbert transform by real variable techniques (Titchmarsh [130] and Marcinkiewicz [92]). Calderón and Zygmund [32] introduced the Calderón-Zygmund decomposition, which became the basis for the theory of singular integral operators.

The classes A_p appeared in Rosenblum [117]. The relation with the Hardy-Littlewood maximal function (Corollary 2.3.7) is due to Muckenhoupt [101]. The bounds for singular integrals of Theorem 2.3.8 (together with Theorem 2.2.11) were obtained by Coifman and Fefferman [38] after some earlier results by Hunt, Muckenhoupt and Wheeden [73] and Muckenhoupt [102]. Coifman and Fefferman [38] characterized A_∞. The characterization of A_1 is due to Coifman and Rochberg [40]. The quotient representation is from Jones [77] and Coifman, Jones and Rubio de Francia, [39]. Rokhman and Solonnikov [116] studied A_p–weights in a polyhomogeneous setting. The interested reader should consult Christ [35], Garcia-Cuerva and Rubio de Francia [63], Stein [125] and Strömberg and Torchinsky [127] for a more complete development of the theory, including the Non-Euclidean setting.

Chapter 3

Fundamental solutions

We construct and estimate several fundamental solutions to a number of different problems: elliptic equations, elliptic equations with drift, Oseen operator, homogenization, Kohn-Laplacian and a Kolmogorov equation. The chapter has a three-fold purpose: it illustrates the theory of the last chapter, it shows its relevance and usefulness for other problems, and it prepares for the Gaussian estimates and their applications of the next chapter.

3.1 Preliminaries

The Fourier transform

Functions in this part are assumed to have complex values. We define for $f \in L^1$ its Fourier transform by

$$\hat{f}(k) = F(f)(k) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-ik\cdot x} f(x)\, dx$$

and the cotransform $\check{\mathrm{f}} = F^{-1}(f)$ as the adjoint operator. Both map the Schwartz space to itself, they are the inverse of each other, and hence by duality they map its dual, the space of tempered distributions to itself. The Fourier transform is an isometry on $L^2(\mathbb{R}^n)$.

Sobolev spaces

We define the Bessel-potential spaces $H^{s,p}(\mathbb{R}^n)$ for $s \in \mathbb{R}$ and $1 < p < \infty$ by their norm

$$\|f\|_{H^{s,p}(\mathbb{R}^n)} = \|F^{-1}((1 + |\xi|^2)^{s/2} \hat{f})\|_{L^p(\mathbb{R}^n)}.$$

Let k be a positive integer, $1 \le p \le \infty$ and $\Omega \subset \mathbb{R}^n$ be open. The Sobolev spaces $W^{k,p}(\Omega)$ are defined as the spaces of distributions for which the norm

$$\|f\|^p_{W^{k,p}(\Omega)} = \sum_{|\alpha|\le N} \int_\Omega |D^\alpha f|^p \, dx$$

is finite. There are obvious modifications if $p=\infty$. For $1<p<\infty$ we have $W^{k,p}(\mathbb{R}^n) = H^{k,p}(\mathbb{R}^n)$ as topological vector spaces (i.e. the identity on C_0^∞ extends to a unique continuous invertible linear operator from one space to the other).

The Fourier transform changes convolution to multiplication and vice versa:

$$F(fg) = (2\pi)^{-n/2}\hat{f} * \hat{g}, \qquad F(f*g) = (2\pi)^{n/2}\hat{f}\hat{g}.$$

Sobolev imbedding

Let $x_0 \in \mathbb{R}^n$ and $Q = \{x \in \mathbb{R}^n : \max_i |x^i - x_0^i| \le R/2\}$. Suppose that $u \in W^{n,1}(Q)$. Then

$$\|u\|_{L^\infty(Q)} \le \sum_{j=0}^n R^{j-n} \sum_{i_1<\cdots<i_j} \|\partial^{i_1}\ldots\partial^{i_j}u\|_{L^1(\mathbb{R}^n)}. \tag{3.1}$$

This follows from an iterative application of the fundamental theorem of calculus, plus an integration. We give the argument in more detail in two dimensions:

$$\begin{aligned} u(x_1,x_2) &= u(y_1,x_2) + \int_{y_1}^{x_1} \partial_1 u(z_1,x_2)dz_1 \\ &= u(y_1,y_2) + \int_{y_1}^{x_1} \partial_1 u(z_1,y_2)dz_1 + \int_{y_1}^{x_1}\int_{y_2}^{x_2} \partial^2_{12}u(z_1,z_2)dz_1\,dz_2. \end{aligned}$$

An integration over the cube with respect to y gives the estimate.

If $1 \le p < q < \infty$ and $1/q + 1/n = 1/p$ then the Sobolev inequality

$$\|u\|_{L^q(\mathbb{R}^n)} \le c(n)q\|\nabla u\|_{L^p(\mathbb{R}^n)} \tag{3.2}$$

holds. See Friedman [61] or Chapter 4.

Complex interpolation

Let C be the complex strip $\{z = x+iy \in \mathbb{C} : 0 \le y \le 1\}$. Let $1 \le p_0, p_1, q_0, q_1 \le \infty$. Suppose that $C \ni z \to T(z) \in L(L^{p_1}\cap L^{p_2}, L^{q_1} \cup L^{q_2})$ is a uniformly bounded map, holomorphic in the interior of C. Let

$$\frac{1}{p_t} = \frac{1-t}{p_0} + \frac{t}{p_1}, \quad \frac{1}{q_t} = \frac{1-t}{q_0} + \frac{t}{q_1}.$$

Suppose that $T(is)$ extends to a bounded operator from L^{p_0} to L^{q_0}, uniformly bounded by c_0 and $T(1+is)$ extends to a bounded operator from L^{p_1} to L^{q_1}, uniformly bounded by c_1 for all $s \in \mathbb{R}$. Then, if $0 \le t \le 1$, $T(t) : L^{p_t} \to L^{q_t}$ with operator bound $c_0^{1-t} c_1^t$. A proof is contained in Stein and Weiss [126].

3.2 The theorem of Mihlin-Hörmander

3.2.1 Kernels

Let $\phi \in C_0^\infty(\mathbb{R}^n)$ be a radial nonnegative function supported in $[1/4, 1]$ such that $\sum \phi(2^{-j}\xi) = 1$. Let $s > n$. A function $m \in L^\infty(\mathbb{R}^n)$ is called a *Mihlin-Hörmander* Fourier multiplier if

$$\|\phi m(\lambda * .)\|_{H^{s,2}(\mathbb{R}^n)} \le c_0 < \infty \tag{3.3}$$

for all positive λ with c_0 independent of λ. In particular m is bounded and

$$L^2(\mathbb{R}^n) \ni f \to F^{-1}(m\hat{f}) =: T_m f \in L^2(\mathbb{R}^n)$$

is a bounded operator.

Theorem 3.2.1. *The operator T_m is a singular integral operator with*

$$\|T_m f\|_{L^2(\mathbb{R}^n)} \le \|m\|_\infty \|f\|_{L^2(\mathbb{R}^n)}$$

the kernel of which satisfies (2.9) *and* (2.12) *with respect to the metric in* $\mathbb{R}^n$ *and the Lebesgue measure.*

Remark 3.2.2. This has the consequence that Theorem 2.2.8 and Corollary 2.3.9 hold for these operators.

Proof. We decompose T into dyadic pieces,

$$T = \sum_{j\in\mathbb{Z}} T_j \quad \text{ with } T_j = T_{m\phi(2^{-j}*.)}.$$

The operator T_j has a kernel K_j which satisfies by assumption

$$\int (1 + |2^j x|^2)^s |D_x^k K_j|^2 dx = 2^{-jn+2jk} \int (1 + |z|^2)^s |D_z^k K_j(2^{-j}z)|^2 dz \le c_1 2^{jn+2jk}.$$

We claim that, with a constant c_1 depending on n and s,

$$2^{-j} |\nabla K_j(x)| + |K_j(x)| \le c_1 2^{jn} (1 + 2^j |x|)^{-s}.$$

Suppose that this is true. Then

$$|K(x)| \le c_1 \Big[\sum_{j \le -\ln_2 |x|} 2^{jn} + \sum_{j > -\ln_2 |x|} |x|^{-s} 2^{j(n-s)} \Big] \le c_2 |x|^{-n}$$

and, if $s < n+1$, $|x-y| \le |x|/2$, with $L = -\ln_2(|x-y|)$

$$\begin{aligned} |K(x) - K(y)| &\le |x-y| \sum_{j \le L} \|\nabla K_j\|_{L^\infty(B(x,|x|/2))} + 2 \sum_{j>L} \|K_j(x)\|_{L^\infty(B(x,|x|/2))} \\ &\le c|x-y| \Big[\sum_{j \le -\ln_2 |x|} 2^{j(n+1)} + |x|^{-s} \sum_{-\ln_2 |x| < j \le L} 2^{j(n+1-s)} \Big] \\ &\quad + c|x|^{-s} \sum_{j>L} 2^{j(n-s)} \\ &\le c_2 \left(|x-y||x|^{-n-1} + |x-y|^{s-n} |x|^{-s} \right), \end{aligned}$$

which is the desired estimate.

It remains to prove the bounds for K_j and ∇K_j. We apply Sobolev's inequality (3.1) in cubes $Q = \{x : \max |x_i - x_i^0| \le 2^{-j-1}\}$:

$$\begin{aligned} |K_j(x^0)| &\le \sum_{k=0}^{n} 2^{j(n-k)} \sum_{i_1 < \dots < i_k} \|\partial_{i_1} \dots \partial_{i_k} K_j\|_{L^1(Q)} \\ &\le \sum_{k=0}^{n} 2^{j(n/2-k)} \sum \|\partial_{i_1} \dots \partial_{i_k} K_j\|_{L^2(Q)} \\ &\le c_3 2^{nj} (1 + |2^j x^0|)^{-s} \sum_k 2^{-j(n/2+k)} \|(1+|2^j x|^2)^{s/2} D^k K_j\|_{L^2(\mathbb{R}^n)} \\ &\le c_4 2^{nj} (1 + |2^j x^0|)^{-s} \end{aligned}$$

which implies the first half of the desired estimates. Here D^k denotes the vector of all derivatives of order k. The derivatives can be estimated in the same way. □

There are many examples of operators given by Fourier multipliers. We list some of them.

The Hilbert transform

Here $n = 1$, $m(\xi) := \begin{cases} -i & \text{if } \xi \ge 0 \\ i & \text{if } \xi < 0 \end{cases}$. The kernel is $g(x) = \frac{1}{\pi x}$, which can be computed by calculating the Fourier cotransform

$$\frac{1}{\sqrt{2\pi}} F^{-1}(e^{-\varepsilon|\xi|} m(\xi)) = \frac{1}{2\pi} \left(\frac{1}{x + i\varepsilon} + \frac{1}{x - i\varepsilon} \right)$$

and taking the limit $\varepsilon \to 0$.

The Riesz operators

The Riesz operators R_i are defined by the Fourier multipliers $m_i(\xi) = i\xi_i/|\xi|$. The kernels are

$$g_i(x) = \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}} \frac{x_i}{|x|^{n+1}}.$$

The identity $\partial_i \partial_j = R_i R_j \Delta$ can be verified by applying the Fourier transform. Thus we obtain

$$\|\partial_i \partial_j u\|_p \le c \frac{p^2}{p-1} \|\Delta u\|_p.$$

We refer to Stein [124] for more details.

Weights

We have seen that Mihlin-Hörmander multipliers define bounded operators from $L^p(\omega\, dx)$ to itself if $\omega \in A_p$ with respect to the standard metric and the Lebesgue measure. The following are examples of elements in A_p.

1. $|x|^\sigma$ for $-n < \sigma < n(p-1)$.
2. Suppose that f and g are nonnegative measures for which Mf and Mg are finite almost everywhere. If moreover $0 \le s < 1$ then

   $$\omega = (Mf)^s (Mg)^{s(1-p)} \in A_p.$$

   This is a consequence of Proposition 2.3.11 and Proposition 2.3.12.

Other scaling

We may assign different homogeneities to different variables. There are only marginal changes in the arguments. We consider

$$(\partial_t - \Delta) u = f \qquad \text{in } \mathbb{R}^n \times \mathbb{R}.$$

We are interested in the parabolic Riesz operator given by the Fourier multiplier $\frac{\xi_i \xi_j}{i\tau + |\xi|^2}$. We denote its kernel by $K_{ij}(x,t)$. The relevant scaling is $(\xi, \tau)_\lambda = (\lambda x, \lambda^2 \tau)$, the distance is

$$d(x,t;y,s) = \sqrt{|x-y|^2 + |t-s|}$$

and, again with the Lebesgue measure,

$$V(x,t;y,s) = cd(x,t,y,s)^{n+2}.$$

The kernel estimates are

$$K_{ij}(x,t,y,s) \leq cV(x,t,y,s)^{-1},$$

$$|K_{ij}(x,t,y,s) - K_{ij}(\tilde{x},\tilde{t},\tilde{y},\tilde{s})| \leq c\frac{d(x,t,\tilde{x},\tilde{t}) + d(y,s,\tilde{y},\tilde{s})}{d(x,t;y,s)} V(x,t,y,s)^{-1}$$

for $d(x,t,\tilde{x},\tilde{t}) + d(y,s,\tilde{y},\tilde{s}) \leq d(x,t,y,s)/4$. Of course these estimates are weak: they do not capture the fact that the heat kernel vanishes in the half space $\{t < 0\}$. The L^2-estimate is obtained through the Fourier transform.

An example of an A_p-weight is $|t|^\sigma$ with $-1 < \sigma < p-1$. The characteristic function of the half space $t < 0$ is not an A_p weight: it is not doubling. Nevertheless the weighted estimates with this weight hold:

$$\|\chi_{t\leq 0} T_m f\|_p \leq \|T_m \chi f\|_p \leq c(p)\|\chi f\|_{L^p}.$$

There seems to be no characterization of weights for which weighted estimates hold in this setting. This is an open question.

3.2.2 Variants

Periodic functions

There are versions of the Mihlin-Hörmander theorem on the torus $\mathbb{T}^n := (\mathbb{S}^1)^n$ where $\mathbb{S}^1$ is the unit circle in $\mathbb{C}$. In that case the multipliers are functions on the lattice $\mathbb{Z}^n$. Using the same notation as above the Mihlin-Hörmander condition has the form

$$\sum_{k\leq n+1} \lambda^{2k-n} |D^k(\phi(\lambda^{-1}\xi)m(\xi))|^2$$

where D^k denotes all products of length k of the operators $f \to f(a+e_i) - f(a)$ with $a \in \mathbb{Z}^n$. The distance on the torus is again the standard distance and the measure is the Lebesgue measure. The proof requires only marginal changes.

Mihlin-Hörmander: the lattice

This is the dual to the situation above. Here we consider the problem on the lattice $\mathbb{Z}^n$ and the Fourier multipliers are bounded measurable functions on the torus $\mathbb{T}^n$, which we identify with 2π–periodic functions in all variables on $\mathbb{R}^n$. The Mihlin-Hörmander condition has the form $D^k m \in L^2$ away from $2\pi\mathbb{Z}^n$ for $k \le n+1$ and

$$\sup_{\lambda > 2\pi} \|\phi m(\lambda * .)\|_{H^{n+1,2}(\mathbb{R}^n)} < \infty. \tag{3.4}$$

The metric in $\mathbb{R}^n$ induces a metric in $\mathbb{Z}^n$. The measure is the counting measure, which is doubling. The proof that m defines a singular integral operator is the same as above.

A particular example is given by $L = \sum_j \partial_j^- \partial_j^+$ where $\partial_j^\pm$ are the forward/backward differences in direction j, and $M = \partial_j^+ \partial_k^+ L^{-1}$. The Fourier multiplier of L is

$$l(\xi) = \sum_j e^{i\xi_j} - 2 + e^{-i\xi_j} = -4 \sum_j \sin^2(\xi_j/2)$$

and the Fourier multiplier defined by M is $(e^{i\xi_j} - 1)(e^{i\xi_k} - 1) l(\xi)^{-1}$. This function satisfies the Mihlin-Hörmander condition (3.4). We obtain the following result.

Proposition 3.2.3. *The discrete Laplacian satisfies*

$$\|\partial_j^+ \partial_k^+ u\|_{l^p(\mathbb{Z}^N)} \le c \|Lu\|_{l^p(\mathbb{Z}^N)}.$$

Again this result remains true for weighted spaces. A particular example of a weight is $\omega = (1 + |x|)^s$ for $-n < s < (p-1)n$.

The next examples show that very different metrics occur naturally in connection with partial differential equations.

3.3 The operator $\Delta u - 2\partial_n u$

Here it is the drift term which leads to new phenomena. This is, among other things, a model operator for Navier-Stokes with drift at infinity. It has been studied by Farwig [54]. Diffusion is dominant on small scales and the drift becomes important on large scales.

The kernel

We define the fundamental solution as the unique tempered distribution which satisfies $\Delta g - 2\partial_n g = \delta$ and $g(x) \to 0$ as $x \to \infty$. The theorem of Malgrange-Ehrenpreis (see Nirenberg [109]) ensures existence of a fundamental solution. It follows from Hörmander [71], Theorem 10.2.1 that it is a tempered distribution. Let g_1 and g_2 be two fundamental solutions which are tempered distributions. The Fourier transform of their difference is supported at 0. Hence the difference is a polynomial and the decay at infinity gives uniqueness of the fundamental solution. It remains to show the decay property for one fundamental solution.

We shall take a different more elementary approach. We claim that

$$g(x) = c_n e^{x_n} |x|^{-\frac{n-2}{2}} K(|x|)$$

where K is a MacDonald or modified Bessel function. We make the Ansatz $g = e^{x_n} u$ where u satisfies $\Delta u - u = \delta$ and $|x|^N u(x) \to 0$ as $x \to \infty$ for all $N > 0$. Existence and uniqueness of a fundamental solution which is a tempered distribution are again clear. We shall see that there exists a nontrivial radially symmetric function u with these properties: it is $u(x) = r^{(2-n)/2} k(r)$ where $r = |x|$ and where k satisfies the modified Bessel differential equation

$$k^{(2)} + \frac{1}{r} k' - \left(1 + \frac{n-2}{2r^2}\right) k = 0 \qquad \text{for} \qquad r > 0,$$

together with suitable conditions at $r = 0$ and ∞. We search for a solution with $k(r) \to 0$ as $t \to \infty$ and $r^{\frac{n-2}{2}} k(r) \to \frac{\Gamma(n/2)}{2(n-2)\pi^{n/2}}$ as $r \to 0$ if $n \neq 2$ and $-(\ln r)^{-1} k(r) \to (2\pi)^{-1}$ if $n = 2$. This implies that $k(r) = \tilde{c}_n K_{\frac{n-2}{2}}(r)$, a multiple of the Macdonald function. It satisfies, for all $k \geq 0$,

$$|\partial_r^k (e^r \sqrt{1+r} k(r))| \leq c_k (1+r)^{-k},$$

which follows from

$$K_\nu(z) = \left(\frac{\pi}{2z}\right)^{1/2} e^{-z} \left(\sum_{m=0}^{N} \frac{\prod_{j=1}^{m} (4\nu^2 - (2j-1)^2)}{m!(8z)^m} + O(|z|^{-N-2}) \right)$$

by the Cauchy integral formula. The product is defined to be 1 if $m = 0$. This formula holds for z in a sector of the complex plane containing $\mathbb{R}^+$, see Watson [138], paragraph 7.23.

The sum on the right hand side is finite if ν is the half of an odd integer. In that case we obtain an identity if $N = \nu - 1/2$. If $n = 1$ then $u(x) = \frac{|x|}{2} e^{-|x|}$, if $n = 3$ then $u(x) = \frac{1}{4\pi|x|} e^{-|x|}$ and if $n = 5$ then

$$u(x) = c_5 |x|^{-2} e^{-|x|} (1 + 1/|x|) = \frac{1}{8\pi^2} e^{-|x|} (|x|^{-3} + |x|^{-2}).$$

Let $s = |x| - x_n$. We easily obtain the crucial estimate of derivatives of the fundamental solution

$$|\partial^\alpha g(x)| \leq c_\varepsilon e^{-(1-\varepsilon)s} (1 + |x|)^{-|\alpha|/2 - \alpha_n/2} \tag{3.5}$$

because derivatives falling not on s decay faster. We are interested in the operator $f \to \partial_{ij} g * f$ where we take the precaution to require $f \in C_0^\infty$. Let $E_{ij}(x) = \partial_{ij} g$ away from the origin.

The corresponding Fourier multiplier is $m_{ij}(\xi) = \frac{\xi_i \xi_j}{|\xi|^2 + 2i\xi_n}$. It satisfies the assumptions of multiplier theorems of Marcinkiewicz (see [124]) or Lizorkin type [91] but not the assumptions of Theorem 3.2.1. Hence this multiplier defines a bounded operator in $L^p(\mathbb{R}^n)$, $1 < p < \infty$. This gives almost no information on the kernel, and we do not get estimates in weighted spaces from these multiplier theorems. Below we shall see that the operator is a singular integral operator with respect to a different metric.

The metric

Our experience has shown that we should try to understand mapping properties of operators in terms of metrics and the Lebesgue measure. The translation invariant metric is defined by (with $\hat{x} = (x_1, \dots, x_{n-1})$),

$$d(x,0) = \max\left\{ |\hat{x}|, \frac{|x_n|}{\sqrt{1+|x_n|}} \right\}. \tag{3.6}$$

Let c_n be the volume of an Euclidean ball in $\mathbb{R}^n$ of radius 1. The volume of a ball of radius R in the metric d is

$$|B_R(0)| = c_{n-1}|R|^n \left(R + \sqrt{R^2+4} \right)$$

where c_n is the measure of the n-dimensional ball. Thus the Lebesgue measure is doubling.

We write $a \sim b$ if there exists c with $c^{-1}a < b < ca$. Then, if $V(x,0)$ is the volume of the ball of radius $d(x,0)$,

$$V(x,0) \sim \left(|\hat{x}|^2 + \frac{x_n^2}{1+|x_n|} \right)^{\frac{n}{2}} \left(1 + \sqrt{|\hat{x}|^2 + \frac{x_n^2}{1+|x_n|}} \right)$$

and, for $d(x,0) > 1$

$$V(x,0) \sim |\hat{x}|^{n+1} + |x|^{\frac{n+1}{2}} \sim \left(|x|(1+|x|-|x_n|) \right)^{\frac{n+1}{2}}.$$

Hence $|E_{ij}| \le cV(x,0)^{-1}$ and

$$|E_{ij}(x) - E_{ij}(y)| \le c \frac{d(x,y)}{d(x,0)V(x,0)}$$

if $d(x,y) \le \frac{1}{2} d(0,x)$. We have introduced all the data of Theorem 2.2.11 and Theorem 2.3.8. The first part of the following result is a consequence.

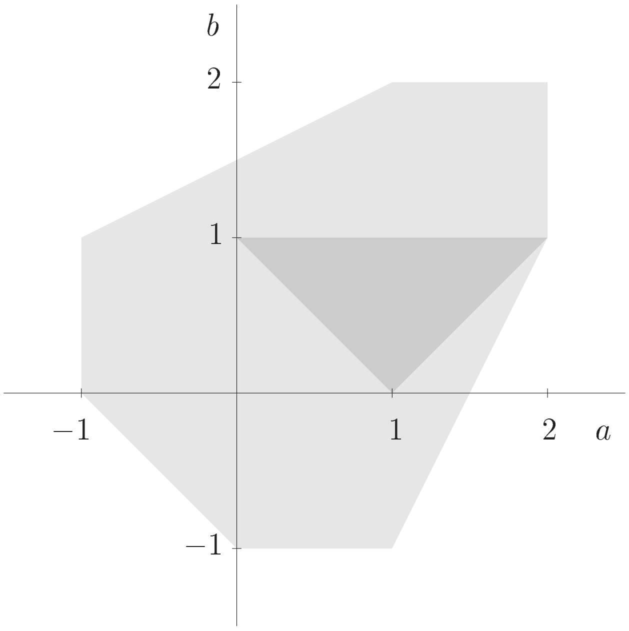

Figure 3.1: The region of Prop. 3.3.1 for $p = 2$ and of Farwig [55], Theorem 5.6

Proposition 3.3.1. *The following estimate holds for $u = \partial_{ij} g * w$:*

$$\|u\|_{L^p(\omega)} \leq c\|w\|_{L^p(\omega)}$$

for $\omega \in A_p$ where A_p is defined with respect to the Lebesgue measure and the metric d. The same estimate holds for the weight functions

$$\omega_{ab} = (1 + |x|)^a (1 + |x| - x_n)^b$$

if $|a - b| < 1$, $-1 < a, b, a + b < 2$, $2a - b, 2b - a < 3$, $p = 2$ and $n = 3$ (see figure 3.1).

Remark 3.3.2. It is not hard to see that $\omega_{a,b} \notin A_p$ if $a + b \geq 4(p - 1)$. We shall see similar arguments below.

Muckenhoupt weights

Here we construct elements of A_p with metric d.

Lemma 3.3.3. *The following functions are elements of A_1 if $-1 < \sigma \leq 0$: $V(x, 0)^\sigma$, $(1 + V(x, 0))^\sigma$, $|x_n|^\sigma$, $(1 + |x_n|)^\sigma$, $|\hat{x}|^{(n-1)\sigma}$, $|x|^\sigma$, $(1 + |x|)^\sigma$, $(1 + |x| - |x_n|)^{\frac{n-1}{2}\sigma}$ and $(1 + |x| - x_n)^\sigma$.*

Proof. The Hardy-Littlewood maximal operator (with respect to d) applied to the Dirac measure is comparable to $V(x)^{-1}$. Hence, by Proposition 2.3.11, $V(x)^\sigma \in A_1$ for $-1 < \sigma \leq 0$. Application of the maximal operator to the $n-1$ dimensional Dirac-measure $\delta_{\mathbb{R}^{n-1}}$ gives $|\hat{x}|^{-n-1}$. Thus $\omega = |\hat{x}|^{\sigma(n-1)} \in A_1$. Similarly, when we apply the maximal function to $\delta(x_n)$ we obtain $|x_n|^\sigma \in A_1$ for $-1 < \sigma \leq 0$.

We turn to $\omega = |x|^\sigma$. We shall show that $M\omega \leq c\omega$. It is easy to see that the A_1 condition is satisfied for $|x|^\sigma$ for $R \leq 1$. Hence we consider $R \geq 1$.

We may and do use the equivalent norm $\bar{d}(x,y) = \max\{|\hat{x} - \hat{y}|, \sqrt{|x_n - y_n|}\}$ for points with $d(x,y) \geq 1$. With this norm $B(x^0, R) = \{x : |\hat{x} - \hat{x}^0| \leq R, |x_n - x_n^0| \leq R^2\}$. If $R < \frac{1}{2} d(x,0)$ then $|x| \sim |y|$ for $y \in B(x,R)$. Then $|B(x,R)|^{-1} \int_{B(x,R)} |y|\, dy \sim |x|$. Similarly we may assume that $x = 0$ if $d(x,0) < R/2$. Then $\int_{B(0,R)} \omega dx \leq \omega(1)|B(0,R)|$ since ω assumes its maximum at $x = 0$ and the A_1 condition for those balls is satisfied with constant 1. It remains to consider the case $R/2 \leq d(x,0) \leq 2R$, $R \geq 1$. Then

$$
\begin{aligned}
|B(x,R)|^{-1} \int_{B(x,R)} (1+|x|)^\sigma dx &\sim R^{\sigma-1} + R^{2\sigma} \\
&\sim (|\hat{x}| + \sqrt{|x_n|})^{\sigma-1} + (|\hat{x}| + \sqrt{x_n})^{2\sigma} \leq c|x|^\sigma
\end{aligned}
$$

if and only if $\sigma > -1$. Hence $\omega \in A_1$ if $-1 < \sigma \leq 0$.

We turn to $\omega(x) = (1 + |x| - |x_n|)^{\frac{n-1}{2}\sigma}$: The case $|x_n| \leq R^2/2$ follows from the same arguments as above. One easily computes

$$
|x| - |x_n| = \sqrt{x_n^2 + |\hat{x}|^2} - |x_n| = |x_n|(\sqrt{1 + |\hat{x}|^2/x_n^2} - 1) \sim |\hat{x}|^2/|x_1|
$$

for $|\hat{x}| \leq c|x_n|$. We choose large balls centered not too far from the x_n-axis: they show that necessarily $-1 < \sigma \leq 0$. We leave the verification that this is the only restriction as well as the verification for the remaining weights to the interested reader. □

One may construct more weights in A_1 out of these by the following observation: if $\omega_1, \omega_2 \in A_1$ then $\omega_1^\lambda \omega_2^{1-\lambda} \in A_1$ for $0 \leq \lambda \leq 1$. This is a consequence of Hölder's inequality.

Proposition 2.3.12 allows to construct elements in A_p. In particular

$$
(1+|x|)^a (1 + |x| - |x_n|)^b \in A_2
$$

if $n = 3$, $|a|, |b| < 2$, $|a - b| < 1$.

Weights not in A_p

Here we restrict ourselves to $n = 3$. The previous results are not completely satisfactory: g decays like $|x_3|^{-4}$ on the negative n^{th} coordinate axis, which is stronger than necessary for a kernel function of a singular integral. This is the reason that the second part of Proposition 3.3.1 holds.

Proof of Proposition 3.3.1, completion: Suppose the weighted estimates hold for ω_1 and ω_2. Then they are true for the weight $\omega_1^{1-\lambda}\omega_2^{\lambda}$, $0 < \lambda < 1$. This is a consequence of complex interpolation applied to

$$T_z = \omega_1^{z/p}\omega_2^{(1-z)/p} T \omega_1^{-z/p}\omega_2^{-(1-z)/p}.$$

Hence the set of a and b in Proposition 3.3.1 is convex and it suffices to prove the weighted estimates close to the corners of the set. We have $\omega_{a,b} \in A_p$ if $-1 < a, b, a+b < p-1$, by Proposition 2.3.12 and Lemma 3.3.3. This gives four of the corners.

Now let $a = b$ and $0 \le a < 2(p-1)$. The estimate

$$\begin{aligned} \|\chi_{x_3>0} T f\|_{L^p(\omega_{a,b})} &\le \|Tf\|_{L^2((1+|x|)^a(1+|x|-|x_n|)^a)} \\ &\le c\|f\|_{L^2((1+|x|)^a(1+|x|-|x_n|)^a)} \le c\|f\|_{L^p(\omega_{a,b})} \end{aligned}$$

is a consequence of the first part of Proposition 3.3.1 and Lemma 3.3.3.

The truncated operator $\tilde{T}$ whose kernel is zero for $|x-y| \le 1$ is bounded on L^p by Theorem 2.2.9. Moreover $T - \tilde{T}$ is bounded in $L^p(\omega_{a,b})$ for $a, b \in \mathbb{R}$ because the kernel is supported in $|x-y| \le 1$. It remains to show boundedness of $\chi_{x_3<0}\tilde{T}$ in weighted spaces.

We do this by interpolation from endpoint estimates for $p = \infty$ and $p = 1+\varepsilon$. More precisely we define

$$T_z := \omega_{a,a}^{z} \chi_{x_3<0} \tilde{T} \omega_{a,a}^{-z}.$$

The weighted estimate for $\omega_{a,a}$, $0 < a < 2$ in Proposition 3.3.1 follows by complex interpolation if we establish for $\lambda \in \mathbb{R}$, $1 < q < \infty$ and a close to 2

$$\|T_{i\lambda} f\|_{L^q(\mathbb{R}^n)} \le c\|f\|_{L^q(\mathbb{R}^n)} \quad \text{and} \quad \|T_{1+i\lambda} f\|_{L^\infty(\mathbb{R}^n)} \le c\|f\|_{L^\infty(\mathbb{R}^n)}.$$

We have seen that the first estimate holds true and it remains to prove the L^∞ estimate. This reduces to a kernel estimate. It is true if for $x_3 < 0$

$$\begin{aligned} \int_{\mathbb{R}^3} &\left[(1+|x-y|^2)^{-2} + \chi_{\{x_3>y_3\}}(1+|x-y|)^{-2}\left(1+\frac{|\hat{x}-\hat{y}|^2}{x_3-y_3}\right)^{-2}\right] \\ &\times \left[(1+|y|^2)^{-2} + \chi_{\{y_3>0\}}(1+|y|)^{-2}\left(1+\frac{|\hat{y}|^2}{y_3}\right)^{-2}\right]^{a/2} dy \le c(1+|x|^4)^{-a/2}. \end{aligned}$$

We expand the sums, which leads to three terms. The first one is (with a close to 2)

$$I_1(x) = \int_{\mathbb{R}^3} (1+|x-y|^4)^{-1}(1+|y|^4)^{-a/2} dy \le c(1+|x|^4)^{-a/2},$$

which can be seen by splitting the domain of integration into balls of radius $|x|/2$ centered at 0, at x and the complement of their union. The second one

$$\begin{aligned} I_2(x) &= \int_{y_3<x_3} (1+|x-y|)^{-2}\left(1+\frac{|\hat{x}-\hat{y}|^2}{x_3-y_3}\right)^{-2}(1+|y|^2)^{-a}\,dy \\ &= \int_{y_3>0} (1+|x-y|^2)^{-a}(1+|y|)^{-2}\left(1+\frac{|\hat{y}|^2}{|y_3|}\right)^{-2} dy \end{aligned}$$

is larger than the third one

$$I_3(x) = \int_{y_3>0} (1+|x-y|^2)^{-2}(1+|y|)^{-a}\left(1+\frac{|\hat{y}|^2}{y_3}\right)^{-a} dy$$

and it suffices to bound $I_2(x)$. Now

$$\int_{\mathbb{R}^2} (1+|\hat{y}|^2/|y_3|)^{-2} d\hat{y} = |y_3| \int_{\mathbb{R}^2} (1+|\hat{y}|^2)^{-2} d\hat{y}$$

and the bound for $x = x_3 e_3$ follows from

$$\int_0^\infty (1+x_3^2+y_3^2)^{-a}(1+y_3)^{-1} dy_3 \le c(1+|x_3|^2)^{-a}.$$

This proves the claim if $|\hat{x}| \le x_3$. Vice versa: suppose that $x_3 = 0$. We decompose the domain of integration into $J_1 = \{y : |\hat{x}-\hat{y}| \le |\hat{x}|/2\}$ and its complement J_2. The integral over J_2 is estimated exactly as above. The L^∞ estimate is completed by

$$\begin{aligned} \int_{J_1} \dots dy \;&\le |\hat{x}|^{-4} \int_{y_3, |\hat{y}| \le |\hat{x}|} (1+|y|^2)^{-a} dy + |\hat{x}|^{1/2} \int_{|\hat{x}|}^{|\hat{x}|^{5/4}} (1+y_3^2)^{-1-a}\,dy_3 \\ &\quad + |\hat{x}|^2 \int_{|\hat{x}|^{5/4}}^\infty (1+y_3)^{-2-2a}\,dy_3 \\ &\le c(1+|\hat{x}|^2)^{-a}. \end{aligned}$$

for $|\hat{x}| \ge 1$. This gives the weighted estimates of Proposition 3.3.1 if $a = b$ and $-1 < a < 2$.

The endpoints $a = 2(p-1)-1-\varepsilon$, $b = 2(p-1)-\varepsilon$ resp. $a = 2(p-1)-\varepsilon$, $b = 2(p-1)-1-\varepsilon$ follow by interpolation between a modified L^∞ estimate and the L^1 estimate with the weights $(1+|x|)^{-1+\varepsilon}$ and $(1+|x|-x_n)^{-1+\varepsilon}$. The modified L^∞ estimate follows from the same type of arguments. This completes the proof. □

Remark 3.3.4. We shall see that the same estimates hold for the Oseen operator. In fact we have organized the proofs in a way which applies immediately to the Oseen operator. Proposition 3.3.1 has three purposes. It shows that one does not always get optimal conditions for weights - the heuristics behind the theory of weights lead further than the rigorously proven results. Second it extends the range of parameters for which such weighted estimates are known. Since it is easy to produce an immense quantity of weighted estimates with the methods at hand it is crucial to do the comparison with results obtained by other means. The condition $a + b < 4$ (for $p = 2$) is necessary. This can be seen by applying the operator to suitable bump functions.

3.4 The Oseen operator

Consider the problem

$$\begin{aligned} k\partial_1 u - \nu\Delta u + \nabla p &= w \\ \nabla \cdot u = 0. \end{aligned} \tag{3.7}$$

The basic result is the following (using again the notation of the previous section):

Theorem 3.4.1. *Let $T : w \to \partial_{ij} u$. Then the following estimates hold:*

$$\|Tw\|_{L^p(\omega)} \le c\|w\|_{L^p(\omega)}$$

if $\omega \in A_p$, defined as in Proposition 3.3.1 or if $n = 3$ and $p = 2$

$$\omega = (1+|x|)^a (1+s)^b$$

and $-1 < a, b, a+b < 2$, $2b - a, 2a - b < 3$ (see Fig. 3.3).

Proof. Let g be the kernel function of T. We claim that

$$|g(x)| \le c|V(x)|^{-1}$$

$$|g(x) - g(y)| \le c|V(x)|^{-1} \frac{d(x,y)}{d(x,0) + d(y,0)}$$

$$|g(x)| \leq c|x|^{-n-1} \text{ and } |\nabla g(x)| \leq c|x|^{-n-2} \text{ for } x_n < 0.$$

Theorem 3.4.1 is a consequence of these kernel estimates because we only used these estimates for Proposition 3.3.1.

There is an explicit formula for the fundamental solution if $n = 3$. Let $s(x) = |x| - x_1$ and

$$\Phi(x) = \frac{1}{4\pi k} \int_0^{ks/2\nu} \frac{1 - e^{\alpha}}{\alpha} d\alpha.$$

Then (see Oseen [113]),

$$E_{ij}(x) = (\delta_{ij}\Delta - \partial_i\partial_j)\Phi.$$

In that case the kernel estimates can be verified directly. Otherwise the kernel estimates can be obtained from the kernel estimates of the previous section. We omit the details of the kernel estimates for other dimensions. □

Remark 3.4.2. Remark 3.3.4 applies here as well. One obtains the same estimates for exterior domains, see Farwig [55].

3.5 Elliptic operators

We consider various elliptic operators: measurable coefficients, lattice equations, periodic coefficients, continuous coefficients and mixed problems. The previous results and techniques have a strong impact here. We review classical results, explain their relation to singular integrals and weighted estimates, and derive some new results and give alternative proofs for periodic coefficients. Besides its intrinsic interest and importance this part is motivated by the fact that we shall use similar arguments in a different context later.

3.5.1 Measurable coefficients

Let a^{ij} be bounded measurable functions on $\mathbb{R}^n$ with

$$\lambda|\xi|^2 \leq a^{ij}\xi_i\xi_j \quad \text{and} \quad |a^{ij}| \leq \Lambda \quad \text{almost everywhere.}$$

We study weak $H_0^1(\Omega)$ solutions to

$$\partial_i(a^{ij}\partial_j u) = \partial_i f^i + w$$

in the bounded domain $\Omega \subset \mathbb{R}^n$ with smooth boundary Γ. The Lemma of Lax-Milgram ensures existence of a unique weak solution u with

$$\|\nabla u\|_2^2 \leq \lambda^{-1}\|f\|_2^2 + c(\Omega)\|w\|_2.$$

Proposition 3.5.1. *Suppose that $B(x,R)$ is a ball either contained in the interior, or with center x at the boundary, with radius $R<1$, $u \in H_0^1(\Omega)$ with $\partial_i(a^{ij}\partial_j u)=0$ in B. Then the following is true:*

1. **Cacciopoli inequality:** *If the ball B is contained in the interior then*

$$\begin{aligned}\|\nabla u\|_{L^2(B(x,R/2))} &\leq c/R \inf_{\kappa} \|u-\kappa\|_{L^2(B(x,3R/4))} \\ &\leq cR^{-n/p+n/2}\|\nabla u\|_{L^p(B(x,R))}\end{aligned}$$

   if $\min\{1,\frac{2n}{n+2}\}<p\leq 2$. *The* inf *and* κ *are omitted if the ball is centered at the boundary.*

2. **Moser iteration:** *The following estimate holds for $r<R$ and $p>1$*

$$\|u\|_{L^{\frac{p(n+2)}{n}}(B(x,r))} \leq (c(p,\lambda,\Lambda)|R-r|)^{-1/p}\|u\|_{L^p(B(x,R))}$$

3. **BMO:** *If u is positive in B then*

$$\|\nabla \ln u\|_{L^2(B(x,R/2))} \leq cR^{n-1}.$$

   In particular $\ln u \in BMO_{loc}(B)$.

4. *If u is positive on the ball then, if $p<0$*

$$\|u^p\|_{L^{\frac{(n+2)}{n}}(B(x,r))} \leq (c|R-r|)^{-1}\|u^p\|_{L^1(B(x,R))}.$$

We refer to Stein [125] for a definition of BMO.

Sketch of proof: The first part follows from using $\chi(u-c)$ as test function, combined with Poincaré's inequality. For the second and the fourth part we use $\chi|u|^{p-1}$ as test function. The third part follows by testing by χu^{-1}. This is contained in Moser [99]. □

Corollary 3.5.2. *Under the same assumptions as in Proposition 3.5.1 the following is true if $\partial_i(a^{ij}\partial_j u)=\partial_i f^i$ in the weak energy sense:*

1. *There exists $p>2$ such that*

$$\|\nabla u\|_{L^p(\Omega)} \leq c\|f\|_{L^p(\Omega)}.$$

2. *There exists c such that, if u is a nonnegative solution to the homogeneous equation on the interior ball $B(x,R)$ then the Harnack inequality*

$$u(x) \leq c \inf_{B(x,R/2)} u$$

   holds.

3. *There exists* $\alpha > 0$ *such that, if* $\partial_i(a^{ij}\partial_j u) = 0$ *in the ball* $B(x,R)$ *and* $R \le 1$ *then*

$$\|u\|_{C^\alpha(B(x,R/2))} \le cR^{-\alpha-n}\|u\|_{L^1(B(x,R))}.$$

4. *Let* g *be the fundamental solution. Let* α *be as above. Then*
   - $|g(x,y)| \le c|x-y|^{2-n}$
   - $|g(x,y) - g(x,z)| \le c\frac{|y-z|^\alpha}{|x-y|^{n-2+\alpha}}$ *for* $|y-z| \le 1/2|x-y|$
   - $\nabla_x g \in L^{n/(n-1),\infty}$.

Proof. The first part is a consequence of the Cacciopoli inequality, and a reverse Hölder inequality, which is a small modification of Proposition 3.5.1. The reverse Hölder inequality is due to Gehring [64]. The approach to the L^p estimates above can be found in Giaquinta [66]. The Harnack inequality is a consequence of the parts 2-4 of Proposition 3.5.1 and the John-Nirenberg inequality, see Moser [99]. The last part is due to Grüter and Widman, [67]. The space $L^{p,\infty}$ is the Lorentz space or weak L^p_w space. □

Remark 3.5.3. We shall later derive similar results in Section 4 for a different problem.

Equation on a lattice

For simplicity we restrict ourselves to a domain in $\mathbb{Z}^n$. More precisely: let Ω be a bounded domain with Lipschitz boundary in $\mathbb{R}^n$. Let Ω_ε be lattice points of $\varepsilon\mathbb{Z}^n$ in Ω. Let $l^p(\Omega_\varepsilon)$ be the closed subspaces of $l^p(\varepsilon\mathbb{Z}^n)$ of sequences which are zero outside Ω_ε. We define

$$\|u\|_{l^p(\Omega_\varepsilon)} = \varepsilon^{n/p}\|u\|_{l^p(\mathbb{Z}^n)}$$

and the discrete partial derivative $(\partial_i^+ u)(x) = \varepsilon^{-1}(u(x+\varepsilon e_i) - u(x))$ and ∂^- the backwards difference quotient. The discrete second order derivative is defined by $\partial_i^-\partial_i^+$. Let a^{ij} be uniformly elliptic quadratic forms defined on $\varepsilon\mathbb{Z}^n$. Then we consider the problem

$$\sum_{i,j} \partial_i^-(a^{ij}\partial_j^+ u) = \partial_i^+ f^i.$$

Theorem 3.5.4. 1. *There exist* p *and* c *depending only on ellipticity and* Ω *such that*

$$\|\nabla_\varepsilon u\|_p \le c\|f\|_p.$$

Here ∇_ε *denotes the vector of forward differences.*

2. *The discrete kernel satisfies the same estimates as the continuous one in Corollary 3.5.2.*

3. *Harnack and Hölder inequalities hold.*

Remark 3.5.5. The crucial point here is that the constants do not depend on ε. The arguments for the continuous problem carry over. There are interesting related open questions. Let $a^{ij} = \delta^{ij}$. Is it possible to choose $1 < p < \infty$ in the first part of the theorem? Obviously the operator $f \to \nabla u$ is given by a matrix. Are these operators singular integral operators? I.e. do they satisfy the bounds for singular integral operators independently of ε? These questions will occur again in the more complicated context of homogenization below.

3.5.2 Periodic coefficients

Let $G = \mathbb{Z}^n \subset \mathbb{R}^n$ be the lattice of integers, let a^{ij}_{kl} be continuous G-invariant (periodic functions) which satisfy

$$a^{ij}_{kl}\xi^k_i\xi^l_j \geq \kappa|\xi|^2.$$

We study the problem

$$\partial_i(a^{ij}_{kl}\partial_j u^l) = \partial_m f^{mj}.$$

Theorem 3.5.6. *Let $1 < p < \infty$ and $\omega \in A_p$. Suppose that $\ln\omega$ is Lipschitz continuous (i.e. $\ln[\omega(x)/\omega(y)] \leq c|x-y|$) for $|x-y| \leq 1$. Then the following apriori estimate holds for energy solutions:*

$$\|\nabla u\|_{L^p(\omega)} \leq c(p)\|f\|_{L^p(\omega)}.$$

Remark 3.5.7. Lipschitz continuity of $\ln\omega$ can be removed by some extra work. The estimate has been proven if $\omega = 1$ by Avellaneda and Lin [20] under stronger assumptions on the coefficients and by Caffarelli [29] by different methods. Our proof is a rather immediate application of the theory of singular integrals.

Proof. For notational simplicity we only consider scalar problems, taking care however, that all arguments apply as well to systems. By the lemma of Lax-Milgram there exists a unique weak solution (up to the addition of constants) which satisfies

$$\|\nabla u\|_2 \leq c\|f\|_2.$$

The Cacciopoli inequality

$$\|\nabla u\|_{L^2(B_r)} \leq c(R-r)^{-1}\|u\|_{L^2(B_R)} + c\|f\|_{L^2(B_R)}$$

is obtained by using suitable functions χu as test function.

The next result provides the modified kernel estimates. It says that the gradient decays like R^{-1} if we take an L^2 average over sets of size 1.

Lemma 3.5.8. *Suppose that $\partial_i(a^{ij}\partial_j u) = 0$ on a ball of radius R. Let $1 \leq r \leq R/2$. Then*

$$\|u\|_{L^2(B_r)} \leq c\left(\frac{r}{R}\right)^{n/2}\|u\|_{L^2(B_R)},$$

and

$$\|\nabla u\|_{L^2(B_r)} \leq c\left(\frac{r}{R}\right)^{n/2} R^{-1}\|u\|_{L^2(B_R)}.$$

Proof. We may assume that $R \geq R_0$ where R_0 is at our disposal. Let $\gamma < 1$ be a real number close to 1. Clearly $v = u(.+e_i) - u$ is again a solution. It satisfies by the gradient estimate (using the Cacciopoli inequality)

$$\|\nabla v\|_{L^2(B(x,\gamma^2 R))} \leq cR^{-1}\|v\|_{L^2(B(x,\gamma R))} \leq cR^{-2}\|u\|_{L^2(B(x,R))}.$$

This can be iterated. Let ∂_i^+ be the forward difference operator in direction i and D^α the corresponding iterated operator where α is a multi-index. Then

$$\|D^\alpha \nabla u\|_{L^2(B(x,R/2))} \leq R^{-|\alpha|-1}\|u\|_{L^2(B(x,R))}.$$

The estimates are now a consequence of the discrete Sobolev inequality (which is proven in the same way as (3.1) with functions taking values in L^2 of the unit cube). $\square$

Let $Q \subset \mathbb{R}^n$ be the unit cube. We identify $L^2(\mathbb{R}^n)$ and $l^2(\mathbb{Z}^n, L^2(Q))$ via the invertible map $(Su)_\alpha(y) = u(y+\alpha)$ where $y \in Q$ and $\alpha \in Z^n$. In that way one obtains a matrix representation of $T^{ij} : f^i \to \partial_j u$ with entries labelled by $(\alpha, \beta) \in (\mathbb{Z}^n)^2$ of the form $T^{ij}_{\alpha\beta} : L^2(Q) \to L^2(Q)$.

Lemma 3.5.9. *The following estimates hold for $\alpha, \beta \in \mathbb{Z}^n$*

$$\|T_{\alpha\beta}\|_{L((L^2(Q))^N)} \leq c(1+|\alpha-\beta|)^{-n}$$

and, if $|\alpha-\alpha'| + |\beta-\beta'| < |\alpha-\beta|/4$,

$$\|T_{\alpha\beta} - T_{\alpha'\beta'}\| \leq c\frac{|\alpha-\alpha'|+|\beta-\beta'|}{(1+|\alpha-\beta|)^{n+1}}.$$

Proof. This is more or less a reformulation of the L^2 estimates coming from the lemma of Lax-Milgram and the estimates of Lemma 3.5.8, applied with respect to x. □

We collect our information: the set $\mathbb{Z}^n$ carries the metric as subset of $\mathbb{R}^n$, the counting measure is doubling and T^{ij} is a singular integral operator. A check reveals that the results on singular integral operators remain true if the functions assume values in a Hilbert space and the kernel is operator valued. Hence

$$T^{ij} : l^p(G, L^2(Q)) \to l^p(G, L^2(Q))$$

for $1 < p \leq 2$ and by duality the same holds for $1 < p < \infty$.

Let $\omega \in A_p(\mathbb{R}^n)$ be such that $\ln \omega$ is Lipschitz continuous. We define $\bar{\omega}(\alpha) = \omega(\alpha)$. Then $\bar{\omega} \in A_p(\mathbb{Z}^n)$ with respect to the counting measure and

$$T : l^p(G, \bar{\omega}; L^2(Q)) \to l^p(G, \bar{\omega}; L^2(Q))$$

if $1 < p < \infty$. Up to here we did not need continuity of the coefficients. The proof is complete once we show that we can replace $L^2(Q)$ by $L^p(Q)$. We complete the proof after a discussion of elliptic operators with continuous coefficients.

3.5.3 Continuous coefficients

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary Γ. We consider again weak H^1_0 solutions to

$$\partial_i(a^{ij}\partial_j u) = \partial_i f^i$$

where $a^{ij} \in C(\overline{\Omega})$, $a^{ij}\xi_i\xi_j \geq \kappa|\xi^2|$ with $\kappa > 0$. We assume that the coefficients are continuous. We claim that

$$\|\nabla u\|_{L^p(B_{R/2})} \leq c(R^{n/p-n+1}\|u\|_{L^1(B_R)} + \|f\|_{L^p(B_R)}) \tag{3.8}$$

for $R \leq 1$ and $2 \leq p < \infty$ where the constant depends on the modulus of continuity.

We recall that the estimate

$$\|\nabla u\|_p \leq \kappa(p)\|f\|_p$$

holds for constant coefficients. We fix the smallest κ for the frozen coefficients $a^{ij}(x)$. There exists δ such that

$$(a^{ij}(x) - a^{ij}(y))\xi_i\eta_j \leq (2\kappa)^{-1}|\xi||\eta|$$

if $|x-y| \le \delta$. Now, with χ a suitable cutoff function,

$$\begin{aligned}\partial_i(a^{ij}(x_0)\partial_j(\chi u)) &= \partial_i(\chi f^i) - (\partial_i \chi) f^i + \partial_i \left[(a^{ij}(x_0) - a^{ij})\partial_j(\chi u) \right] \\ &\quad + \partial_i \left[a^{ij}(\partial_j \chi) u \right] + (\partial_i \chi) a^{ij} \partial_j u\end{aligned}$$

hence

$$\|\nabla(\chi u)\|_{L^p} \le \frac{1}{2}\|\nabla(\chi u)\|_{L^p} + c(R)\left(\|f\|_{L^p(B_{2R})} + \||\nabla\chi|u\|_{L^p} + \||\nabla\chi|\nabla u\|_{L^{\tilde p}}\right)$$

where $\max\{1, \frac{pn}{n+p}\} < \tilde p \le p$. We have used the decay of the fundamental solution with constant coefficients and the weak Young inequality. This argument can be iterated. Together with Corollary 3.5.2 and scaling arguments we arrive at (3.8). Similar results are true for systems with continuous coefficients. There we only obtain the weaker estimate

$$\|\nabla u\|_{L^p(B_{R/2})} \le c(R^{n/p-n/2-1}\|u\|_{L^2(B_R)} + \|f\|_{L^p(B_R)}).$$

Proof of Theorem 3.5.6, completion:
The gradient estimate of inequality (3.8) shows that

$$\begin{aligned}\|T^{ij} f\|_{l^p(\bar\omega;L^p(Q))} &\le c_1\Big(\|T^{ij} f\|_{l^p(\bar\omega;L^2(Q))} + \|f\|_{l^p(\bar\omega,L^2(Q))} \\ &\le c_2 \|f\|_{l^p(\bar\omega;L^p(Q))}\end{aligned}$$

if $2 \le p < \infty$. Hence

$$\|T^{ij} f\|_{L^p(\mathbb{R}^n,\omega)} \le c\|f\|_{L^p(\mathbb{R}^n,\omega)}.$$

The case $1 < p \le 2$ follows by duality. □

3.6 The Kohn-Laplacian

We equip $\mathbb{R}^{2n+1}$ with a particular group structure. Let $[x,y,t] \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ be a typical point. Then

$$[x,y,t][\tilde x,\tilde y,\tilde t] = [x+\tilde x, y+\tilde y, t+\tilde t + 2(x\tilde y - y\tilde x)]$$

defines a group structure. This group is the Heisenberg group.

The boundary of the unit ball in $S = \mathbb{C}^{n+1}\backslash(0,\dots,0,1)$ can be identified with the set of biholomorphic mappings of the open unit ball to itself which extend to a homomorphism of the closed ball which keeps the point $(0,\dots,0,1)$ fixed. There is a group isomorphism of the Heisenberg group to S. This relates the Heisenberg group to complex analysis of several variables. It is moreover in some sense the simplest and most typical example of boundary of a so-called strongly pseudoconvex domain in $\mathbb{C}^{n+1}$. As such it carries a so-called Cauchy-Riemann structure.

There are the left invariant vector fields $X_i = \partial_{x_i} + 2y_i\partial_t$, $Y_i = \partial_{y_i} - 2x_i\partial_t$ and $T = \partial_t$ with commutators $[X_i, Y_j] = 4\delta_{ij}T$. One easily calculates

$$\begin{aligned}\Delta_\alpha &:= -\frac{1}{4}\sum_j (X_j^2 + Y_j^2) + i\alpha T \\ &= -\frac{1}{4}\Delta_{\mathbb{R}^{2n}} - (|x|^2 + |y|^2)\partial_t^2 + x_i\partial_{y_i}\partial_t - y_i\partial_{x_i}\partial_t + i\alpha\partial_t.\end{aligned}$$

The geodesics are defined by the Hamiltonian function

$$H(x, y, t, p_x, p_y, \tau) = \frac{1}{4}(|p_x|^2 + |p_y|^2) + (|x|^2 + |y|^2)\tau^2 + x_i p_{y_i}\tau - y_i p_{x_i}\tau$$

and there is an associated distance defined in a natural way. The figures (3.2–3.5) show the singular set of the fundamental solution of the wave equation on the Heisenberg group. The outer part is the boundary of the geodesic ball, see Nachman [106].

The associated measure is the Lebesgue measure. Then, according to Stein, Theorem 13.2.1 in [125],

$$\Delta_\alpha\Big[\left(|x|^2 + |y|^2 - it\right)^{-\frac{n+\alpha}{2}}\left(|x|^2 + |y|^2 + it\right)^{-\frac{n-\alpha}{2}}\Big] = \kappa\delta$$

where

$$\kappa = \frac{2^{2-n}\pi^{n+1}}{\Gamma([n+\alpha]/2)\Gamma([n-\alpha]/2)\delta_0}.$$

Hence the fundamental solution is $g(x, y) = g_0(y^{-1}x, 0)$ where g is $1/\kappa$ times the function on the left hand side if $\kappa \neq 0$. If $\alpha = \pm(n + 2k)$ for some nonnegative integer k then there exists no fundamental solution. It is not hard to see that $D^2_{X,Y}g$ and $\partial_T g$ define singular integral operators with respect to a suitable metric.

3.7 A Kolmogorov equation

Here we consider a stochastic process which describes a particle driven by a random force. The momentum is then a Brownian motion and the position is obtained by integrating the velocity. The Kolmogorov equation is

$$Lu := \frac{\partial^2 u}{\partial x^2} + x\frac{\partial u}{\partial y} - \frac{\partial u}{\partial t} = \delta_{x_0,0,0}. \tag{3.9}$$

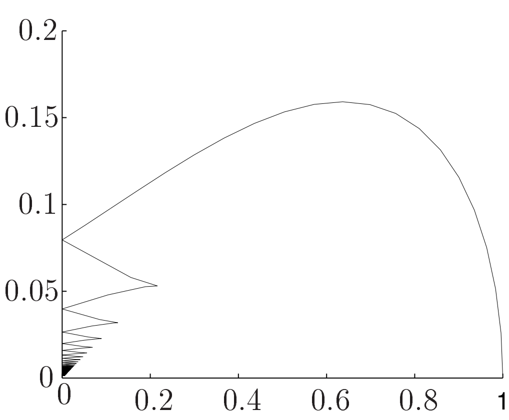

Figure 3.2: Endpoints of geodesics of length one

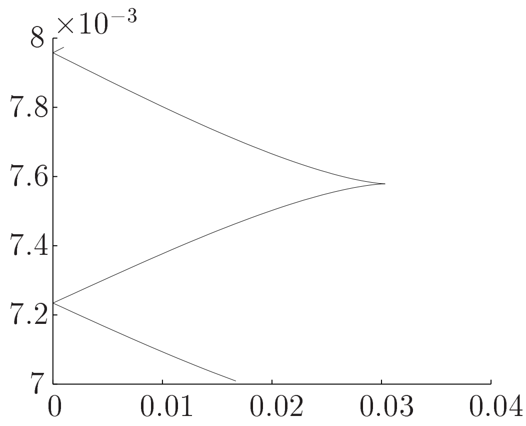

Figure 3.3: The tenth tip

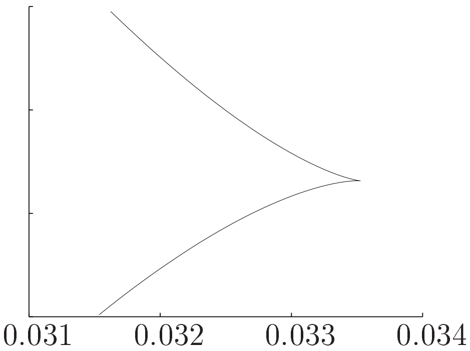

Figure 3.4: Details

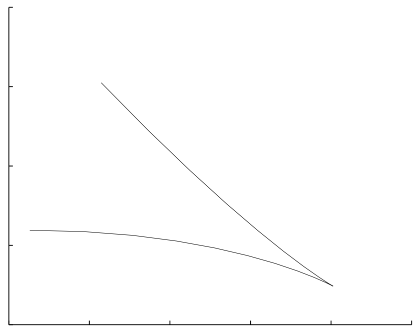

Figure 3.5: Finer details

Here x denotes the velocity and y the position.

This is a particular case of squares of vector fields: The two vector fields $X = \partial_x$ and $Y = -\partial_t + x\partial_y$ satisfy

$$Z = [X, Y] = \partial_y$$

and X, Y, Z span $\mathbb{R}^3$. Hence the Hörmander condition (see [69]) is satisfied and the operator L is subelliptic.

The kernel

On physical grounds it is reasonable to change coordinates to $\tilde{y} = y + xt$. With $\tilde{u}(x, \tilde{y}, t) := u(x, \tilde{y} - xt, t)$ the equation has the form $\tilde{u}_t = (\partial_x + t\partial_{\tilde{y}})^2\tilde{u}$. The Fourier transform of the fundamental solution at fixed time t is

$$\hat{u}(t, s, \xi + \eta t, \eta) = e^{-t\xi^2 - t^2\xi\eta - \frac{1}{3}t^3\eta^2}$$

hence

$$
\begin{aligned}
u(x, y-tx, t) &= \frac{\sqrt{3}}{2\pi t^2} exp\left(-\frac{(x-x_0)^2}{t} + 3\frac{(x-x_0)(y-y_0)}{t^2} - 3\frac{(y-y_0)^2}{t^3}\right)\\
&= \frac{\sqrt{3}}{2\pi t^2} exp\left(-\frac{(x-x_0)^2}{4t} - 3\frac{(y-y_0-\frac{t(x-x_0)}{2})^2}{t^3}\right)
\end{aligned}
$$

and

$$
u(x,y,t) = \frac{3^{1/2}}{2\pi t^2} exp\left(-\frac{|x-x_0|^2}{4t} - 3\frac{(y-y_0-\frac{1}{2}t(x+x_0))^2}{t^3}\right).
$$

This computation follows Hörmander, [70] p211.

The metric

We define the quasi–metric

$$
d_0((x,y,t);(\bar{x},\bar{y},\bar{t})) := \max\left\{|x-\bar{x}|, \sqrt{|t-\bar{t}|}, |y-\bar{y}-(t-\bar{t})(x+\bar{x})/2|^{1/3}\right\}.
$$

We have to check the triangle inequality, which reduces to

$$
\begin{aligned}
&|y-\tilde{y}-(t-\tilde{t})(x+\tilde{x})/2|\\
&\quad \leq \left|y-\bar{y}-(t-\bar{t})(x+\bar{x})/2\right| + \left|\tilde{y}-\bar{y}-(\tilde{t}-\bar{t})(\tilde{x}+\bar{x})/2\right|\\
&\qquad + \left|(t-\bar{t})(x+\bar{x})-(t-\tilde{t})(x+\tilde{x})-(\tilde{t}-\bar{t})(\tilde{x}+\bar{x})\right|
\end{aligned}
$$

and

$$
\begin{aligned}
&|(t-\bar{t})(x+\bar{x})-(t-\tilde{t})(x+\tilde{x})-(\tilde{t}-\bar{t})(\tilde{x}+\bar{x})|\\
&\qquad = |(t-\tilde{t})(\bar{x}-\tilde{x})-(\bar{t}-\tilde{t})(x-\tilde{x})|
\end{aligned}
$$

hence d_0 is a quasi–metric. Let $X=(x,y,t)$ and $X_\lambda=(\lambda x, \lambda^3 y, \lambda^2 t)$. Then $d(X_\lambda, Y_\lambda) = \lambda d(X_1, Y_1)$. There is an equivalent metric: Let γ be a smooth curve. We parametrize it over $[0,1]$ so that

$$
\sqrt{|\gamma_t'| + |\gamma_x'|^2 + |x||\gamma_y'|} = k
$$

is constant. Then we define the length of the curve by k. This defines a metric d on $\mathbb{R}^3$. One may check that $d(X_\lambda, Y_\lambda) = \lambda d(X_1, Y_1)$ and $d(X,Y) \sim d_0(X,Y)$.

It is not hard to see that $\|\partial_x u\|_2 \leq c\|f\|_{L^2(\mathbb{R}^3)}$ if $Lu = w$ and if $w = \partial_x f$. Moreover the kernel functions k given by

$$\partial_x \partial_{\bar{x}} g(x, y, t, \bar{x}, \bar{y}, \bar{t}),\ \partial_x^2 g,\ (\partial_t - x\partial_y)g$$

all satisfy the kernel estimates (namely $|k| \leq cV(x, y, t, \bar{x}, \bar{y}, \bar{s})^{-1}$ and the Dini condition). The estimates

$$\|u_t - x\partial_y u\|_{L^2(\mathbb{R}^3)} + \|\partial_{xx}^2 u\|_{L^2(\mathbb{R}^3)} \leq c\|w\|_{L^2(\mathbb{R}^2)}$$

are special cases of Theorem 16 in Rothschild and Stein [118], where these estimates are proven for compactly supported u. Scaling gives then the global estimate. Hence, if $\omega \in A_p$ with respect to d then

$$\|u_t - x\partial_y u\|_{L^p(\omega)} \leq c\|Lu\|_{L^p(\omega)}$$

and

$$\|u_{xx}\|_{L^p(\omega)} \leq c\|Lu\|_{L^p(\omega)}.$$

Examples of weights in A_p are the following: $|x|^\sigma$, $|t|^\sigma$ and $(|x|^2 + |t|)^\sigma$ for $-1 < \sigma < p - 1$. One may ask whether one obtains similar estimates for $\partial_t u$ in terms of the right hand side. . The best estimate of u_t in terms of $\|Lu\|_{L^2}$ and $\|u\|_{L^2}$ is contained in Theorem 16 of [118]: there is a loss of 2/3 derivatives.

3.8 Notes

The characterization for L^p multipliers can be improved. The question, how much regularity one needs for a function to be a Fourier multiplier in L^p is a deep one. Particular cases are related to the most famous problems in analysis like the boundedness of Bochner-Riesz operators, restriction theorem and local smoothing for the wave equation. The survey paper Müller [103] contains a discussion of these questions, together with many references.

The study of elliptic operators with drift is motivated by discussions and hints by Farwig (see also [54]), which the author thankfully acknowledges. The hydrodynamic potentials for the Oseen operator have been studied by Bemelmans [23] and Faxen [56] without weights. A smaller family of weights has been considered by Farwig [55]. The weighted estimates of Solonnikov [123] are of a different nature. The importance here is that the weights are closely related to the fundamental solution. The Navier-Stokes equations with drift at infinity still pose subtle problems. See Galdi, Heywood and Shibata [62] for a related and surprisingly difficult problem.

The estimates for the fundamental solution to elliptic problems with measurable coefficients as well as the techniques are taken from Moser [99] and Grüter and Widman [67]. The relation between periodic coefficients and singular integrals in the form presented here seems to be new.

The Heisenberg group is a model for operators in complex analysis and nilpotent groups. The existence of explicit fundamental solutions on the Heisenberg group allowed us to avoid many technicalities, while it is still exhibiting the essence of (part of) the problem. It is needless to say that we lost at the same time many of the beautiful aspects of the theory nilpotent groups and the Kohn-Laplace operator. The interested reader should consult Christ [36], d'Angelo [46], Stein [125] and Varopoulos, Saloff-Coste and Coulhon [133] and the references therein for more information.

The Kolmogorov equation has been introduced by Kolmogorov [84]. It is a particular case of the general subelliptic problems. Again existence of an explicit fundamental solution allows us to avoid technical difficulties.

Chapter 4

Gaussian estimates

4.1 Introduction

The heat kernel

$$g(x,t,y,s) = \begin{cases} \dfrac{1}{(4\pi(t-s))^{n/2}} e^{-\frac{|x-y|^2}{4(t-s)}} & \text{if } t > s \\ 0 & \text{otherwise,} \end{cases}$$

together with its generalizations is among the most interesting functions in partial differential equations. It links geometric and analytic properties, which is obvious if we write it in the form

$$g(x,t,y,s) = \begin{cases} \dfrac{1}{c(n)|B(x,t^{1/2})|} e^{-\frac{d(x,y)^2}{4(t-s)}} & \text{if } t > s \\ 0 & \text{otherwise,} \end{cases} \tag{4.1}$$

where $|B(x,R)|$ denotes the Lebesgue-measure of the ball around x of radius R and d denotes the distance in $\mathbb{R}^n$. It also links the heat equation to the theory of singular integral operators.

Another motivation, besides problems in geometry and partial differential equations, to study heat kernel estimates comes from the interpretation as transition probabilities in stochastic processes: the heat equation is the (forward) Kolmogorov equation of a Wiener process.

For uniformly parabolic equations with measurable coefficients a somewhat weaker statement remains true: the fundamental solution is bounded from above and below by terms of the type of the right hand side of (4.1). These estimates were first obtained by Aronson [7]. They are intimately related to the Harnack inequality and Hölder continuity of solutions. This is a crucial point in the regularity theory of quasi-linear parabolic equations: the derivative of a solution satisfies a parabolic equation with measurable coefficients, which, under suitable assumptions, is uniformly parabolic. Thus first derivatives of solutions are Hölder continuous. Further regularity can be derived by the device of frozen coefficients as soon as one establishes estimates for the fundamental solution to model problems. To do this one may start once more with kernel estimates and combine them with estimates for derivatives of solutions. These estimates, together with simpler L^2 estimates, are exactly the estimates required for singular integral operators and then one may apply the whole machinery of singular integral operators and weights.

It is no surprise that these concepts suggest and admit a vast number of generalizations and variations. The requirement of uniform positive-definiteness can be considerably relaxed. Under suitable, much weaker assumptions all we said remains true. The generalization of the Euclidean theory of singular integral operators applies in particular to kernels occuring in the connection with sub-elliptic differential equations: The metric is the Carnot-Carathéodory metric and the measure is a natural measure of the object under study.

We shall follow the strategy suggested by this introduction. First we derive kernel estimates. The kernel estimates are estimates from above and below of the type (4.1), interpreted in the natural measure and metric. To get these estimates we need some Hardy-Sobolev inequalities. Then we discuss metric and measure in Section 4.3. After these preparations we turn to the kernel estimates themselves. Here we choose the approach of Fabes and Stroock [53] for uniformly parabolic equations, which is closer to the early paper by Nash [108] than to Moser's [100] or de Giorgi's [50] approach.

In the second part of this chapter we use the kernel estimates to derive supremum estimates, a Harnack inequality, and Hölder regularity of solutions to linear equations with measurable coefficients. The second application of the kernel estimates is to the model operator (4.43) and to mapping properties.

Our particular problem fits into the general picture of the relation between metric, weights, geometry, heat kernel and operator estimates. A unique feature is the appearance of a one-parameter family of natural weights.

4.2 Weighted Sobolev inequalities

In this section we study Sobolev inequalities in $\mathbb{R}^n$ and weighted Hardy-Sobolev inequalities.

Lemma 4.2.1. *Let $\|\nabla u\|_{L^p} = \| \, |\nabla u| \, \|_{L^p}$. The following inequality holds for $u \in W^{1,1}(\mathbb{R}^n)$:*

$$\|u\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq c_n \|\nabla u\|_{L^1(\mathbb{R}^n)}.$$

Proof. This is the standard Sobolev inequality. The optimal constant c_n is the constant in the isoperimetric inequality, i.e. $c_n = \frac{|B(0,1)|^{\frac{n-1}{n}}}{|\mathbb{S}_1^{n-1}(0)|_{H^{n-1}}} \leq \frac{1}{2}$ where H^{n-1} denotes the $n-1$-dimensional Hausdorff measure and $\mathbb{S}_1^{n-1}(0)$ is the $n-1$-dimensional sphere. We sketch a proof with the sharp constant and a simpler second one which gives a sharp constant, however for a different inequality. The estimate by $\frac{1}{2}$ is a consequence of the simpler estimate. We may clearly assume that $u \geq 0$ and $\|u\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} < \infty$. The sets $\{x \in \mathbb{R}^n : u(x) > \lambda\}$ are measurable since u is measurable. By Tschebyscheff's inequality

$$|\{x \in \mathbb{R}^n : u(x) > \lambda\}| \leq \lambda^{-\frac{n}{n-1}} \|u\|^{\frac{n}{n-1}}_{L^{\frac{n}{n-1}}(\Omega)}$$

for all $\lambda > 0$. Let u^* be the (Schwarz)-symmetrization of u. By this we mean the unique radial and radially decreasing function which satisfies

$$|\{x \in \mathbb{R} : u^*(x) > \lambda\}| = |\{x \in \mathbb{R}^n : u > \lambda\}|.$$

The function u^* is measurable by construction. Moreover $\|u^*\|_p = \|u\|_p$ and $\|\nabla u^*\|_p \leq \|\nabla u\|_p$ for $1 \leq p \leq \infty$ for which we refer to Kawohl [78]. Hence it suffices to prove the estimate for radial and radially decreasing nonnegative functions. Any such function can be approximated by step functions of the same type (for which the L^1 norm has to be replaced by the mass of a measure defined by the distributional gradient). It suffices to prove the estimate for step functions. It is not hard to see that the worst case is the characteristic function of the ball, in which case the constant is obtained.

A second possibility is as follows: we shall prove

$$\|u\|^n_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq 2^{-n} \prod_{i=1}^{n} \|\partial_i u\|_{L^1(\mathbb{R}^n)}. \tag{4.2}$$

We prove (4.2) by induction on the dimension. The case $n = 1$ is obvious. Suppose inequality (4.2) is true for $n \leq k-1$. Then

$$\begin{aligned}
\|u\|^{\frac{k}{k-1}}_{L^{\frac{k}{k-1}}(\mathbb{R}^k)} &\leq \int_{\mathbb{R}} \|u(x_1,.)\|^{\frac{1}{k-1}}_{L^1(\mathbb{R}^{k-1})} \|u(x_1,.)\|_{L^{\frac{k-1}{k-2}}(\mathbb{R}^{k-1})} \, dx_1 \\
&\leq \sup_{x_1} \|u(x_1,.)\|^{\frac{1}{k-1}}_{L^1(\mathbb{R}^{k-1})} \int_{\mathbb{R}} \|u(x_1,.)\|_{L^{\frac{k-1}{k-2}}(\mathbb{R}^{k-1})} \, dx_1 \\
&\leq 2^{-\frac{1}{k-1}} \|\partial_1 u\|^{\frac{1}{k-1}}_{L^1(\mathbb{R}^k)} \int_{\mathbb{R}} \left(2^{-(k-1)} \prod_{i=2}^{k} \|\partial_i u(x_1,.)\|_{L^1(\mathbb{R}^{k-1})} \right)^{1/(k-1)} dx_1.
\end{aligned}$$

where we used Hölder's inequality and the induction hypothesis. We take the power $k-1$ of the inequality, apply Hölder's inequality and arrive at

$$\|u\|^k_{L^{\frac{k}{k-1}}(\mathbb{R}^k)} \le 2^{-k} \prod_{i=1}^k \|\partial_i u\|_{L^1(\mathbb{R}^k)}.$$

Inequality (4.2) is sharp: we get equality for $e^{-|x|^2}$ and for all products of smooth functions of each coordinate with exactly one critical point. □

We shall need a weighted version of Sobolev's inequality. Let H be the half space $\{x \in \mathbb{R}^n : x_n \ge 0\}$ and let $C_0^\infty(H)$ be the space of restriction of functions in $C_0^\infty(\mathbb{R}^n)$ to H.

Theorem 4.2.2. *Let $1 \le p \le q < \infty$, $s > -1/p$ and we assume that $\sigma \le 1$ satisfies*

$$\frac{1}{q} - \frac{1-\sigma}{n} = \frac{1}{p}.$$

Then

$$\|x_n^s u\|_{L^p(H)} \le c(n,s,p,q)\|x_n^{s+\sigma}\nabla u\|_{L^q(H)}$$

for the closure of $C_0^\infty(H)$ in the norm of the right hand side.

Remark 4.2.3. The condition on s is sharp. The left hand side diverges for $s \le -1/p$ for continuous functions which do not vanish at the boundary of H.

Proof. We begin with the case $q = 1$. Let $u \in C_0^\infty(H)$ and $v(x) = u(.,2x_n)$. Let $R > 0$, $A = \{x : R < x_n < 2R\}$ and $B = \{x_n : R < x_n < 4R\}$. Then

$$\|x_n^s(v-u)\|_{L^{\frac{n}{n-1}}(A)} \le c\|x_n^s \nabla u\|_{L^1(B)} \tag{4.3}$$

because, with $w = u - v$ and $L^p = L^p(\mathbb{R}^{n-1})$

$$\begin{aligned}\|x_n^s w\|^{\frac{n}{n-1}}_{L^{\frac{n}{n-1}}(A)} &\le \int_R^{2R} \|x_n^s w(.,x_n)\|_{L^1}^{\frac{1}{n-1}} \|x_n^s w(.,x_n)\|_{L^{\frac{n-1}{n-2}}}\, dx_n \\ &\le \sup_{R<x_n<2R} \|x_n^s w(.,x_n)\|^{\frac{1}{n-1}}_{L^1(\mathbb{R}^{n-1})} \int_R^{2R} x_n^s \|w(.,x_n)\|_{L^{\frac{n-1}{n-2}}}\, dx_n.\end{aligned}$$

Inequality 4.3 follows from $\|w(.,x_n)\|_{L^{\frac{n-1}{n-2}}} \le \|\nabla u(.,x_n)\|_{L^1} + \|\nabla u(.,2x_n)\|_{L^1}$ and

$$\|x_n^s w(.,x_n)\|_{L^1} \le x_n^s \int_{x_n}^{2x_n} \|\partial_{x_n} u(z)\|_{L^1}\, dz$$

in the same way as for Lemma 4.2.1. The second endpoint is a Hardy-type inequality. Let w be as above. Then

$$\begin{aligned} \|x_n^s w\|_{L^1(A)} &\le \left\| x_n^s \int_{x_n}^{2x_n} \partial_{x_n} u(z)\, dz \right\|_{L^1(A)} \\ &\le \int_R^{4R} \|\partial_{x_n} u(.,z)\|_{L^1} \int_{z/2}^{z} x_n^s dx_n\, dz \\ &\le \frac{1}{s+1} \|x_n^{s+1}\partial_{x_n} u\|_{L^1(B)}. \end{aligned} \tag{4.4}$$

Let $1 < p < \frac{n}{n-1}$, $s > -1/p$ and $\sigma = 1 + \frac{n}{p} - n$ and $A_i = \{x : 2^i \le x_n \le 2^{i+1}\}$. Then

$$\|x_n^s w\|_{L^p(H)} \le \sum_{i \in \mathbb{Z}} \|x_n^s w\|_{L^p(A_i)}$$

and

$$\|x_n^s w\|_{L^p(A_i)} \le c \|x_n^s w\|^{\sigma}_{L^1(A_i)} \|x_n^s w\|^{1-\sigma}_{L^{\frac{n}{n-1}}(A_i)} \le c \sum_{j=i}^{i+1} \|x_n^{s+\sigma} \nabla u\|_{L^1(A_j)}$$

by inequalities (4.3) and (4.4) (restricted to A_i). Summation implies

$$\|x_n^s w\|_{L^p(H)} \le c \|x_n^{s+\sigma} \nabla u\|_{L^1(H)}.$$

Hence

$$\begin{aligned} \|x_n^s u\|_{L^p(H)} &\le \|x_n^s v\|_{L^p(H)} + c\|x_n^{s+\sigma}\nabla u\|_{L^1(H)} \\ &= 2^{-s-1/p} \|x_n^s u\|_{L^p(H)} + c \|x_n^{s+\sigma} \nabla u\|_{L^1(H)}. \end{aligned}$$

We subtract $2^{-s-1/p}\|.\|_{L^p(H)}$ from both sides and arrive at the desired conclusion,

$$\|x_n^s u\|_{L^p(H)} \le c \|x_n^{s+\sigma} \nabla u\|_{L^1(H)}.$$

The general case is a rather immediate consequence. Let p, q, s and σ as in the theorem and

$$\frac{1}{r} := \sigma + (1-\sigma)\frac{n-1}{n} = 1 + \frac{1}{p} - \frac{1}{q}.$$

Then $1 \le r \le \frac{n}{n-1}$ and with $L^p = L^p(H)$,

$$\|x_n^s u\|_{L^p}^{p/r} = \|x_n^{sp/r} |u|^{p/r}\|_{L^r} \le c \|(x_n^s |u|)^{p/r-1} x_n^{s+\sigma} \nabla u\|_{L^1} \le c \|x_n^s u\|_{L^p} \|x_n^{s+\sigma}\nabla u\|_{L^q}.$$

because $(\frac{p}{r} - 1)\frac{q-1}{q} = p$. $\square$

We shall use the following consequence

Lemma 4.2.4. *Suppose that $s > -1$ and $n + s > 1$. Then*

$$\|u\|^{2+\frac{2}{n+s}}_{L^{2+\frac{2}{n+s}}(x_n^s)} \leq c\|u\|^{\frac{2}{n+s}}_{L^2(x_n^s)}\|\nabla u\|^2_{L^2(x_n^{s+1})}.$$

Proof. It follows from Theorem 4.2.2 that

$$\|u\|_{L^{2\frac{n+s}{n+s-1}}(x_n^s)} \leq c\|\nabla u\|_{L^2(x_n^{s+1})}.$$

The statement follows then by Hölder's inequality:

$$\|u\|^{2+\frac{2}{n+s}}_{L^{2+\frac{2}{n+s}}(x_n^s)} \leq c\|u\|^{\frac{2}{n+s}}_{L^2(x_n^s)}\|u\|^2_{L^{2\frac{n+s}{n+s-1}}(x_n^s)}.$$

□

Remark 4.2.5. The condition $n + s > 1$ is always satisfied if $n \geq 2$. It is likely that the conclusion holds without this restriction.

4.3 Geometry

We equip the half plane H with the Riemannian metric

$$< v, w >_x = x_n^{-1}(v, w).$$

The metric of H is invariant with respect to orthogonal transformations of $\mathbb{R}^{n-1}$ and it scales as follows: $< \lambda v, \lambda w >_{\lambda x} = \lambda < v, w >_x$. Geodesics γ which are parametrized by arc-length satisfy the Euler-Lagrange equations of

$$\frac{1}{2}\int \gamma_n^{-1}|\gamma'|^2 ds.$$

The function $E = \frac{1}{2}\gamma_n^{-1}|\gamma'|^2$ is constant on geodesics parametrized by the arclength s. E is equal to $1/2$ because γ is parametrized by arc-length. The Euler-Lagrange equations are

$$0 = \left(\gamma_n^{-1}\gamma_n'\right)' + \frac{1}{2}\gamma_n^{-1} \tag{4.5}$$

and, for $1 \leq i \leq n-1$, $0 = (\gamma_n^{-1}\gamma_i')'$ which is equivalent to (if $\gamma_n \neq 0$)

$$\gamma_i' = \eta_i\gamma_n \tag{4.6}$$

for some $(\eta_i) \in \mathbb{R}^{n-1}$. We define $\eta = \sqrt{\sum_{i=1}^{n-1} \eta_i^2}$. Then

$$1 = \gamma_n^{-1}((\gamma_n')^2 + (\gamma_i')^2) = \gamma_n^{-1}\left[(\gamma_n')^2 + \eta^2\gamma_n^2\right]$$

hence $(\gamma_n')^2 = \gamma_n - \eta^2\gamma_n^2$. One easily checks that

$$\gamma_n(s) = \frac{1}{\eta^2}\sin^2\left(\frac{\eta}{2}(s-k)\right) = \frac{1}{2\eta^2}\left[1 - \cos(\eta(s-k))\right] \tag{4.7}$$

is the general solution. We integrate (4.6):

$$\begin{aligned} \gamma_i(s) &= \gamma_i(0) + \frac{\eta_i}{2\eta^2}\int_0^s [1 - \cos(\eta(s-k))]ds \\ &= \gamma_i(0) + \frac{\eta_i}{2\eta^2}\left(s - \frac{1}{\eta}\left[\sin(\eta(s-k)) - \sin(-\eta k)\right]\right). \end{aligned} \tag{4.8}$$

For any two interior points $x = (\hat{x}, x_n)$ and $y = (\hat{y}, y_n)$ there exists exactly one geodesic connecting them which does not touch the boundary. This geodesic realizes the distance, which can easily be seen by looking at γ_n.

If $\eta = 0$ then the differential equation simplifies: $\gamma_n' = \pm\sqrt{\gamma_n}$ hence $\gamma_n(s) = \frac{1}{4}(s-k)^2$ and the distance between two points x, y with $x - y \in \mathbb{R}e_n$ is

$$2|x_n^{1/2} - y_n^{1/2}|. \tag{4.9}$$

If $k = 0$ and $s = 2\pi/\eta$ then the geodesic begins and ends at the boundary. Let $|.|$ be the Euclidean distance. Evaluation of (4.8) shows that

$$|\gamma(2\pi/\eta) - \gamma(0)| = \frac{\pi}{\eta^2}$$

and hence the distance between these two points is $2\pi/\eta$. Thus, if x and y are at the boundary the distance is

$$d(x,y) = 2\sqrt{\pi|x-y|}. \tag{4.10}$$

We shall use this exact information to obtain manageable estimates from above and below of the distance. The first step is contained in the next lemma. We use the notation $\hat{x} = (x_1, \ldots x_{n-1})$.

Lemma 4.3.1. *The following holds.*

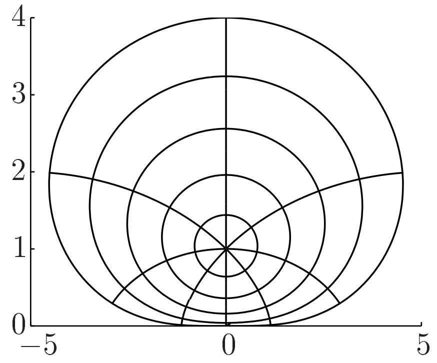

Figure 4.1: Image of radial grid under the exponential map: $r = 2$.

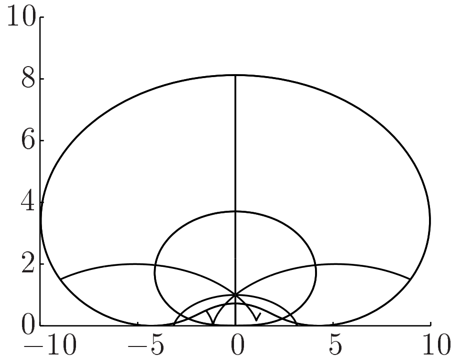

Figure 4.2: $r = 4$

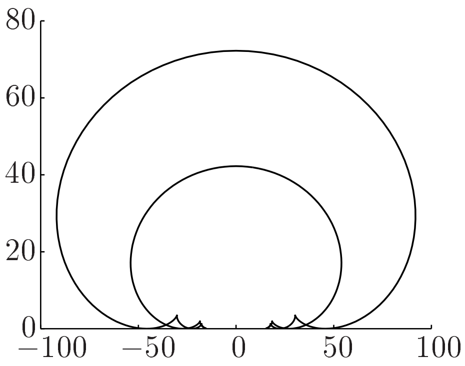

Figure 4.3: $r = 16$

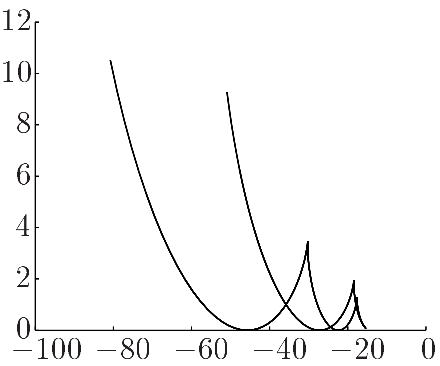

Figure 4.4: Detail of figure 4.3

1. *Estimates from above:*

$$d(x,y) \le 2\sqrt{x_n} + 2\sqrt{y_n} + 2\sqrt{\pi|\hat{x}-\hat{y}|} \tag{4.11}$$

$$d(x,y) \le 2|\sqrt{x_n} - \sqrt{y_n}| + \frac{|\hat{x}-\hat{y}|}{\max\{\sqrt{x_n}, \sqrt{y_n}\}} \tag{4.12}$$

2. *Let* $z_0 = \max\{x_n, y_n\}$. *Then*

$$d(x,y) \ge \min_{z \ge z_0} \left\{ 2\sqrt{z} - \sqrt{\min\{x_n, y_n\}}, \frac{|x-y|}{\sqrt{z}} \right\}. \tag{4.13}$$

Proof. Estimate (4.11) follows from the triangle inequality

$$d(x,y) \le d(x,(\hat{x},0)) + d((\hat{x},0),(\hat{y},0)) + d(y,(\hat{y},0))$$

together with the estimates (4.9) and (4.10). Similarly estimate (4.12) follows from the triangle inequality by

$$d(x,y) \le d(x,(\hat{y},x_n)) + d((\hat{y},x_n),y),$$

where we assume $x_n \ge y_n$, and the trivial inequality $d(x,z) \le x_n^{-1/2}|x-z|$ if $x_n = z_n$. Estimate (4.13) is a consequence of

$$\gamma_n^{-1/2}(s)|\gamma'(s)| \ge \frac{|\gamma'(s)|}{\sup\sqrt{\gamma_n}}$$

and

$$|\gamma(0)-\gamma(1)| \ge 2\left(\sqrt{\sup\gamma_n} - \sqrt{\min\{\gamma_n(0),\gamma_n(1)\}}\right).$$

□

We combine these estimates into a single expression.

Proposition 4.3.2. *The following estimate holds for all $x, y \in H$:*

$$\frac{1}{12}d(x,y) \le \frac{|x-y|}{\sqrt{x_n}+\sqrt{y_n}+\sqrt{|x-y|}} \le d(x,y).$$

Proof. We begin with the first inequality. If $\sqrt{x_n}+\sqrt{y_n} \ge \sqrt{|x-y|}$ then we use

$$2|\sqrt{x_n}-\sqrt{y_n}| + \frac{|\hat{x}-\hat{y}|}{\max\{\sqrt{x_n},\sqrt{y_n}\}} \le 5\frac{|x-y|}{\sqrt{x_n}+\sqrt{y_n}}.$$

Otherwise

$$2(\sqrt{x_n}+\sqrt{y_n}+\sqrt{\pi|\hat{x}-\hat{y}|}) \le 6\sqrt{|x-y|} \le 12\frac{|x-y|}{\sqrt{x_n}+\sqrt{y_n}+\sqrt{|x-y|}}.$$

For the second inequality we may assume that the infimum in (4.13) is assumed at $\sqrt{z} \ge \sqrt{x_n}+\sqrt{y_n}+\sqrt{|x-y|}$ since otherwise at the infimum

$$\frac{|x-y|}{\sqrt{x_n}+\sqrt{y_n}+\sqrt{|x-y|}} \le \frac{|x-y|}{\sqrt{z}} \le d(x,y).$$

Hence

$$\frac{|x-y|}{\sqrt{x_n}+\sqrt{y_n}+\sqrt{|x-y|}} \le \sqrt{|x-y|} \le \sqrt{z}-\sqrt{x_n}-\sqrt{y_n} \le d(x,y)$$

and the proof is complete. □

We call the set $B(x,R) := \{y \in H | d(x,y) \le R\}$ ball with radius R centered at x. Let $\sigma > -1$ and $\mu_\sigma = x_n^\sigma\, dx$. We denote by $|A|_\sigma$ the measure of A with respect to the measure μ_σ.

Lemma 4.3.3. *Let $R > 0$. The following inclusions hold:*

$$\{y : 2000|x-y| \le R(R+\sqrt{x_n})\} \subset B(x,R)$$
$$\subset \{y : |x-y| \le R(R+2\sqrt{x_n})\}.$$

There exists c depending only on n and σ such that

$$c^{-1}R^n(R+\sqrt{x_n})^{n+2\sigma} \le |B(x,R)|_\sigma \le cR^n(R+\sqrt{x_n})^{n+2\sigma}. \tag{4.14}$$

Proof. The inclusions are immediate consequences of Proposition 4.3.2. Let $Q(x,r) = \{y \in H : \max_i |x_i - y_i| < r\}$. The Euclidean ball $\tilde{B}(x,r)$ contains the Euclidean cube $Q(x, r/\sqrt{n})$ and it is contained in the cube $Q(x,r)$. If $x_n \ge r$

$$|Q(x,r)|_\sigma = r^{n-1}\int_{x_n-r/2}^{x_n+r/2} x_n^\sigma\, dx_n = \frac{r^{n-1}}{1+\sigma}\left((x_n+r/2)^{1+\sigma} - (x_n-r/2)^{1+\sigma}\right)$$

and, if $x_n < r$,

$$|Q(x,r)| = r^{n-1}\int_0^{x_n+r/2} x_n^\sigma\, dx_n = \frac{r^{n-1}}{1+\sigma}(x_n+r/2)^{1+\sigma}.$$

Estimate (4.14) is a consequence of these two identities. □

Corollary 4.3.4. *1. x_n^σ is doubling if and only if $\sigma > -1$.*

2. $x_n^s \in A_p(x_n^\sigma)$ if and only if $-1 < s < p(1+\sigma) - 1$.

Proof. It follows from (4.14) that

$$|B(x,3R)|_\sigma \le c(n,\sigma)^2 3^{2n+2\sigma}|B(x,R)|_\sigma$$

which implies the doubling property. Necessity is clear because the volume of a ball containing 0 is infinite if $\sigma \le -1$. Necessity of $s > -1$ in the second part is a consequence of the requirement of being doubling.

If $x_n^s \in A_p(\mu_\sigma)$ then $x_n^{\sigma - \frac{s-\sigma}{p-1}} \in A_{p'}(\mu_\sigma)$ where $1/p' + 1/p = 1$. In particular it is doubling hence $s < p(1+\sigma) - 1$. Now let $s < p(1+\sigma) - 1$. Then

$$\frac{|B|_\rho}{|B|_\sigma}\left(\frac{|B|_{\sigma - \frac{s-\sigma}{p-1}}}{|B|_\sigma}\right)^{p-1} \le c(n,\rho)c(n,\sigma)^p c\left(n, \sigma - \frac{s-\sigma}{p-1}\right)^{p-1}$$

for all d-ball B by (4.14). □

Definition 4.3.5. *We define the parabolic cylinder*

$$Q(x,t,R) = B(x,R) \times (t - R^2, t].$$

We define $\mu_{\sigma,0} = \mu_\sigma \, dt$ and $|A|_{\sigma,0}$ is the $\mu_{\sigma,0}$ measure of A. The metric is $d_0(x,t;y,s) = \sqrt{d(x,y)^2 + |t-s|}$.

Lemma 4.3.6. *There exists c depending only on n and σ such that the following estimate holds*

$$c^{-1}|Q(x,R)|_\sigma \leq (x_n^{1/2} + R)^{n+2\sigma} R^{n+2} \leq c|Q(x,R)|_\sigma.$$

Moreover $\mu_{\sigma,0}$ is doubling with respect to d_0 if and only if $\sigma > -1$ and $x_n^s \in A_p(x_n^\sigma)$ if and only if $-1 < s < p(1+\sigma) - 1$.

Proof. This follows from Lemma 4.3.3 and Corollary 4.3.4. □

4.4 Gaussian estimates

Let a^{ij} be a matrix with measurable coefficients which satisfy

$$\lambda^{-1}|\xi|^2 \leq a^{ij}\xi_i\xi_j \quad \text{and } |a^{ij}\eta_i\xi_j| \leq \lambda|\xi|\,|\eta|, \tag{4.15}$$

for all $\xi, \eta \in \mathbb{R}^n$ and let $\sigma > -1$. We study the operator

$$Lu = u_t - \partial_i(x_n a^{ij}\partial_j u) - \sigma a^{nj}\partial_j u. \tag{4.16}$$

Let $w \in L^1(\mathbb{R}, L^2(\mu_\sigma))$ and $f \in L^2(H \times \mathbb{R}, \mu^0_{\sigma+1})$. The equation

$$Lu = w + x_n^{-\sigma}\partial_i(x_n^{1+\sigma} f^i) \tag{4.17}$$

has a unique weak energy solution $u \in C_b(\mathbb{R}, L^2(\mu_\sigma))$ with $|\nabla u| \in L^2(\mu^0_{\sigma+1})$ in the following sense:

Definition 4.4.1. *The function u with $u, \nabla u \in L^2_{loc}(\mu^0_{\sigma+1})$ is called an* energy *solution to equation* (4.17) *if the identity*

$$\int_{H\times\mathbb{R}} -x_n^\sigma u\phi_t + x_n^{\sigma+1}a^{ij}\partial_i u \partial_j\phi \, dx\, dt = \int_{H\times\mathbb{R}} x_n^\sigma w\phi - x_n^{1+\sigma} f^i \partial_i\phi \, dx\, dt$$

holds for all $\phi \in C_0^\infty(\mathbb{R}^n \times \mathbb{R})$.

We would like to use $\chi_{(t_1,t_2)}u$ as test function. This would lead to

$$\begin{aligned}\frac{1}{2}\int_H x_n^\sigma |u(.,t_2)|^2\,dx &+ \int_{t_1}^{t_2}\int_H x_n^{1+\sigma} a^{ij}\partial_i u \partial_j u\,dx\,dt\\ &= \frac{1}{2}\int_H x_n^\sigma |u(.,t_1)|^2\,dx + \int_{t_1}^{t_2}\int_H x_n^\sigma wu - x_n^{1+\sigma} f^i \partial_i u\,dx\,dt.\end{aligned}\tag{4.18}$$

The solution is constructed by a Galerkin approximation: We choose a suitable increasing sequence of finite dimensional spaces V_j of functions u on H which satisfy $u \in L^2(x_n^\sigma)$ and $\nabla u \in L^2(x_n^{\sigma+1})$ and look for functions $u_j \in C^{0,1}(\mathbb{R}, V_j)$ for which the weak formulation holds for test functions ϕ which are differentiable with values in V_j. The problem reduces to an ordinary differential equation, for which there is a unique solution which satisfies $u_j(t) \to 0$ as $t \to -\infty$. More precisely, by Cauchy-Schwarz

$$\|u_j(t)\|_{L^2(\mu_\sigma)} \leq c\int_{-\infty}^t \|w(s)\|_{L^2(\mu_\sigma)}\,ds + c\|f\|_{L^2(H\times(-\infty,t),\mu^0_{\sigma+1})}.$$

The limit $j \to \infty$ poses no difficulties. There is only one point to check: continuity in time. For that we check that $\phi(t)h^{-1}\int_0^h u(x,t+\tau)d\tau$ with $\phi \in C_0^\infty$ is a legitimate test function. The limit $h \to 0$ does not pose difficulties. Then we choose ϕ which approximates the characteristic function of an interval. This gives (4.18) for every weak solution. As a consequence the norm of $u(.,t)$ is continuous. By construction $t \to u(.,t)$ is weakly continuous and hence $t \to u(.,t) \in L^2(x_n^\sigma)$ is continuous. The same procedure can be used to solve the initial value problem.

These arguments can be used also with cutoff functions in space. Let $B(R) = B(x^0, R)$ and $Q(R) = Q(x^0, t^0, R)$. Then, if $Lu = 0$ in $Q(R)$ in the energy sense,

$$\sup_{t_0-r^2\leq t\leq t_0} \|u(t)\|_{L^2(B(r),\mu_\sigma)} + \|\nabla u\|_{L^2(Q(r),\mu^0_{\sigma+1})} \leq c(R-r)^{-1}\|u\|_{L^2(Q(R),\mu^0_\sigma)}\tag{4.19}$$

for $r < R$. This follows formally by using ϕu as testfunction for a suitable function ϕ. The justification is the same as above. One has to use the fact that $x_n|\nabla d|^2 = 1$ away from $x_n = 0$.

Finally we shall use weighted forms of the energy inequality, with test functions $|u|^{p-2}u$. This is justified in exactly the same way as the energy inequality itself.

We define the kernel distribution $g(x,t;y,s)$ as the kernel of the operator $f \to u$ where u is the unique function which satisfies $Lu = f$ in the sense defined above and which tends to zero in $L^2(H,\mu_\sigma)$ as $t \to -\infty$. This last condition is not really needed, but we do not prove that the null space of L is trivial (i.e. $Lu = 0$,and u an energy solution implies $u = 0$) and hence we have to use an assumption which implies uniqueness with the energy estimates discussed here. The distribution g is actually a function on $(H \times \mathbb{R})^2$ as we shall see in the sequel.

Theorem 4.4.2. *Suppose that $\sigma > -1$. The fundamental solution g is in $L^1_{loc}(H \times \mathbb{R}, H \times \mathbb{R})$. There exist $c_1 = c_1(n, \sigma, \lambda)$ and $c_2 = c_2(\lambda)$ such that*

$$g(x,t,y,s) \le c_1 \frac{{y_n}^{\sigma}}{|B(y,(t-s)^{1/2})|_\sigma} e^{-\frac{d(x,y)^2}{c_2(t-s)}} \tag{4.20}$$

and

$$g(x,t,y,s) \ge c_1^{-1} \frac{y_n^{\sigma}}{|B(y,(t-s)^{1/2})|_\sigma} e^{-\frac{c_1 d(x,y)^2}{t-s}} \tag{4.21}$$

for $t > s$.

Proof. The proof consists of three steps: the estimate from above, a weak estimate from below and the final estimate from below. It suffices to prove the estimates for $0 = s < t$ which we assume in the sequel whenever it is convenient. We denote constants depending only on λ by c_2 and constants depending on n, σ and λ by c_1. We follow the approach of Fabes and Stroock [53] resp. Kusuoka and Stroock [88].

Step 1: It suffices to consider $n + \sigma > 1$, since then we obtain the estimates for $n = 1$ and $s \le 0$ by introducing a dummy space variable and by integrating that variable in the end. We claim that

$$g(x,t,y,s) \le c_1 \frac{y_n^{\sigma}}{t^{n+\sigma}} e^{-\frac{d(x,y)^2}{c_2 t}}. \tag{4.22}$$

Suppose this claim is true. Then estimate (4.20) holds if $\max\{d(x,y), \sqrt{t}\} \ge \sqrt{y_n}$. To see this we consider two cases. Lemma 4.3.3 shows that $|B(x,R)|_\sigma \le c(n,\sigma) R^n (R + \sqrt{y_n})^{n+2\sigma}$. If $t \ge y_n$ then $|B(x,\sqrt{t})|_\sigma \le c(n,\sigma) t^{n+\sigma}$ hence

$$t^{-n-\sigma} e^{-\frac{d(x,y)^2}{c_2 t}} \le c(n,\sigma) |B(y,\sqrt{t})|_\sigma^{-1} e^{-\frac{d(x,y)^2}{c_2 t}}.$$

If $d(x,y) \ge \sqrt{y_n} \ge \sqrt{t}$ then $B(y,\sqrt{t})/t^{n+\sigma} \le c(n,\sigma)(y_n/t)^{\frac{n+\sigma}{2}}$ and

$$t^{-n-\sigma} e^{-\frac{d(x,y)^2}{c_2 t}} \le c(n,\sigma) |B(y,\sqrt{t})|_\sigma^{-1} e^{-\frac{d(x,y)^2}{2c_2 t}} \left[\left(\frac{d(x,y)^2}{t} \right)^{\frac{n+\sigma}{2}} e^{-\frac{d(x,y)^2}{2c_2 t}} \right]$$

where the last factor is uniformly bounded.

We postpone a proof of the estimate (4.20) in the uniformly parabolic case $\max\{d(x,y),\sqrt{t}\} \le \sqrt{y_n}$ and turn to estimate (4.22).

In the sequel we assume that all norms which occur are finite and that all derivatives and integrals make sense. Even if this is not true the calculations can easily be justified.

Let $\alpha \in \mathbb{R}^n$. We define the auxiliary function

$$\Psi = \frac{\alpha\cdot(x-y)}{\sqrt{|x-y|+x_n+y_n}}.$$

A straight forward computation shows that

$$x_n\left|\nabla\Psi_x\right|^2 \le 4|\alpha|^2. \tag{4.23}$$

We compute with $v = e^{\Psi(x,y)}u$ and fixed t,

$$\begin{aligned}
\frac{d}{dt}\int_H |v|^p\,d\mu_\sigma \; &= p\int_H e^\Psi|v|^{p-2}vu_t\,d\mu_\sigma = p\int_H e^\Psi|v|^{p-2}v\partial_i(x_n^{1+\sigma}a^{ij}\partial_j u)dx\\
&= -p\int_H e^\Psi[v\partial_i\Psi+(p-1)\partial_i v]a^{ij}|v|^{p-2}\partial_j u\,d\mu_{\sigma+1}\\
&= \int_H a^{ij}\Big[p|v|^p(\partial_i\psi)(\partial_j\psi) - \frac{4(p-1)}{p}(\partial_i|v|^{p/2})(\partial_j|v|^{p/2})\Big]\,d\mu_{\sigma+1}\\
&\quad -\int_H a^{ij}\Big[2|v|^{p/2}(\partial_i\psi)\partial_j|v|^{p/2}+2(p-1)|v|^{p/2}(\partial_j\psi)\partial_i|v|^{p/2}\Big]\,d\mu_{\sigma+1}\\
&\le -\frac{4(p-1)}{\lambda p}\|\nabla|v|^{p/2}\|^2_{L^2(\mu_{\sigma+1})} + 4|\alpha|\lambda p\,\|v\|^{p/2}_{L^p(\mu_\sigma)}\|\nabla|v|^{p/2}\|_{L^2(\mu_{\sigma+1})}\\
&\quad +4p\lambda|\alpha|^2\|v\|^p_{L^p(\mu_\sigma)}\\
&\le -\frac{2(p-1)}{p\lambda}\|\nabla|v|^{p/2}\|^2_{L^p(\mu_{\sigma+1})} + 2|\alpha|^2\Big(\frac{(p\lambda)^3}{p-1}+2p\lambda\Big)\|v\|^p_{L^p(\mu_\sigma)}.
\end{aligned}$$

The first inequality is a consequence of Cauchy-Schwarz and (4.23). Hence, if $p \ge 2$, with $\beta = -8|\alpha|^2p^2\lambda^3$,

$$e^{-\beta t}\|v(t)\|^p_{L^p(\mu_\sigma)} + \frac{2(p-1)}{\lambda p}\int_0^t e^{-\beta s}\|\nabla v^{p/2}\|^2_{L^2(\mu_{\sigma+1})}\,ds \le \|v(0)\|^p_{L^p(\mu_\sigma)}. \tag{4.24}$$

There is a simple consequence of this formula: $e^{-\beta t/p}\|v(t)\|_{L^p(\mu_\sigma)}$ is nonincreasing. Moreover there exists t_0 with $0 \le t_0 \le t$ and

$$\|\nabla|v|^{p/2}(t_0)\|^2_{L^2(\mu_{\sigma+1})} \le e^{\beta t}\frac{\lambda p}{2(p-1)t}\|v(0)\|^p_{L^p(\mu_\sigma)}.$$

The Hardy-Sobolev inequality, Proposition 4.2.2 can be written in the form

$$\|v(t)\|^p_{L^{\frac{p(n+\sigma)}{n+\sigma-1}}(\mu_\sigma)} \le c(n,\sigma)\|\nabla|v(t)|^{p/2}\|^2_{L^2(\mu_{\sigma+1})},$$

provided that $n+\sigma>1$. Hence

$$\|v(t_0)\|_{L^{p\frac{n+\sigma}{n+\sigma-1}}(\mu_\sigma)} \le c(n,\sigma)^{1/p}\|\nabla|v(t_0)|^{p/2}\|^{2/p}_{L^2(\mu_{\sigma+1})}$$

and, using again that $e^{-\beta t/p}\|v(t)\|_{L^p(\mu_\sigma)}$ is non-increasing,

$$\|v(t-s)\|_{L^{p\frac{n+\sigma}{n+\sigma-1}}(\mu_\sigma)} \le e^{\beta(t-s)/p}(c(n,\sigma)\lambda/(t-s))^{1/p}\|v(s)\|_{L^p(\mu_\sigma)} \tag{4.25}$$

if $t>s$. We apply estimate (4.25) iteratively. Let $p_i = 2\left(\frac{n+\sigma}{n+\sigma-1}\right)^i$, $t_i = t(1-4p_i^{-2})$ and $b=\lambda c(n,\sigma)e^{64|\alpha|^2\lambda^3 t}$. Inequality (4.25) reads as

$$\|e^{\Psi/2}u(t_{i+1})\|_{L^{p_{i+1}}(\mu_\sigma)} \le b^{\frac{1}{p_i}}|t_{i+1}-t_i|^{-\frac{1}{p_i}}\|e^\Psi u(t_i)\|_{L^{p_i}(\mu_\sigma)}$$

hence, if $i\ge j$,

$$\|e^\Psi u(t_i)\|_{L^{p_i}(\mu_\sigma)} \le \prod_{k=j}^{i-1} b^{\frac{1}{p_k}}|t_{k+1}-t_k|^{-\frac{1}{p_k}}\|e^\Psi u(t_j)\|_{L^{p_j}(\mu_\sigma)}.$$

We estimate the products: $\prod_{k=0}^\infty (b/t)^{1/p_k} = (b/t)^{\frac{n+\sigma}{2}}$. Since $t_{k+1}-t_k = 2t\gamma^{-2(k+1)}(\gamma^2-1)$ where $\gamma = \frac{n+\sigma}{n+\sigma-1}$ we have with $c=c(n,\sigma)$

$$\prod_{k\ge 0}((t_{k+1}-t_k)/t)^{-\frac{1}{p_k}} \le c_0 \exp\left(\ln(\gamma)\sum_{k=0}^\infty 2(k+1)\gamma^{-k}\right) = c.$$

After these preparations we let i tend to ∞:

$$\|e^\Psi u(t)\|_{L^\infty(H)} \le c(n,\sigma)\left(\frac{e^{64|\alpha|^2\lambda^3(t-s)}}{t-s}\right)^{\frac{n+\sigma}{2}}\|e^\Psi u(s)\|_{L^2(\mu_\sigma)}. \tag{4.26}$$

Let again $v=e^\Psi u$. The adjoint of the operator

$$L^2(\mu_\sigma)\ni v_0 \to v(t)\in L^2(\mu_\sigma)$$

is the operator

$$L^2(\mu_\sigma)\ni w_0 \to w(t)\in L^2(\mu_\sigma)$$

where $w = e^{-\Psi}\tilde{u}$ and $\tilde{u}$ is a solution to $\tilde{L}\tilde{u} = 0$ with $\tilde{L}$ the operator with a^{ij} replaced by a^{ji}. Hence, by duality

$$\|e^{\Psi}u(.,t)\|_{L^2(\mu_\sigma)} \le c(n,\sigma,\lambda)e^{32(n+\sigma)|\alpha|^2\lambda^3(t-s)}(t-s)^{-(n+\sigma)/2}\|e^{\Psi}u(.,s)\|_{L^1(\mu_\sigma)}$$

and

$$e^{\Psi(x,y)}|u(x,t)| \le c(n,\sigma,\lambda)t^{-n-\sigma}e^{32(n+\sigma)\lambda^3|\alpha|^2 t}\int e^{\Psi(x,y,0)}|u(y,0)|d\mu_\sigma(y).$$

In particular we can solve the initial value problem with data in $L^1(\mu_\sigma)$, and choosing a suitable Dirac sequence we obtain the kernel estimate, with $\rho = 64(n+\sigma)\lambda^3$:

$$|g(x,t,y,s)| \le c(n,\lambda,\sigma)e^{-\Psi(x,y)+\rho|\alpha|^2 t}t^{-n-\sigma}y_n^\sigma.$$

This inequality holds for all α. We optimize it with respect to α. For that purpose we fix x and y. The function

$$\Phi(x,y,t) = \frac{\alpha\cdot(x-y)}{\sqrt{x_n+y_n+|x-y|}} - \rho|\alpha|^2 t$$

is maximal if

$$\alpha = \frac{x-y}{2\rho\sqrt{|x-y|+x_n+y_n}t}$$

which gives $\Phi(x,y,t) = \frac{|x-y|^2}{4\rho t(x_n+y_n+|x-y|)} \ge \frac{d(x,y)^2}{c_2 t}$ and hence

$$|g(x,t,y,s)| \le c(n,\sigma,\lambda)t^{-n-\sigma}y_n^\sigma e^{-\frac{d(x,y)}{c_2(\lambda)t}}.$$

We have seen that this implies

$$|B(y,\sqrt{t})|_\sigma|g(x,t,y,0)| \le c(n,\sigma,\lambda)y_n^\sigma e^{-\frac{d(x,y)}{c_2(\lambda)t}}$$

if $\max\{d(x,y),\sqrt{t}\} \ge \sqrt{y_n}$ and it remains to prove the estimate in

$$Q = \{(x,y) : d(x,y) \le \sqrt{y_n}, |t| \le y_n\}$$

for fixed y uniformly in y. In Q we have $y_n/4 \le x_n \le \frac{9}{4}y_n$ and we are in a uniformly parabolic setting. Let

$$\Gamma = \{(x,t) : d(x,y) = \sqrt{y_n}, 0 \le t \le y_n\}. \tag{4.27}$$

On Γ our kernel estimates give

$$g(x,t,y,s) \leq c \sup_{0\leq s\leq t} s^{-n/2} e^{-\frac{y_n}{c_2(\lambda)s}} =: v(t). \tag{4.28}$$

There is a comparison principle for local solutions to $Lu = 0$. Let Ω be a bounded domain in $\{x : 0 \leq x_n < \delta^{-1}\}$ and $H_0^1(\Omega)$ the closure of C_0^∞ in the norm

$$\|v\|^2 = \|v\|^2_{L^2(\Omega,\mu_\sigma)} + \|\nabla v\|^2_{L^2(\Omega,\mu_{\sigma+1})}.$$

We say that $w \in H^1(\Omega)$ is nonnegative on the boundary $\partial\Omega$ if the negative part $(-w)_+$ is in $H_0^1(\Omega)$.

Lemma 4.4.3. *Suppose that $u \in C([0,T], L^2(\Omega))$, $\nabla u \in L^2(\Omega \times [0,T])$, $u(.,t)$ nonnegative at the boundary of Ω for almost all t, $u(x,0) \geq 0$ almost everywhere in Ω and $Lu = 0$. Then u is nonnegative in $\Omega \times (0,T)$.*

Proof. We only sketch the proof. One verifies that the energy estimate holds for the negative part $u_- = -(-u)_+$ of u. This implies that $u_- = 0$ and hence $u \geq 0$ almost everywhere. $\square$

It suffices to consider

$$u_t - x_n^{-\sigma}\partial_i(x_n^{1+\sigma} a^{ij}\partial_j u) = 0$$

with $\frac{1}{2}y_n \leq x_n \leq 2y_n$ and homogeneous Dirichlet boundary data $u = 0$ at the boundary. We proceed exactly as above until we reach the point where we used the Sobolev inequality. We define $v = ue^\Psi$. Here we use the standard Sobolev inequality

$$\|v\|^p_{L^{p\frac{2n}{n-2}}(\mu_\sigma)} \leq c(n,\sigma,\lambda) y_n^{-1/2} \|\nabla |v|^{p/2}\|_{L^2(\mu_{\sigma+1})}.$$

There is no change in the other arguments and we arrive at kernel estimates for the homogeneous Dirichlet problem. Let us call this kernel $\bar{g}$. Then

$$\bar{g}(x,t,y,0) \leq c(n,\sigma,\lambda) t^{-n/2} e^{-\frac{|x-y|^2}{c_2 y_n t}}.$$

Now $g - \bar{g}$ is a solution to the homogeneous problem which is zero at $t = 0$ and less than $v(t)$ at the boundary of the ball. We apply Lemma 4.4.3 on $\Omega = B(y, \sqrt{y_n})$, $T = t$ to $v(t) - g(.,.,y,0) + \bar{g}(.,.,y,0)$ and hence

$$g(x,t,y,0) \leq v(t) + \bar{g}(.,.,y,0) \leq c(n,\sigma,\lambda) t^{-n/2} e^{-\frac{|x-y|^2}{c_2 y_n t}}$$

with different constants. This is exactly the upper bound (4.20).

Step 2: Here we prove estimate (4.21) for $d(x,y)^2 \le t$. There are several ways to deal with the estimate in the opposite case $d(x,y) \ge \sqrt{t}$. We shall derive the estimate from below later as a consequence of the Harnack inequality. The proof of the Harnack inequality requires only this weaker estimate from below.

The argument of Lemma 4.4.3, together with an approximation of the Dirac measure by a Dirac sequence shows that g is nonnegative. In the sequel we assume that g is positive for $t > 0$. Under this assumption we shall derive a lower estimate. To remove that assumption we apply the arguments to $g + \varepsilon$ and let ε tend to zero. We begin with a Poincaré-type inequality.

Lemma 4.4.4. *Let $\sigma > -1$, $\kappa > 0$ and*

$$\rho(x,y) = \frac{|x-y|^2}{\sqrt{|x-y|^2 + x_n^2 + y_n^2}}.$$

There exists $c(n,\sigma) > 0$ such that

$$\inf_{\bar{u}\in\mathbb{R}} \|u - \bar{u}\|_{L^2(e^{-\kappa\rho(.,x_0)}\mu_\sigma)} \le \frac{c(n,\sigma)}{\sqrt{\kappa}} \|\nabla u\|_{L^2(e^{-\kappa\rho(.,x_0)}\mu_{\sigma+1})}$$

whenever the right hand side is bounded.

Remark 4.4.5. This can be understood as an eigenvalue estimate for the Hermite operator in our setting. Unfortunately there is no explicit description of eigenvalues and eigenfunctions available.

The proof of Lemma 4.4.4 is postponed until the end of the proof of the estimates from below. We fix $x^0 \in H$ and $t^0 > 0$ and consider $y \in B(x^0, \sqrt{t^0})$. Let

$$h(t_0) =: \int_H e^{-\frac{\rho(x,x_0)}{t_0}} d\mu_\sigma(x).$$

Then

$$e^{-1} h(R^2) \le |B(x_0,R)|_\sigma \le c(n,\sigma) h(R^2). \tag{4.29}$$

For the second inequality one restricts the integration to $B(x_0,R)$. The first estimate is obtained by integrating on $B(x_0,R)$ and $B(x_0,2^iR)\backslash B(x_0,2^{i-1}R)$: Then the obvious estimate of the integrand and the volume estimate of Lemma 4.3.3 allow to sum up the parts. Let

$$G(t) := h(t_0)^{-1} \int e^{-\frac{\rho(x,x_0)}{t_0}} \ln\left(\frac{h(t_0)g(x,t,y,0)}{y_n^\sigma}\right) d\mu_\sigma(x).$$

We want to show that there exists $D(n, \sigma, \lambda)$ with $G(t_0) \geq -D(n, \sigma, \lambda)$. The time derivative of G can be estimated as follows:

$$
\begin{aligned}
\frac{d}{dt}G(t) &= \frac{1}{h(t_0)} \int_H e^{-\frac{\rho(x,x_0)}{t_0}} \frac{g_t}{g}\, d\mu_\sigma(x) \\
&= \frac{1}{h(t_0)} \int_H e^{-\frac{\rho}{t_0}} a^{ij} (\partial_i \ln g)(\partial_j \ln g)\, d\mu_{\sigma+1}(x) \\
&\quad + \frac{1}{t_0 h(t_0)} \int_H e^{-\frac{\rho}{t_0}} [\partial_i \rho] a^{ij} \partial_j \ln g\, d\mu_{\sigma+1}(x) \\
&\geq \frac{1}{2\lambda h(t_0)} \int_H e^{-\frac{\rho}{t_0}} \left| \nabla \ln \left(\frac{h(t_0) g}{y_n^\sigma} \right) \right|^2 d\mu_{\sigma+1}(x) - \frac{c(n,\sigma)\lambda^3}{2t_0} \\
&\geq \frac{1}{c(n,\sigma)\lambda t_0 h(t_0)} \int_H e^{-\frac{\rho}{t_0}} \left| \ln \left(\frac{h(t_0) g}{y_n^\sigma} \right) - G(t) \right|^2 d\mu_\sigma - \frac{c(n,\sigma)\lambda^3}{2t_0}.
\end{aligned}
\tag{4.30}
$$

We did use Cauchy-Schwarz and the obvious estimate $x_n |\nabla \rho|^2 \leq c\rho$ (since the norm of the gradient of the distance in a Euclidean manifold is always 1) for the first inequality, plus

$$\int_H \frac{\rho}{t_0} e^{-\rho/t_0} d\mu_\sigma \leq 2 \sup h(t_0/2).$$

The second inequality is a consequence of Poincaré's inequality, Lemma 4.4.4. Now

$$\frac{d}{du} \left[\frac{(\ln u - G(t))^2}{u} \right] \leq 0 \qquad \text{if } u \geq e^{2+G(t)}$$

and

$$\frac{h(t_0) g(x,t,y,0)}{y_n^\sigma} \leq K(n,\sigma,\lambda) \qquad \text{for } t > t_0/2$$

because of (4.20) and (4.29). Let $U_t := \left\{ x : \frac{h(t_0)g(x,t,y,0)}{y_n^\sigma} \geq e^{2+G(t)} \right\}$. Then

$$t_0 G'(t) \geq -\frac{\lambda^3}{2} + \frac{1}{c(n,\sigma)\lambda} \frac{(\ln K - G(t))^2}{K} \int_{U_t} e^{-\frac{\rho(x,x_0)}{t_0}} y_n^{-\sigma} g(x,t,y,0)\, d\mu_\sigma.$$

But

$$
\begin{aligned}
\int_{U_t} e^{-\frac{\rho(x_0,y)}{t_0}} y_n^{-\sigma} g(x,t,y,0)\, d\mu_\sigma(x) \ &\geq \int_H e^{-\frac{\rho}{t_0}} y_n^{-\sigma} g\, d\mu_\sigma(x) - e^{2+G(t)} \\
&\geq e^{-R} \int_{\{x:\frac{\rho(x,x_0)}{t_0}\leq R\}} y_n^{-\sigma} g\, d\mu_\sigma(x) - e^{2+G(t)} \\
&\geq e^{-R}\Big(1 - \int_{\{x:\frac{\rho}{t_0}\geq R\}} y_n^{-\sigma} g\, d\mu_\sigma(x)\Big) - e^{2+G(t)} \\
&\geq \frac{e^{-R}}{2} - e^{2+G(t)}
\end{aligned}
\tag{4.31}
$$

if R is sufficiently large and $t > t_0/2$ because of (4.20). For the third inequality we have used that $\int y_n^{-\sigma} g(x,t,y,s)\, d\mu_\sigma(y) = 1$ for all t – which is equivalent to the fact that constants are solutions. If $G(t) \leq -4R$ and $R \geq 2$ then

$$
t_0 G' \geq -c(n,\sigma,\lambda) + \frac{1}{c(n,\sigma,\lambda)} \frac{(\ln K - G(t))^2}{K}
$$

and hence, if in addition $G(t) \leq \min\{2\ln K, 0\}$

$$
t_0 G'(t) \geq -c(n,\sigma,\lambda) + \frac{G(t)^2}{c(n,\sigma,\lambda)}
$$

where we included the dependence on K, which is itself a function of n,σ and λ, into the constant. If $G(t) \geq \min\{-4R, 2\ln K\}$ for one $t > t_0/2$ then the desired estimate clearly holds. Let us hence suppose that

$$
G(t) \leq \min\{-4R, 2\ln K, -2c(n,\sigma,\lambda)\} \qquad \text{for all } t_0/2 \leq t \leq t_0.
$$

Then $t_0 G'(t) \geq \frac{G(t)^2}{2c(n,\sigma,\lambda)}$ and hence

$$
G(t_0) \geq -D(n,\sigma,\lambda).
$$

Step 3: The estimate $G(t_0) \geq -D(n,\sigma,\lambda)$ implies a pointwise estimate. The $L^2(\mu_\sigma)$ adjoint of the map $u(0) \to u(t)$ is a solution operator of the equation where a^{ij} is replaced by a^{ji}. Hence, if we denote the kernel of the adjoint equation by $\tilde{g}$,

$$
\tilde{g}(x,t,y,0) := \Big(\frac{y_n}{x_n}\Big)^{\sigma} g(y,t,x,0),
$$

it satisfies the same upper and lower estimates. We claim that

$$
g(x,t,y,0) \geq c(n,\sigma,\lambda) \frac{y_n^\sigma}{|B(x,t^{1/2})|_\sigma}
\tag{4.32}
$$

if $d(x,y)^2 \le t$. The function g is the kernel of the evolution operator. Thus

$$\frac{h(t)g(x,2t,y,0)}{y_n^\sigma} = h(t)^{-1}\int_H \frac{h(t)g(z,t,y,0)}{y_n^\sigma}\frac{h(t)g(x,2t,z,t))}{z_n^\sigma} d\mu_\sigma(z)$$

and, by Jensen's inequality (with $t_0 = t$, $x_0 = y$, $d(x,y) \le t^{1/2}$)

$$\begin{aligned}
\ln\left(\frac{h(t)g(x,2t,y,0)}{y_n^\sigma}\right) &\ge \ln\left(h(t)^{-1}\int_H e^{-\frac{\rho(z,y)}{t_0}}\dots d\mu(z)\right)\\
&\ge h(t)^{-1}\int_H e^{-\frac{\rho}{t_0}}\ln\left(\frac{h(t)g(x,2t,z,t)}{z_n^\sigma}\right) d\mu_\sigma(z)\\
&\quad +h(t)^{-1}\int_H e^{-\frac{\rho}{t_0}}\ln\left(\frac{h(t)g(z,t,y,0)}{y_n^\sigma}\right) d\mu_\sigma(z)\\
&\ge -c(n,\sigma,\lambda)
\end{aligned} \tag{4.33}$$

which immediately implies claim (4.32) and hence the estimate from below in the region where $d(x,y) \le \sqrt{t}$. The kernel estimates for $d(x,y)^2 > t$ could be derived directly from these weaker estimates. We prefer however to postpone this point until we have proven the Harnack inequality, which will technically simplify the proof considerably while still using the same idea. □

Proof of Lemma 4.4.4: We may assume that $\kappa = 1$ and $\hat{y} = 0$. This can be achieved by an affine transformation. The function ρ has then the form

$$\rho(x,y) = \frac{|\hat{x}|^2 + (x_n - y_n)^2}{\sqrt{|\hat{x}|^2 + (x_n - y_n)^2 + x_n^2 + y_n^2}}.$$

Let $A = \{x : x_n - y_n/2 > |\hat{x}|\}$. This is the larger cone in Figure 4.5. Let

$$C = \Big\{x \in H : |\hat{x}| \ge \min\{x_n/2 - y_n/3, 8(y_n - x_n)\}\Big\}$$

and let $D = C^0 \cup A$. The set C is the exterior of the union of the smaller cone in Figure 4.5 and a large cylinder. Let $B = B(y,2)$ and suppose that $y_n > 2$. We claim three inequalities:

$$\int_H e^{-\rho(x,y)}|u|^2 d\mu_\sigma(x) \le c\int_D e^{-\rho}|u|^2\, d\mu_\sigma + c\int_H e^{-\rho}|\nabla u|^2\, d\mu_{\sigma+1} \tag{4.34}$$

$$\int_D e^{-\rho(x,y)}|u|^2 d\mu_\sigma(x) \le c\int_B e^{-\rho}|u|^2\, d\mu_\sigma + c\int_D e^{-\rho}|\nabla u|^2\, d\mu_{\sigma+1} \tag{4.35}$$

$$\inf_{\bar{u}}\int_B |u-\bar{u}|^2\, dx \le c(\sigma)\int_B |\nabla u|^2 dx. \tag{4.36}$$

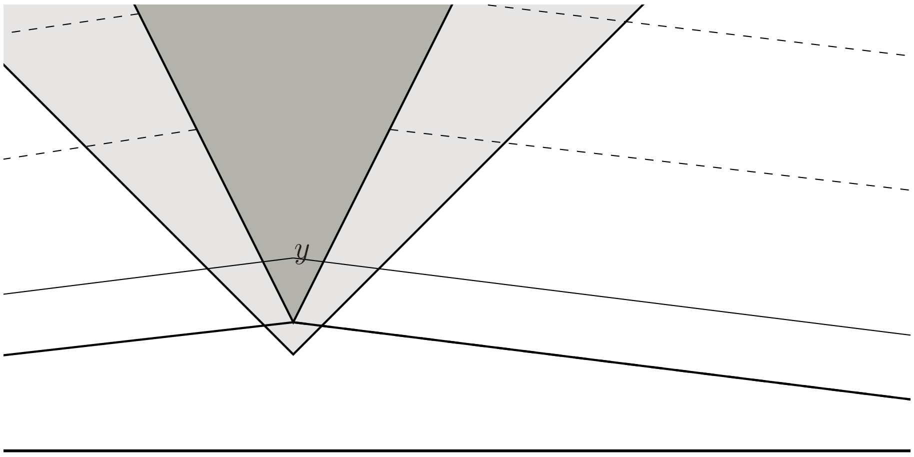

Figure 4.5: The regions

We postpone their proof. Then

$$\begin{aligned}\inf_{\bar u}\int_H e^{-\rho}(u-\bar u)^2 d\mu_\sigma &\le c y_n^\sigma \inf_{\bar u}\int_B (u-\bar u)^2 d\mu_\sigma + c\int_H e^{-\rho(x,y)}|\nabla u|^2\, d\mu_{\sigma+1}\\ &\le \tilde c\int_H e^{-\rho(x,y)}|\nabla u|^2\, d\mu_{\sigma+1}\end{aligned}$$

where we used all three inequalities. To remove the assumption $y_n > 2$ we remark that ρ is Lipschitz continuous in both variables. Hence

$$\begin{aligned}\inf_{\bar u}\int_H e^{-\rho}(u-\bar u)^2 d\mu_\sigma &\le c_1 \inf_{\bar u}\int_H e^{-\rho(.,.+2e_n)}(u-\bar u)^2 d\mu_\sigma\\ &\le c_2\int_H e^{-\rho(.,.+2e_n)}|\nabla u|^2\, d\mu_{\sigma+1}\\ &\le c_3\int_H e^{-\rho}|\nabla u|^2\, d\mu_{\sigma+1}.\end{aligned}$$

By this we have reduced the proof to a verification of the claims. Inequality (4.36) is Poincaré's inequality. We turn to the proof of inequality (4.34).

We shall use cylindrical coordinates $x = (r-8t)\theta + te_n$ and apply the transformation formula to reduce (4.34) to a family of one-dimensional integrals, which we obtain as inner integrals if we do the t as the inner most integration. We define $f_1(r) = \frac{r}{10} + \frac{2y_n}{15}$ and $f_2(r) = \frac{r}{9} + \frac{y_n}{18}$. They give the value of t when the lines intersect the two cones.

Let $g(t) = u((r-8t)\theta + te_n)$. By an abuse of notation we write $\rho(t)$ for

$\rho((r-8t)\theta+te_n,y)$. In B we have $1 \le \frac{r}{r-8t} \le 60/7$. We shall prove that

$$\int_0^{f_1(r)} t^\sigma e^{-\rho(t)}|g|^2 dt \le c\int_0^{f_1(r)} t^{\sigma+1}e^{-\rho(t)}|g'|^2\,dt + c\int_{f_2(r)}^{f_1(r)} t^\sigma e^{-\rho(t)}|g|^2\,dt, \tag{4.37}$$

which gives (4.34). A calculation shows that

$$\sum_{i=1}^{n-1} x_i/|\hat{x}|\partial_i\rho(x,y) - \frac{1}{8}\partial_{x_n}\rho(x,y) \ge \frac{1}{100} \qquad \text{in } C.$$

By the fundamental theorem of calculus (with $R=f_1(r)$ and κ small)

$$g(t) = g(R-\kappa(R-t)) - \int_t^{R-\kappa(R-t)} g'(s)ds$$

hence

$$\begin{aligned}\int_0^R t^\sigma e^{-\rho(t)}|g|^2 dt \ &\le \kappa^{-1}\int_{(1-\kappa)R}^R (R-\kappa(R-t))^\sigma e^{-\rho(R-\kappa(R-t))}|g(t)|^2\,dt \\ &\quad + c(\sigma)\int_0^R t^\sigma e^{-\rho(t)}\Big(\int_t^{R-\kappa(R-t)}|g'(s)|ds\Big)^2\,dt\end{aligned}$$

where $c_0 = \int_0^R t^\sigma e^{-\rho}\,dt$. This yields the desired inequality because

$$\sup_{(1-\kappa)R\le t\le R} \frac{R^\sigma}{t^\sigma} e^{\rho(t)-\rho(R-\kappa(R-t))} < \infty$$

and the second term is controlled by dividing the s integral up into a sum of integrals of length at most one and then by changing the order of integration. The exponential weight allows to drop all cross terms. This completes the proof of claim (4.34).

In D we have

$$\sum_i \frac{x_i-y_i}{|x-y|}\partial_i\rho(x,y) \ge \frac{1}{3}\left(\frac{|x|}{y_n+|x|}\right).$$

We use polar coordinates centered at y and denote $\rho(r) = \rho(y+r\theta,y)$, $g(r) = u(y+r\theta)$. It suffices to show,

$$\begin{aligned}\int_0^{R(\theta)} r^{n-1}(y_n+r\theta_n)^\sigma e^{-\rho(r)}g^2\,dr \ &\le \int_0^{R(\theta)} r^{n-1}(y_n+r\theta_n)^{\sigma+1}e^{-\rho(r)}|g'|^2\,dr \\ &\quad + \int_0^1 r^{n-1}(y_n+r\theta_n)^\sigma e^{-\rho(r)}g^2 dr.\end{aligned}$$

We prove this in the same way as above by setting (if $R(\theta) \geq 1$, otherwise there is nothing to show)

$$g(r) = g(r/(1+r)) + \int_{r/(1+r)}^{r} g'(r).$$

We omit the details. This completes the proof of (4.35). $\square$

4.5 Hölder continuity

In this section we shall use the Gaussian estimates to derive Hölder continuity and a Harnack inequality. Clearly all this has to be done in the Carnot-Carathéodory metric. The first step, which is more or less a repetition of the estimates of the fundamental solution, is a supremum estimate. The second step is the derivation of Harnack's inequality. The third step consists of a combination with certain kernel estimates to obtain Hölder estimates.

Proposition 4.5.1 (Supremum estimates). *Suppose that*

$$u_t - \partial_i(x_n a^{ij} \partial_j u) - \sigma a^{nj} \partial_j u = 0 \text{ in } Q(x^0, t^0, 2R) \tag{4.38}$$

in the sense of Definition 4.4.1. Then

$$\|u\|_{L^\infty(Q(x^0,t^0,R))} \leq c(n, \sigma, \lambda) |Q(x^0, t^0, 2R)|_{\sigma,0}^{-1} \int_{Q(x^0,t^0,2R))} |u| \, d\mu_\sigma^0.$$

Proof. We may and do assume that $n + \sigma > 1$. If not we introduce dummy variables, which do not change the assertion. We use a Moser iteration. The main difficulty is caused by our need to have the L^1 norm on the right hand side. The constants in (4.24) deteriorate if $p \to 1$. In particular we may not set $p = 1$ there. So we approach $p = 1$ step by step, while controlling the constants. We balance the deteriorating constants by restricting to small improvements near $p = 1$. We first need a version of Nash's inequality.

Lemma 4.5.2 (Nash's inequality). *Let $n + \sigma > 1$, let $B(x^0, R)$ be a ball in H and $v \in L^p(B(x^0, R), \mu_\sigma)$. Then*

$$\|v\|_{L^{2+\frac{2}{n+\sigma}}(B(x^0,R),\mu_\sigma)}^{2+\frac{2}{n+\sigma}} \leq c\Big(R^{-2}\|v\|_{L^2(B,\mu_\sigma)}^2 + \|\nabla v\|_{L^2(B(x^0,R),\mu_{\sigma+1})}^2\Big) \|v\|_{L^2(B(x^0,R),\mu_\sigma)}^{2/(n+\sigma)}.$$

Proof. We apply Lemma 4.2.4 to ϕv where ϕ is a suitable cutoff function. $\square$

Let the nonnegative function $\chi \in C_0^\infty(H \times \mathbb{R})$ satisfy $\chi(x,t) = 0$ if $t \le t_1$ and $1 < p < \infty$. Then, with κ depending only on λ, the bound for the ellipticity of the coefficients,

$$\sup_{t_1 \le t \le t_0} \frac{1}{p} \int_H \chi^2 |u(.,t)|^p d\mu_\sigma + \frac{p-1}{\kappa p^2} \int_{H \times (t_1,t_0)} \chi^2 |\nabla u^{p/2}|^2 d\mu_{\sigma+1}\, dt \\ \le \int_{H \times (t_1,t_0)} \Big(\frac{1}{p} \chi |\chi_t| + \frac{\kappa}{p-1} x_n |\nabla \chi|^2 \Big) |u|^p \, d\mu_\sigma \, dt. \tag{4.39}$$

The localized energy-type inequality (4.39) and Lemma 4.5.2 imply a gain of integrability on smaller subdomains. To make use of this we construct a sequence of cut–off functions. Let $\phi_0 \in C^\infty(\mathbb{R})$ be as above: $\phi_0(s) = 1$ for $s \le 0$ and $\phi_0(s) = 0$ for $s \ge 1$. We define for $N \ge 0$

$$\phi_N(x) := \phi_0 \Big(2(N+2)^2 \Big[\frac{d^2(x,x^0)}{R^2} - 1 - \frac{1}{2(N+2)} \Big] \Big)$$

which satisfies $\phi_N(x) = 1$ if $d(x,x^0)^2 \le (1 + 1/(2(N+2)))R^2$ and $\phi_N(x) = 0$ if $d(x,x^0)^2 \ge (1 + 1/(2(N+1)))R^2$. If $N < 0$ we set

$$\phi_N(x) := \phi_0 \Big(2(|N|+1)^2 \Big[\frac{d^2(x,x^0)}{R^2} - 2 + \frac{1}{2|N|} \Big] \Big)$$

with $\phi_N(x) = 1$ if $d(x,x^0)^2 \le (2 - 1/(2|N|))R^2$ and $\phi_N(x) = 0$ if $d(x,x^0)^2 \ge (2 - 1/(2(|N|+1))R^2$. In both cases $x_n |\nabla \phi_N|^2 \le c(\phi_0, R)(|N|+1)^4$. We define $p_N = 2\left(\frac{n+\sigma+1}{n+\sigma}\right)^N$ if $N \ge 0$ and the dual exponent $p_N = \frac{p_{-N}}{p_{-N}-1}$ of p_{-N} if $N < 0$. Let $t_N = t^0 - (1 + 1/(2(N+1)))R^2$ if $N \ge 0$, $t_N = t^0 - (2 - 1/(2|N|))R^2$ if $N < 0$ and let $\tilde\phi_N(t) = 1$ for $t \ge t_N$ and $\tilde\phi_N(t) = 0$ if $t \ge t_{N-1}$ and affine between t_N and t_{N-1}. We use inequality (4.39) with $p = p_N$ and $\chi = \phi_N(x)\tilde\phi_N(t)$. The function χ is not smooth, but nevertheless it is a legitimate cutoff function. Let $B_N = \{x \in H : \phi_N(x) = 1\}$ and

$$I(N) = \Big(\int_{t_N}^{t_0} \int_{B_N} |u(x,s)|^{p_{N+1}} \, d\mu_\sigma(x) \Big)^{1/p_{N+1}}.$$

Then, because of Lemma 4.5.2 and (4.39), if $N \ge 0$

$$I(N) \le (c(n,\lambda,\sigma,R)(1+N)^4)^{1/p_N} I(N-1).$$

The case $N < 0$ is a bit more technical. There by Hölder's inequality, with $\alpha = \frac{n+\sigma+1}{n+\sigma}$,

$$I(N) \le I(N-1)^{1 - \frac{1}{2\alpha^{|N|+1} - \alpha}} \|u\|_{L^{\frac{n+\sigma+1}{n+\sigma} p_N}(B_N \times (t_N, t^0), \mu_\sigma)}^{\frac{1}{2\alpha^{|N|+1} - \alpha}}$$

since

$$\frac{1}{p_{N+1}} = \frac{2\alpha^{|N|+1}-1}{2\alpha^{|N|+1}} = \left[1-\frac{1}{2\alpha^{|N|+1}-\alpha}\right]\frac{2\alpha^{|N|}-1}{2\alpha^{|N|}} + \frac{1}{2\alpha^{|N|+1}-\alpha}\frac{2\alpha^{|N|+1}-\alpha}{2\alpha^{|N|}}$$
$$= \left[1-\tfrac{1}{2\alpha^{|N|+1}-\alpha}\right]p_N^{-1} + \left[\tfrac{1}{2\alpha^{|N|+1}-\alpha}\right](\alpha p_N)^{-1}.$$

Hence, with $\beta = \frac{1}{2\alpha^{|N|+1}-\alpha}$,

$$I(N) \leq \left[c\frac{\kappa^2}{(p_N-1)^2}|N|^4\right]^\beta I(N-1) \leq \left[c\kappa^2\alpha^{2|N|+2}(1+|N|^4)\right]^\beta I(N-1)$$

and induction gives $I_N \leq \rho I_{-N}$ with ρ independent of N. The limit $N \to \infty$ gives the desired conclusion, with constant depending on R. The correct dependence on R follows by scaling. □

Theorem 4.5.3 (Harnack inequality). *Suppose that u is nonnegative in $Q(x_0, t_0, 2R)$ and satisfies* (4.38)*. Then there exists $\gamma > 0$, depending only on n, σ and λ, such that*

$$\inf_{(y,s)\in Q(x_0,t_0,R/2)} u(y,s) \geq \gamma \sup_{(y,s)\in Q(x_0,t_0-R^2,R/2)} u(y,t).$$

Proof of Theorem 4.5.3: Proposition 4.5.1 implies that

$$\|u\|_{L^\infty(Q(x_0,t_0-R^2,R/4))} \leq c|Q(x_0,t_0-R^2,R/2)|_\sigma^{-1}\int_{Q(x_0,t_0-R^2,R/2)}|u|d\mu_\sigma^0.$$

Let $v = tu$. Then

$$v_t - \partial_i(x_n a^{ij}\partial_j v) - \sigma a^{nj}\partial_j v = u. \tag{4.40}$$

We assume first that u is a nonnegative solution in $\tilde{Q} = Q(x_0, t_0, NR)$ for some large N. Let v^1 be the solution to (4.40) which is zero at the parabolic boundary of $\tilde{Q}$ in the sense that $v^1(x,s) = 0$ for almost all $x \in B(x_0, NR)$ and for almost all t we have $v^1 \in H_0^1(B(x_0, Nr) \cap \{x_n \geq \delta\})$ for all $\delta > 0$ (here we have used the closed halfplane $\{x_n \geq \delta\}$ so that v^1 is not required to vanish at $x_n = \delta$). Existence of v^1 can be shown by the same arguments as in H. Then $tu = v \geq v^1$ almost everywhere by the comparison principle Lemma 4.4.3 and it suffices to bound v^1 from below.

Let g^1 be the fundamental solution for the Dirichlet problem in its weak formulation. The result follows if we prove

$$\frac{|B(x_0,\sqrt{t_0-s})|_\sigma g^1(x_0,t_0,y,s)}{y_n^\sigma} \geq c_0 \tag{4.41}$$

for $d^2(x_0, y) \leq (R/4)^2$ and $R^2 \leq |t-s| \leq 2R^2$ from the representation formula. Inequality (4.41) holds for the fundamental solution in the half space. The maximum principle, applied to the difference of the two fundamental solutions shows

$$\frac{|B(x_0,\sqrt{t_0-s})|g^1(x_0,t_0,y,s)}{y_n^\sigma} \geq \frac{|B(x_0,\sqrt{t_0-s})|g(x_0,t_0,y,s)}{y_n^\sigma} - e^{-N^2/c}$$

for $(y,s) \in Q(x^0,t^0,2R)$ where we used the bound (4.20) from Theorem 4.4.2 for $g(x,t,y,s)$ on $\partial B(x^0, NR) \times [s,t^0]$. This yields the desired estimate for N sufficiently large by Theorem 4.4.2.

It remains to remove the restriction on the size of the cylinder. This is done by a straightforward iteration in the same way as in the proof of the Gaussian lower bound of the fundamental solution below. We restrict ourselves to that proof because it is more interesting and it easily specializes to the Harnack inequality. □

Proof of Theorem 4.4.2, completion: Here we prove the lower bound of fundamental solutions for $t - 2d(x,y)^2 \leq s \leq t$. Let $g(x,t,y,s)$ be the fundamental solution and $t > s$. We denote by γ the shortest geodesic connecting y and x, parametrized such that $\gamma_n^{-1}|\gamma'|^2 = d^2(x,y)$, $\gamma(0) = y$ and $\gamma(1) = x$.

Then $d(\gamma(r),\gamma(t)) = d(x,y)|r-t|$. We choose the smallest integer such that $\frac{d^2(x,y)}{N(t-s)} \leq 1$. Let $x_i = \gamma(i/N)$ and $t_i = s + \frac{i(t-s)}{N}$ for $0 \leq i \leq N$. Then $\frac{d^2(x_{i+1},x_i)}{t_{i+1}-t_i} = \frac{d^2(x,y)}{N(t-s)} \leq 1$. We have already proven that

$$\frac{|B(y,\sqrt{(t-s)/N})|_\sigma g(x_1,t_1,y,s)}{y_n^\sigma} \geq c_0$$

and, by the proven part of the Harnack inequality, Theorem 4.5.3

$$g(x_{i+1},t_{i+1},y,s) \geq c_1 g(x_i,t_i,y,s)$$

altogether, with $c = c(n,\sigma,\lambda)$,

$$\frac{|B(y,\sqrt{t-s})|_\sigma g(x,t,y,s)}{y_n^\sigma} \geq \frac{|B(y,\sqrt{t-s})|_\sigma}{|B(y,\sqrt{(t-s)/N})|_\sigma} c_0 c_1^{N-1} \geq c(n,\sigma) c_1^N N^{\frac{n+\sigma}{2}}$$

where we used the estimates of Lemma 4.3.3. We arrive at

$$\frac{|B(y,\sqrt{t-s})|_\sigma g(x,t,y,s)}{y_n^\sigma} \geq c(n,\sigma) e^{-c(n,\sigma,\lambda)\frac{d(x,y)^2}{t-s}} \left(\frac{d(x,y)^2}{t-s}\right)^{\frac{n+\sigma}{2}}$$

for $d(x,y)^2 \geq t-s \geq 0$. The last term on the right hand side can be estimated by the exponential and we arrive at the desired estimate by increasing the constant in the exponential. □

Harnack inequality and Hölder regularity are more or less equivalent. We could prove them in any order. Here we derive Hölder regularity from the Harnack inequality. Before that we define Hölder norms with respect to the Carnot-Carathéodory metric.

Definition 4.5.4.

$$\|u\|_{\dot{C}^\alpha(\Omega_T)} = \sup_R \left\{ \sup_{d^2(x,y)+|t-s|\le R^2} \frac{|u(x,t)-u(y,s)|}{R^\alpha} \right\}$$

and $\|u\|_{C^\alpha(\Omega_T)} = \max\{\|u\|_\infty, \|u\|_{\dot{C}^\alpha}\}$. *Moreover* C^α *is the subspace of the space of continuous functions for which the norm* $|.|_{C^\alpha}$ *is finite.*

Theorem 4.5.5 (Hölder regularity). *Let* γ *as in Theorem 4.5.3 and* $\alpha = -\ln_4(1-\gamma)$. *If* u *satisfies* (4.38) *in the sense of Definition 4.4.1 then* $u \in C^\alpha(Q(x^0,t^0,R))$ *and*

$$\|u\|_{\dot{C}^\alpha(Q(x^0,t^0,R))} \le c(n,\sigma,\lambda) R^{-\alpha} |Q|_\sigma^{-1} \|u\|_{L^1(Q((x,t),2R),\mu_\sigma^0)}.$$

Proof. We use again the notation $Q(R)=Q(x^0,t^0,R)$ and define $\omega(u,Q(r))=\sup_{(x,t),(y,s)\in Q(r)} |u(x,t)-u(y,s)|$. Hölder continuity will follow by induction from

$$\omega(Q(R/2),u) \le (1-\gamma)\omega(Q(2R),u). \tag{4.42}$$

Let $M(Q,u)$ be the supremum and $m(Q,u)$ the infimum on the cylinder Q. Let $M = M(Q(2R)$ and $m = m(Q(2R)$. Then $u-m$ and $M-u$ are positive solutions in $Q(2R)$. The Harnack inequality implies (with γ from Theorem 4.5.3)

$$m(Q(R/2), M-u) \ge \gamma M(Q(x^0, t^0-R^2, R/2), M-u)$$

and

$$m(Q(R/2), u-m) \ge \gamma M(Q(x^0, t^0-R^2, R/2), u-m).$$

This gives, by adding the inequalities

$$\begin{aligned} M(Q(R/2)-m(Q(R/2),u) &\le M-m-\gamma\Big[M(Q(x^0,t^0-R^2,R/2),u-m) \\ &\qquad -M(Q(x^0,t^0-R^2,R/2),M-u)\Big] \\ &\le (1-\gamma)(M-m). \end{aligned}$$

This is inequality (4.42). Let (x,t),$(y,s) \in Q(x_0,t_0,R/2)$ with $s<t$ and set $d_0 = \max\{d(x,y), \sqrt{|t-s|}\}$. Then

$$\begin{aligned}|u(x,t)-u(y,s)| &\le \omega(Q(x,t,d_0))\\ &\le (1-\gamma)^{\ln_4(R/(16d_0))}\omega(Q(x_0,t_0,3R/2))\\ &\le (1-\gamma)^{-2}(R/d_0)^{\ln_4(1-\gamma)}\|u\|_{L^\infty(Q(x_0,t_0,3R/2)}\\ &\le c(R/d_0)^{\ln_4(1-\gamma)}|Q(x_0,t_0,2R)|^{-1}_{\sigma,0}\|u\|_{L^1(Q(x_0,t_0,2R),\mu^0_\sigma)}.\end{aligned}$$

where we used Proposition 4.5.1 for the last step. □

Theorem 4.5.6 (First Schauder type estimate). *Suppose that $\sigma>0$, $0<\beta<\alpha=\ln_4(1/(1-\gamma))$, $R\le 1$ and denote $Q(R)=Q(x^0,t^0,R)$. Suppose that*

$$u_t - x_n^{-\sigma}\partial_i(x_n^{1+\sigma}a^{ij}\partial_j u) = x_n^{-\sigma}\partial_i(x_n^\sigma f^i) \quad \text{in} \quad Q(2R)$$

in the energy sense. Then, with $c=c(n,\lambda,\sigma,\beta)$,

$$\|u\|_{\dot C^\beta(Q(R))} \le c\left(R^{-\beta}|Q(2R)|^{-1}_{\sigma,0}\int_{Q(x^0,t^0,2R)}|u|\,d\mu^0_\sigma + \|f\|_{\dot C^\beta(Q(2R))}\right).$$

Proof. We begin with a preliminary result. The kernel estimates and the energy inequality imply uniform boundedness of the L^1 norm of the derivative of the fundamental solution in dyadic annuli.

Lemma 4.5.7. *Suppose that $\sigma>0$. Then*

$$\sup_{x,t,R}\int_{R^2\le\max\{d^2(x,y),|t-s|\}\le 2R^2}|y_n^\sigma\nabla_y(y_n^{-\sigma}g(x,t,y,s))|\,dy\,ds \le c(n,\sigma,\lambda).$$

We postpone the proof. Let $(x,t)\in Q(R)$ and

$$v(x,t) = |\int_{Q(2R)} y_n^\sigma\Big\{\nabla_{y_i}[y_n^{-\sigma}g(x,t,y,s)]\Big\}(f^i(x,t)-f^i(y,s))\,dy\,ds,$$

$$|v(x,t)| \le cR^\beta\|f\|_{\dot C^\beta(Q(2R))}$$

by the kernel estimates of Lemma 4.5.7 and a summation over dyadic annuli $(Q(x,t,2^{-j}R)\backslash Q(x,t,2^{-j-1}R))\cap Q(2R)$. Hence, by Proposition 4.5.1, omitting x^0 and t^0, applied to $u-v$,

$$\|u\|_{L^\infty(Q(R))} \le c\left(|Q(2R)|^{-1}_\sigma\|u\|_{L^1(Q(2R),\mu^0_\sigma)} + R^\beta\|f\|_{\dot C^\beta(Q(2R))}\right).$$

Let (x_1, t_1) and (x^2, t^2) be two points in $Q(R)$ with distance

$$d_0 = \sqrt{d(x^2, x^1)^2 + |t_2 - t_1|} < R/2$$

and $t_1 \le t_2$. Let $\bar f = \chi_{Q(x^2,t^2,2d_0)} f$. Let $\bar u$ be the solution with right hand side $\bar f$ in the half space. Then, by the same argument $\|\bar u\|_{L^\infty} \le c\|f\|_{\dot C^\beta} d_0^\beta$.

Let $A_j = Q(x^2, t^2, 2^{-j}R)\backslash\big[Q(x^2, t^2, 2^{-j-1}R) \cup Q(x_2, t_2, 2d_0)\big]$ and $f_j = \chi_{A_j} f$. Let u_j be the solution with right hand side f_j on the whole half space. Then, applying the previous estimate once more,

$$\|u_j\|_{L^\infty(Q(2R))} \le c\|f_j\|_\infty \le c(2^{-j}R)^\beta$$

by the bound on the fundamental solution in Lemma 4.5.7 and by the definition of f_j. Let $M = \ln_2 R - \ln_2 d_0 - 1$. By construction $u_j = 0$ for $j > M$. If $j \le M$ we apply by Theorem 4.5.5 in $Q(x^2, t^2, 2^{-j-1}R)$:

$$|u_j(x,t) - u_j(y,s)| \le c\left(\frac{d_0}{2^{-j}R}\right)^\alpha (2^{-j}R)^\beta = c d_0^\beta \left(\frac{d_0}{2^{-j}R}\right)^{\alpha-\beta} \|f\|_{\dot C^\beta}.$$

Let $u_M = \sum_{j=0}^M u_j$. Then, since $\sum_{j=-\infty}^M \left(\frac{d_0}{2^{-j}R}\right)^{\alpha-\beta} \le 2$,

$$\frac{|u_M(x_2,t_2) - u_M(x_1,t_1)|}{d_0^\beta} \le \sum_{j=0}^M \frac{|u_j(x_2,t_2) - u_j(x_1,t_1)|}{d_0^\beta} \le 2c\|f\|_{\dot C^\beta}.$$

The final estimate follows if we compose all the parts and use Theorem 4.5.5 once more:

$$\begin{aligned} \|u\|_{\dot C^\alpha(Q)} &\le \|u_M\|_{\dot C^\alpha(Q)} + \|\bar u\|_{\dot C^\alpha(Q)} + \|u - u_M - \bar u\|_{\dot C^\alpha(Q)} \\ &\le c\big[\|f\|_{\dot C^\beta} + R^{-\beta}|Q|_{\sigma,0}^{-1}\|u - u_M - \bar u\|_{L^1(Q(x_0,t_0,2R),\mu_\sigma)}\big] \\ &\le \bar c\big[\|f\|_{\dot C^\beta} + R^{-\beta}|Q|_{\sigma,0}^{-1}\|u\|_{L^1(Q(x_0,t_0,2R),\mu_\sigma)}\big]. \end{aligned}$$

This completes the proof. □

Proof of Lemma 4.5.7: The kernel estimates of Theorem 4.4.2 imply, with $\rho^2 = d(x,y)^2 + |t - s|$,

$$y_n^{-\sigma}(y) g(x,t,y,s) \le c(n,\sigma,\lambda)|B(x,\rho)|_\sigma^{-1}.$$

Moreover $u(y,s) = \left(\frac{x_n}{y_n}\right)^\sigma g(x,t,y,s)$ satisfies

$$\partial_s u + \partial_{y_i}(y_n a^{ji}\partial_{y_j} u) + \sigma a^{jn}\partial_{y_j} u = 0.$$

It is the fundamental solution to the $L^2(\mu_\sigma)$-adjoint problem. The estimate follows from an energy estimate and the kernel estimates. More precisely: Let $\tilde{Q} = Q(x^0, t^0, 2R)\backslash Q(x^0, t^0, R)$ and $\bar{Q} = Q(x^0, t^0, 3R)\backslash Q(x^0, t^0, R/2)$. The energy estimate gives as above an estimate of the $L^2(\mu^0_{\sigma+1})$ norm of the derivative in terms of the L^∞ norm:

$$\int_{\tilde{Q}} |\nabla u|^2 d\mu^0_{\sigma+1} \le c_1 R^{-2} \int_{\bar{Q}} |u|^2 d\mu^0_\sigma \le c_2 x_n^{2\sigma} R^{-2} |B(x,R)|_\sigma^{-2} |\bar{Q}|_{\sigma,0}.$$

Using the volume estimates of Lemma 4.3.3 several times, together with the previous estimates,

$$\begin{aligned}
\int_{\tilde{Q}} |y_n^\sigma \nabla_y (y_n^{-\sigma} g(x,t,y,s))|\, dy\, ds &\le x_n^{-\sigma} \int_{\tilde{Q}} |\nabla_y u(y,s)|\; d\mu^0_\sigma(y,s) \\
&\le |\tilde{Q}|^{1/2}_{\sigma-1,0} x_n^{-\sigma} \|\nabla_y u\|_{L^2(\tilde{Q},\mu_{\sigma+1})} \\
&\le c(n,\sigma,\lambda) |\tilde{Q}|^{1/2}_{\sigma-1,0} |\tilde{Q}|^{1/2}_{\sigma,0} R^{-1} |B(x,R)|_\sigma^{-1} \\
&\le c(n,\sigma,\lambda) R |B(x,R)|^{1/2}_{\sigma-1} |B(x,R)|_\sigma^{-1/2} \\
&\le c(n,\sigma,\lambda) \frac{R}{R+\sqrt{x_n}} \le c(n,\sigma,\lambda),
\end{aligned}$$

which is the assertion of Lemma 4.5.7. □

4.6 The model operator

In the sequel we shall study the degenerate equation with continuous (or even more regular) coefficients. We have reached a turning point: now good estimates for the model operator are an essential tool for further study. Still they rely on the previous estimates and it is not clear whether there is a shortcut to obtain the estimates for the model operator much simpler. Besides that we will rely on some of the estimates proved for measurable coefficients - in contrast to the Euclidean situation we do not have an explicit formula for the fundamental solution.

The estimates for the fundamental solution together with simple L^2-estimates show that the model operator below defines singular integral operators in the same way as the Riesz operators. We consider $Lu = w$, again in the sense of Definition 4.4.1, where

$$Lu := u_t - x_n^{-\sigma} \sum_{i=1}^n \partial_i (x_n^{1+\sigma} \partial_i u). \tag{4.43}$$

For suitable functions w there exists exactly one energy solution. We denote the operator $w \to u =: Tw$. In the sequel we shall derive estimates for T. They have to be understood in the sense that T extends to an operator on various spaces.

We want to show that certain operators are singular integral operators. For that we need two types of estimates: kernel estimates, and L^2 operator estimates. We begin by studying the latter.

Lemma 4.6.1. *Suppose that $\sigma > -1$. There exists c depending only on n and σ such that*

$$\|\nabla_x T\big(x_n^{-\sigma}\partial_i(x_n^{1+\sigma}f)\big)\|_{L^2(\mu^0_{\sigma+1})} \le c\|f\|_{L^2(\mu^0_{\sigma+1})},$$

and

$$\|x_n D_x^2 Tw\|_{L^2(\mu^0_\sigma)} + \|\nabla_{x,t}Tw\|_{L^2(\mu^0_\sigma)} \le c\|w\|_{L^2(\mu^0_\sigma)}.$$

Proof. We provide formal proofs. A rigorous justification is not difficult. Suppose that $u = T[x_n^{-\sigma}\partial_i(x_n^{1+\sigma}f)]$. Multiplication of $Lu = x_n^{-\sigma}\partial_i(x_n^{1+\sigma}f^i)$ by $x_n^\sigma u$ and integration gives

$$\|\nabla u\|^2_{L^2(\mu_{\sigma+1})} \le \|f\|_{L^2(\mu_{\sigma+1})}\|\nabla u\|_{L^2(\mu_{\sigma+1})}$$

which implies the first estimate. Integrating by parts on $\int Lux_n^\sigma u_t\, dx\, dt$ and the Cauchy-Schwarz inequality imply the bound for $\partial_t(Tw)$.

We compute

$$\int_H [\nabla(x_n^{1+\sigma}\nabla u)]\partial_n u dx = -\int_H x_n^{1+\sigma}\partial_i u\partial_i\partial_n u\, dx = \frac{1+\sigma}{2}\int_H |\nabla u|^2\, d\mu_\sigma$$

which together with the bound for $\partial_t Tw$ implies the bound for $\nabla_x Tw$. The bound for the second order derivatives follows from the ones for first order derivatives and

$$\begin{aligned}\|x_n\Delta u\|^2_{L^2(\mu_\sigma)} &= -\int x_n^{2+\sigma}(\partial_j u)\partial_j\Delta u\, dx - (2+\sigma)\int x_n^{1+\sigma}\partial_n u\Delta u\, dx\\ &= \sum_{i,j}\|x_n\partial^2_{ij}u\|^2_{L^2(\mu_\sigma)} + (2+\sigma)\Big[2\int x_n^{1+\sigma}(\partial_j u)\partial^2_{jn}u + (1+\sigma)\int x_n^\sigma(\partial_n u)^2 dx\Big]\\ &= \sum_{i,j}\|x_n\partial^2_{ij}u\|^2_{L^2(\mu_\sigma)} - (2+\sigma)(1+\sigma)\|\nabla_{\mathbb{R}^{n-1}}u\|_2^2.\end{aligned}$$

This implies, together with the estimates for first order derivatives, the last inequality. The proof is complete. $\square$

We need however estimates of derivatives of the kernel function. The key is the following simple proposition.

Proposition 4.6.2. *Suppose that $Lu = 0$ in $Q(x^0, t^0, R)$. Then, with c depending on n, σ, m and $|\alpha|$*

$$\left|\partial_t^m \partial_x^\alpha u(x^0, t^0)\right| \le c \left(R^2 + \sqrt{x_n^0} R\right)^{-|\alpha|} R^{-2m} \|u\|_{L^\infty(Q(x^0,t^0,R))}$$

for $0 \le k, m$ and all multi-indices α.

Proof. Scaling reduces the estimate to $Q(0,0,1)$ or $Q((0,\dots,1),0,R)$ with $R \le 1$. In both cases it suffices to estimate the derivative in the center at time $t = 0$. In the second case the assertion follows from the quantitative version of the fact that solutions to parabolic equations with smooth coefficients are smooth. Hence it suffices to consider $x^0 = 0$.

Tangential derivatives commute with the operator L. Thus it suffices to prove boundedness of first order tangential derivatives to obtain boundedness of tangential derivatives of all orders.

The function u is smooth away from $x_n = 0$ and the boundary of the cube since it satisfies a uniformly parabolic equation there. Let $Q(R) = Q(0,0,R)$. We claim that $Lu = w$ in $Q(1)$ implies

$$\|\nabla u\|_{L^2(Q(R),\mu_\sigma^0)} \le c\Big[\|u\|_{L^\infty(Q(1))} + \|w\|_{L^2(Q(1),\mu_\sigma^0)}\Big] \tag{4.44}$$

for all $R < 1$. Let $\psi \in C_0^\infty$. Then

$$L(\psi u) = (L\psi)u - 2\partial_i(x_n(\partial_i\psi)u) + 2[\partial_i(x_n\partial_i\psi)]u + \psi w$$

hence by Lemma 4.6.1

$$\|\nabla u\|_{L^2(Q(R),\mu_{\sigma+1}^0)} \le c\Big(\|u\|_{L^2(Q(1),\mu_\sigma^0)} + \|w\|_{L^2(Q(1),\mu_\sigma^0)}\Big).$$

Moreover $L(\psi u) = (L\psi)u - 2x_n(\partial_i\psi)(\partial_i u) + w$ and, again by Proposition 4.6.1,

$$\|\nabla u\|_{L^2(Q(R),\mu_\sigma^0)} \le c\big[\|w\|_{L^2(Q(1),\mu_\sigma^0))} + \|u\|_{L^\infty(Q(1))} + \|x_n \nabla u\|_{L^2(Q(1),\mu_\sigma^0)}\big],$$

which implies inequality (4.44).

Let $v(x,t) = h^{-1}(u(x + he_j, t) - u(x,t))$. It satisfies $Lv = 0$. Hence, if $R_1 < R_2 < R_3 < 1$,

$$\|v\|_{L^\infty(Q(R_1))} \le c_1\|v\|_{L^2(Q(R_2))} \le c_2\|\nabla u\|_{L^2(Q(R_3))} \le c_3\|u\|_{L^\infty(Q(1))}$$

uniformly in h because of the previous estimates and Proposition 4.5.1. In the limit $h \to 0$ we get

$$\|\nabla_{\mathbb{R}^{n-1}} u\|_{L^\infty(Q(R))} \le c\|u\|_{L^\infty(Q_1)}.$$

This argument can be iterated and implies boundedness of all tangential derivatives.

Theorem 4.5.5 implies that all tangential (spatial) derivatives are Hölder continuous. Let $v^j = \partial^j_{x_n} u$. It satisfies

$$v^j_t - \nabla(x_n \nabla v^j) - (\sigma + j)\partial_n v^j = j\Delta_{\mathbb{R}^{n-1}} v^{j-1}.$$

Inductively we obtain from Theorem 4.5.6 that v^j together with its tangential spatial derivatives is Hölder continuous provided we show inductively that $\nabla v^j \in L^2(Q(R), \mu^0_{\sigma+j+1})$. Only the estimate of the term $\partial_n v^j$ is not obvious. It follows however from the last estimate of Lemma 4.6.1 applied to v^{j-1}. Inductively we obtain bounds for all spatial derivatives, and hence, solving inductively $\partial^j_t Lu = 0$ for $\partial^{j+1}_t u$, for all derivatives. We conclude that $u \in C^\infty(Q(R))$ for all $R < 1$, together with estimates in term of $\|u\|_{L^\infty(Q(1))}$. □

We apply Proposition 4.6.2 to the fundamental solution. Let

$$V(x, t, y, s; \sigma) := |Q(y, s, \sqrt{d(x, y)^2 + |t - s|})|_\sigma$$

and, suppressing the arguments,

$$D := \frac{\sqrt{d(x, \tilde{x})^2 + |t - \tilde{t}|} + \sqrt{d(y, \tilde{y})^2 + |s - \tilde{s}|}}{\sqrt{d(x, y)^2 + |t - s|} + \sqrt{d(\tilde{x}, \tilde{y})^2 + |\tilde{t} - \tilde{s}|}}.$$

Corollary 4.6.3 (Kernel estimates). *Let g be the fundamental solution.*

1. *Let $k(x, t, y, s)$ be either of the following:*

   $$y_n^{-\sigma} \nabla_{x,t} g(x, t, y, s), \qquad or \qquad y_n^{-\sigma} x_n D^2_x g(x, t, y, s).$$

   Then

   $$k(x, t, y, s) \le c(n, \sigma)[V(x, t, y, s; \sigma)]^{-1}$$

   and if $D \le 1/4$,

   $$|k(x, t, y, s) - k(\tilde{x}, \tilde{t}, \tilde{y}, \tilde{s})| \le c(n, \sigma) D / V(x, t, x, s; \sigma).$$

2. *Let* $k(x,t,y,s) = \partial_{x_i}\partial_{y_i}(y_n^{-\sigma} g(x,t,y,s))$. *Then*

$$|k(x,t,y,s)| \le cV(x,t,y,s;\sigma+1)^{-1}$$

and, if $D < 1/4$,

$$|k(x,t,y,s) - k(\tilde{x},\tilde{t},\tilde{y},\tilde{s})| \le c(n,\sigma)D/V(x,t,x,s;\sigma+1).$$

Proof. Let $t > s$ and $d_0 = \max\{d(x,y), \sqrt{t-s}\}$. The kernel estimate (4.20) implies

$$y_n^{-\sigma} g(x,t,y,s) \le c(|B(x,d_0)|_\sigma + |B(y,d_0)|_\sigma)^{-1} e^{-\frac{d(x,y)}{2c_2(t-s)}}. \tag{4.45}$$

This can be easily checked in the three cases $t \ge d(x,y)^2$, $t \le d(x,y)^2$ and $y_n \le d(x,y)^2$, and $t \le d(x,y)^2 \le y_n$. Proposition 4.6.2 implies (applied in $Q(x,t,d_0)$ to the fundamental solution)

$$|\nabla_{x,t} g| \le c d_0^{-2}(|B(x,d_0)|_\sigma + |B(x,d_0)|_\sigma)^{-1} \le cV(x,t,y,s)^{-1}.$$

Similarly

$$\begin{aligned} |D_x^2(y_n^{-\sigma} g)| + |\nabla_x \nabla_y y_n^{-\sigma} g| &\le c\frac{1}{d_0^2(d_0^2 + x_n + y_n)}(|B(x,d_0)|_\sigma + |B(x,d_0)|_\sigma)^{-1} \\ &\le V(x,t,y,s+1)^{-1} \end{aligned}$$

which follows from Proposition 4.6.2, the fact that $(\frac{y_n}{x_n})^\sigma g(y,t,x,s)$ is a solution to the adjoint problem and estimate (4.45). $\square$

Lemma 4.6.1 and Proposition 4.6.3 are all we need to apply the theory of singular integral operators. We consider the problems $Lu = w$ and $Lv = x_n^{-\sigma}\partial_i(x_n^{1+\sigma} f^i)$.

Corollary 4.6.4. 1. *The operators*

$$\begin{array}{ccccccc} L^2(\mu_\sigma^0) & \ni & w & \to & \nabla_{x,t} Tw & \in & L^2(\mu_\sigma^0) \\ L^2(\mu_\sigma^0) & \ni & w & \to & x_n D_x^2 Tw & \in & L^2(\mu_\sigma^0) \end{array}$$

satisfy the assumptions of Theorem 2.3.8 with the measure μ_σ^0 *and the metric*

$$d_0(x,t,y,s) = \sqrt{d(x,y)^2 + |t-s|}.$$

2. The operators

$$L^2(\mu^0_{\sigma+1}) \ni f^i \to \nabla_x T(x_n^{-\sigma}\partial_i(x_n^{1+\sigma} f^i)) \in L^2(\mu^0_{\sigma+1})$$

satisfy the assumptions of Theorem 2.3.8 with metric d_0 and measure $\mu = \mu^0_{\sigma+1}$.

Hence the conclusion of Corollary 2.3.9 is true. This is one of the basic results, which we formulate for later use.

Theorem 4.6.5 (Weighted L^p estimates). *The estimates*

$$\|x_n^\rho \nabla_{x,t} Tw\|_{L^p} + \|x_n^{1+\rho} D_x^2 Tw\|_{L^p} \le c\|x_n^\rho w\|_{L^p}$$

hold if $-1/p < \rho < \sigma + 1 - 1/p$ and

$$\|x_n^\rho \nabla T(x_n^{-\sigma}\partial_i(x_n^{1+\sigma} f^i))\|_{L^p} \le c\|x_n^\rho f\|_{L^p}$$

hold if $-1/p < \rho < \sigma + 2 - 1/p$.

Proof. Sufficiency follows from Corollary 4.6.4, Lemma 4.3.6 and Theorem 2.3.8. ☐

Remark 4.6.6 (Necessity). The conditions on ρ are sharp. The right hand inequality is sharp because $y^{-\rho}k(x,t,y,s)$ has to be in $L^{p'}_{loc}$ for the inequality to hold. This is equivalent to $y_n^{\sigma-\rho} \in L^{p'}_{loc}$ or $\rho - \sigma + 1 - 1/p > 0$. There are solutions with nonconstant derivatives at $\{x_n = 0\}$. Hence the left hand side is sharp for all first order derivatives. For the second derivatives $\rho > -1 - 1/p$ is the necessary condition. There is some slack in the estimate of the second derivatives. This could be used to show that $-1 - 1/p < \rho < \sigma + 1 - 1/p$ is the sharp range for ρ for which the estimate of the second order derivatives is true.

4.7 Continuous coefficients

The kernel estimates and the estimates for the model operator imply estimates for variable coefficient versions of it. This is the topic of this section. We consider the problem

$$\begin{aligned} Lu := u_t - x_n^{-\sigma}\partial_i(x_n^{1+\sigma} a^{ij}(x,t)\partial_j u) &= w + x_n^{-\sigma}\partial_i(x_n^{1+\sigma} f) && \text{in } H \times \mathbb{R}^+ \\ u(.,0) &= u_0 && \text{in } H \end{aligned} \tag{4.46}$$

under various conditions. In particular we require $u, w \in L^2_{loc}(\mu^0_\sigma)$ and $\nabla u, f \in L^2_{loc}(\mu^0_{\sigma+1})$, where loc refers to the relative topology (as always) of the closed halfspace.

Lemma 4.7.1. *We denote by g the fundamental solution of the model problem. Let $\sigma > -1$ and assume that a_0^{ij} is a matrix with $a_0^{ij} + a_0^{ji}$ positive definite. Let $L_0 := \partial_t - x_n^{-\sigma}\partial_i(x_n^{1+\sigma}a_0^{ij}\partial_j .)$. There exists a matrix D and a vector b with $b_n = 0$ such that the fundamental solution of L_0 is given by*

$$g_0 = |D|^{-1}g(Dx + dt, t, Dy + ds, s).$$

Moreover D and b depend analytically on the matrix a_0^{ij}. Here $|D|$ denotes the determinant.

Proof. The operator L_0 can be written as

$$L_0 v = v_t - \partial_i\Big[\frac{x_n(a_0^{ij} + a_0^{ji})}{2}\partial_j v\Big] - \Big[\sigma a_0^{nj} + \frac{a_0^{nj} - a_0^{jn}}{2}\Big]\partial_j v.$$

Let $A = (a_0^{ij} + a_0^{ji})/2$, $b = A^{-1}e_n$, B the matrix with column e_i if $1 \le i < n$ and $b/\sqrt{b_n}$ as n^{th} column. Let $A_0 = B^t AB$. It coincides with A for indices $i, j < n$ and with the identity otherwise. We choose a circle of radius ρ centered at $\rho + \rho^{-1}$ in $\mathbb{C}$, counterclockwise oriented, with ρ large, and we define the real symmetric matrix C, the inverse of the positive definite root of A_0, by the line integral

$$C = \frac{1}{2\pi i}\oint (A_1 - z)^{-1} z^{-1/2} dz.$$

Then $C^{-t}C^{-1} = A_0$ and $1_{\mathbb{R}^n} = C^t B^t ABC$. Let $D = \sqrt{b_n}C^t B^t$ and

$$d_i = \Big(\sigma a^{nj} + \frac{a^{nj} - a^{jn}}{2}\Big)D_{ji} \text{ if } 1 \le i < n \text{ and } 0 \text{ otherwise.}$$

Let $y = Dx + dt$ and $s = t$. Then

$$L_0 v = v_s - y_n^{-\sigma}\partial_{y_i}(y_n^{1+\sigma}\partial_{y_i} v)$$

and hence the fundamental solution is given by the claimed formula. □

4.7.1 L^p- estimates for the gradient

We assume that the functions a^{ij} are uniformly continuous and elliptic and $-1 < \sigma$. Let $x \in H$, $t \in \mathbb{R}$ and $Q(R) = Q(x, t, R)$.

Proposition 4.7.2. *Let $1 < p < \infty$, $-1 < \rho < p - 1 + p\sigma$ and $R \le 1$. We suppose that $w, u \in L^2(Q(2r), \mu_\sigma^0)$ and $\nabla u, f \in L^2(Q(2R), \mu_{1+\sigma}^0)$ and that u is an energy solution to* (4.46) *in $Q(2R)$. Then*

$$\begin{aligned}\|\nabla u\|_{L^p(Q(R),\mu_\rho)} \le c\Big(&\|w\|_{L^p(Q(2R),\mu_\rho)} + \|f\|_{L^p(Q(2R),\mu_\rho)} \\ &+ |Q(2R)|_{\rho,0}^{1/p}[R(R + \sqrt{x_n})]^{-1}|Q(2R)|_{\sigma,0}^{-1}\|u\|_{L^1(Q(2R),\mu_\sigma)}\Big).\end{aligned}$$

Proof. By covering arguments it suffices to prove the estimate for small balls. Hence we may assume that the oscillation of a^{ij} is small. By Lemma 4.7.1 we may assume that $a^{ij}(x^0,t^0)=\delta^{ij}$. Let $b^{ij}=a^{ij}-\delta^{ij}$. As usual we extend w, f and b^{ij} by zero to the whole half space. We claim that there is a unique fixed point of

$$\bar u \to T\Big(x_n^{-\sigma}\partial_i\left[x_n^{1+\sigma}(f^i-b^{ij}\partial_j\bar u)\right]+w\Big).$$

Since b^{ij} is small this follows from a simple fixed point argument using the estimates of Theorem 4.6.5. This reduces the question to the case $w=f=0$.

After scaling and using simple covering arguments it suffices to consider the cases $x=0$, $R=1$ and $x=e_n$ and $R\le 1$. We restrict to the first case since the arguments in the second case are simpler.

Theorem 4.5.1 and energy inequalities imply then

$$\|u\|_{L^\infty(Q(1))}+\|\nabla u\|_{L^2(Q(1)\mu_{\sigma+1})}\le c\|u\|_{L^1(Q(1),\mu_\sigma)}.$$

We use again $L(\psi u)=(L\psi)u-2\partial_i(x_n(\partial_i\psi)u)+2[\partial_i(x_n\partial_i\psi)]u$. The Hardy-Sobolev inequality 4.2.2 (more precisely Lemma 4.2.4) and Theorem 4.6.5 allow to improve upon weight and exponent iteratively. $\square$

4.7.2 Schauder estimates

Let

$$\omega(r):=\sup\left\{|a^{ij}(x,t)-a^{ij}(y,s)|:d(x,y)^2+|t-s|\le r^2\right\}.$$

Proposition 4.7.3. *Suppose that $R\le 1$, $Lu=0$ on $Q(R):=Q(x^0,t^0,R)$ in the sense of Definition 4.18.*

1. *If $\int_0^1\omega(r)\frac{dr}{r}<\infty$, then*
$$|\nabla u(x^0,t^0)|\le c[R(R+\sqrt{x_n})]^{-1}|Q(R)|_\sigma^{-1}\int_{Q(R)}|u|d\mu_\sigma^0. \tag{4.47}$$

2. *Suppose in addition that $0<\alpha<1$ and $\sup_{0\le r\le 1}\omega(r)r^{-\alpha}<\infty$. Then*
$$\sup_{(x,t),(y,s)\in Q(R)}\frac{|\nabla u(x,t)-\nabla u(y,s)|}{(d(x,y)^2+|t-s|)^{\alpha/2}}\le\frac{c}{R^{1+\alpha}(R+\sqrt{x_n})}|Q(2R)|_\sigma^{-1}\int_{Q(2R)}|u|d\mu_\sigma^0.$$

3. *Let g be the fundamental solution. If $d(x,y)+\sqrt{|t-s|}\le 1$ then, with V as in Proposition 4.6.3 and $k(x,t,y,s)=\nabla_x\nabla_y(y_n^{-\sigma}g(x,t,y,s))$,*
$$|k(x,t,y,s)|\le cV(x,t,y,s;\sigma+1)^{-1}.$$

4. Let V, D be as in Proposition 4.4.2 and $0 < \alpha < 1$. If $\omega(r) \le cr^\alpha$ and

$$d(x,\bar x) + d(y,\bar y) + \sqrt{|t-\bar t|} + \sqrt{|s-\bar s|} \le \frac14 d(x,y) + \sqrt{|t-s|} \le 1$$

then

$$|k(x,t,y,s) - k(\bar x,\bar t,\bar y,\bar s)| \le cD^\alpha V(x,t,y,s;\sigma+1)^{-1}.$$

Proof. We provide a proof near the boundary and assume $x_n \le R^2$ and even $x_n = 0$ without loss of generality. Otherwise we are in a uniformly parabolic setting and the proof is simpler. After scaling we may assume that $R = 1$ and we denote $Q = Q(1)$. Theorem 4.5.5 and localized energy estimates allow to replace the L^1 norm on the right hand side of the gradient estimate of Proposition 4.7.3 by $\|u\|_{L^\infty(Q(1))}$. Let $\bar d(x,t) = \sup\{r : Q(x,t,\kappa r) \subset Q\}$ with a large constant κ. We suppose that there exists K such that

$$\sup_{x\in Q} \bar d(\bar d + \sqrt{x_n})|\nabla u(x)| \le K\|u\|_{L^\infty(Q)}$$

and we want to show that then K is bounded in terms of ω, λ, n and σ.

Let $\phi_0 \in C^\infty(\mathbb{R})$ be nonnegative, $\phi_0(s) = 1$ for $s \le 0$ and $\phi_0(s) = 0$ for $s \ge 1$. Let (x^1,t^1) be in the interior of Q. We define, with $d_0 = \bar d(x^1,t^1)$,

$$\phi(x,t) = \phi_0\Big(\frac{|x-x^1|^2}{d_0^2\sqrt{x_n^2 + (x_n^1)^2 + |x-x^1|^2}} - 1/2\Big)\phi_0(\sqrt{|t-t^1|}/d_0 - 1/2).$$

We choose κ so that $2\bar d(x,t) \ge \bar d(x^1,t^1)$ for all (x,t) in the support of ϕ. This is possible with an absolute constant κ by Lemma 4.3.2. Then $|\nabla_{x,t}\phi| + |\partial_i(x_n\partial_j\phi)| \le c\bar d(x^1,t^1)^{-2}$ and the infimum of $\bar d$ is of the same size as its supremum in the support of ϕ.

Using the fundamental solution to the operator L_0 with frozen coefficients $a_0^{ij} = a^{ij}(x^1,t^1)$, g_0 we obtain a representation of $\nabla u(x^1,t^1)$ as follows. Let $v(x,t) = x_n^{-\sigma} g_{(x^1,t^1)}(y,s,x,t)$. Then $\int [L_0(\phi u)]v\, d\mu_\sigma\, dt = \phi(y,s)u(y,s)$ and

$$\begin{aligned}
0 &= \int_Q (Lu)\phi v\, d\mu_\sigma^0(x,t)\\
&= \int_Q (L_0u)\phi v - x_n(\partial_j\phi)(a^{ij} - a_0^{ij})(\partial_j u)v - x_n\phi(a^{ij} - a_0^{ij})\partial_i u\partial_j v\, d\mu_\sigma^0\\
&= \int_Q (L_0(\phi u))v - \Big[(L_0\phi)u - x_n(a_0^{ij} + a^{ij})(\partial_j u)(\partial_i\phi)\Big]v\, d\mu_\sigma^0 + J_2\\
&= \phi(y,s)u(y,s) - J_1 + J_2.
\end{aligned}$$

The term J_1 consists of all summands containing derivatives of ϕ and J_2 those containing the difference $a^{ij} - a_0^{ij}$ We differentiate with respect to y. Then $\nabla u(x^1, t^1) = (\nabla_y J_1 - \nabla_y J_2)|_{y=x^1, s=t^1}$. We estimate the gradient of J_1

$$\begin{aligned}\Big|\int_Q (L_0\phi) u \nabla_y v d\mu_\sigma\Big| &\le c d_0^{-2} \|u\|_{L^\infty(Q)} |Q|_\sigma [d_0(d_0 + \sqrt{x_n^1})]^{-1} |B(x^1, d_0)|_\sigma^{-1} \\ &= c[d_0(d_0 + \sqrt{x_n^1})]^{-1} \|u\|_{L^\infty(Q)}\end{aligned}$$

and, using again localized energy estimates,

$$\begin{aligned}&\Big|\int_Q (a^{ij} + a_0^{ij}) \partial_j u (\partial_i \phi) \nabla_y v d\mu^0_{\sigma+1}\Big| \\ &\quad \le c d_0^{-1} \|\nabla u\|_{L^2(Q,d_0),\mu_{\sigma+1})} |Q(x^1, t^1, d_0)|_{\sigma,0}^{1/2} [d_0(d_0 + \sqrt{x_n^1})]^{-1} |B(x^1, d_0)|_\sigma^{-1} \\ &\quad \le c[d_0(d_0 + \sqrt{x_n})]^{-1} \|u\|_{L^\infty(Q)}.\end{aligned}$$

The estimate of ∇J_2 is more interesting. We decompose $Q(x^1, t^1, d_0)$ into dyadic annuli $A_j = Q(x^0, t^0, 2^{-j} d_0) \backslash Q(x_0, t_0, 2^{-1-j} d_0)$ for $j \ge 1$ and $Q_0 = Q \backslash Q(x^0, t^0, 2^{-N} d_0)$ where N will be determined later. The integral over Q_0 is estimated as above. Now

$$\begin{aligned}|\nabla J_2| \;&\le c \|\nabla u\|_{L^\infty(Q)} \sum_{i=N+1}^\infty \Big(\omega(2^{-i} d_0) |Q(x^1, t^1, 2^{-i} d_0)|_{\sigma+1,0} \\ &\qquad \times 2^j d_0^{-1} [2^{-j} d_0 + \sqrt{x_n^1}]^{-1} |B(x^1, t^1, 2^{-i} d_0)|_\sigma^{-1} \Big) \\ &\qquad + c(N) [d_0(d_0 + \sqrt{x_n^1})]^{-1} \|u\|_{L^\infty(Q)} \\ &\le \bar{c} K d_0^{-2} \|u\|_{L^\infty(Q)} \sum_{i=N+1}^\infty \omega(2^{-i} d_0) + c(N) [d_0(d_0 + \sqrt{x_n})]^{-1} \|u\|_{L^\infty(Q)}\end{aligned}$$

where we used the kernel estimates of Corollary 4.4.2 and the geometric estimates of Lemma 4.3.3 several times. We choose N large so that

$$\bar{c} \sum_{i=N}^\infty \omega(2^{-i}) \le 1/2.$$

Then

$$\bar{d}^2(x,t) |\nabla u(x,t)| \le (K/2 + c) \|u\|_\infty.$$

We obtain the gradient estimate with $K = 3c$. This is an apriori estimate assuming boundedness of the gradient. There are various ways to get around this. By Theorem 4.7.2 we know that $\nabla u \in L^p$ for all p. The argument above can be made rigorous if the coefficients are Hölder continuous, using Theorem 4.7.2. Then an approximation procedure gives the full assertion.

We turn to the proof of the Hölder continuity of the gradient. We use the first part to reduce the problem to proving the estimate

$$\sup_{(x,t),(y,s)\in Q(R/2)} \frac{|\nabla u(x,t)-\nabla u(y,s)|}{(d(x,y)^2+|t-s|)^{\alpha/2}} \le cR^{-\alpha}\|\nabla u\|_{L^\infty(Q(R))}.$$

We assume again $x_n^0=0$ and $R=1$. Let $(x^1,t^1)\in Q(R/2)$. We want to estimate $\nabla u(x^2,t^2)-\nabla u(x^1,t^1)$. It suffices to consider the case $d(x^2,x^0)+\sqrt{|t^2-t^0|}+d(x^1,x^0)+\sqrt{|t^1-t^0|}\le 1/4$. Moreover we may assume that $d(x^2,x^1)+\sqrt{|t^2-t^1|}\le 1/(8\kappa)$, otherwise there is nothing to show. Then, with upper indices for J_1^j resp. J_2^j the terms corresponding to (x^2,t^2) resp. (x^1,t^1),

$$\nabla u(x^2,t^2)-\nabla u(x^1,t^1)=(\nabla_y J_1^2-\nabla_y J_1^1)-(\nabla_y J_2^2-\nabla_y J_2^1).$$

The estimation of the terms involving J_1 is straight forward and we turn to the ones involving J_2. Let $\tilde R=4\sqrt{d(x^1,x^0)^2+|t^1-t^0|}$. Now

$$\int_{Q(x^1,t^1,\tilde R)} |a^{ij}-a_0^{ij}||\nabla u||\nabla v_1|\,d\mu_{\sigma+1}\le c\int_0^{\tilde R}\omega(r)\frac{dr}{r}\|\nabla u\|_{L^\infty(Q)}\le c\tilde R^\alpha/\alpha$$

by previous arguments. The same is true for the term involving $a_2^{ij}=a^{ij}(x^2,t^2)$ and v_2, the equivalent to v for the point (x^2,t^2). Moreover, with $A=Q\backslash Q(x^1,t^1,\tilde R)$ and $a_2^{ij}=a^{ij}(x^2,t^2)$,

$$\left|\int_A\Big[(a^{ij}-a_2^{ij})\partial_j(\nabla_y(v^2-v^1))+(a_2^{ij}-a_1^{ij})\partial_i\nabla_y v^1\Big]\partial_i u\,d\mu_{\sigma+1}\right|\le c\tilde R^\alpha\|\nabla u\|_{L^\infty(Q)}.$$

The estimates for g follow as in Proposition 4.6.3. $\square$

Theorem 4.7.4 (Schauder estimates). *Suppose that $0<\beta<\alpha<1$ and $\omega(r)\le cr^\alpha$ for $0\le r\le 1$, $0\le R\le 1$, $Q(R)=Q(x^0,t^0,R)$. If $f^i\in C^\beta(Q(2R))$, $w\in L^\infty(2Q)$ and $Lu=x_nw+y_n^{-\sigma}\partial_i(y_n^{1+\sigma}f^i)$ in the energy sense then*

$$\begin{aligned}\|\nabla u\|_{\dot C^\alpha(Q(R))}\le\ & c(R)\|w\|_{L^\infty(Q(2R))}+c\|f\|_{\dot C^\beta(Q(2,R))}\\ &+R^{-\alpha}(R(R+\sqrt{x_n^0})^{-1}|Q(2R)|_\sigma^{-1}\int_{Q(2R)}|u|\,d\mu_\sigma^0\end{aligned}$$

Proof. This follows from the pointwise estimates for $\nabla_x\nabla_y[y_n^{-\sigma}g(x,t,y,s)]$ in Theorem 4.7.3 and similar estimates for $\nabla_x g(x,t,y,s)$ in the same way as Theorem 4.5.6 followed from Theorem 4.5.5 and the kernel estimates. $\square$

4.7.3 The Cauchy problem in Hölder spaces

We turn to the inhomogeneous problem in the bounded domain $\Omega \subset \mathbb{R}^n$ with smooth boundary. Let $\psi \in C^2(\overline{\Omega})$ satisfy $|\nabla\psi| + \psi \geq \delta > 0$, $\psi|_{\partial\Omega} = 0$. We assume that $a^{ij} \in C^\alpha(\Omega \times [0,T])$ is such that $a^{ij} + a^{ji}$ is uniformly positive definite. Let $\sigma > -1$. We define $L_\sigma = u_t - \psi^{-\sigma}\partial_i(\psi^{1+\sigma}a^{ij}(x,t)\partial_j u)$ and consider the problem

$$\begin{aligned} Lu &= \psi w + \psi^{-\sigma}\partial_i(\psi^{1+\sigma} f) && \text{in } \Omega \times \mathbb{R}^+ \\ u(.,0) &= u_0 && \text{in } \Omega \end{aligned} \tag{4.48}$$

where $u_0, \nabla u_0 \in C^\alpha(\Omega)$ and $f^i \in C^\alpha$, $w \in L^\infty$. The Hölder norms here are defined with respect to the Carnot-Carathéodory metric or, equivalently, with respect to

$$d(x,y) = \frac{|x-y|}{\sqrt{\psi(x) + \psi(y) + |x-y|}}.$$

Theorem 4.7.5. *There exists a unique energy solution u to problem (4.48) with $u, \nabla u \in C^\alpha(\Omega \times [0,T])$ which satisfies (with the notation $\Omega_T = \Omega \times [0,T]$ and $c = c(n,\sigma,\lambda,\Omega)$),*

$$\|\nabla_x u\|_{C^\alpha(\Omega_T)} \leq c\Big(\|\nabla_x u_0\|_{C^\alpha(\Omega)} + \|w\|_{L^\infty(\Omega_T)} + \|f\|_{C^\alpha(\Omega_T)}\Big).$$

Proof. We begin with several reductions. First, if we consider $v = u - u_0$ instead of u then we are left with homogeneous initial data. Hence we may and do assume that $u_0 = 0$. Second, it suffices to consider the estimate for $t \in \mathbb{R}$ because we can extend the data and the solution to the whole real line by $u(x,t) = 0$ for $t \leq 0$ and any reasonable compactly supported extension for $t \geq T$ of the data. The restriction then gives the desired solution.

Locally we can flatten the boundary. Hence Theorem 4.7.4 implies the local estimate near the boundary in terms of the energy norm of u. In the interior the equation is uniformly parabolic and the simpler arguments give the local estimates there. This completes the proof. $\square$

4.8 Notes

Weighted Sobolev estimates have been studied by Triebel [132] and Opic and Kufner [112]. They do not cover the range of parameters which we need here. The results in this chapter have strong relations to classical results for elliptic and parabolic problems. Our approach to Gaussian estimates and their application to Hölder regularity and Harnack inequality is similar to the one of Nash [108] in the interpretation of Fabes and Stroock [53], but now for a degenerate (subelliptic) parabolic problem.

It is likely that the approaches of DeGiorgi [50], Nash [108] and Moser [100] could be modified for our problem as well.

The parabolic problem at hand was considered by Shimakura [121], who obtained interesting estimates for its fundamental solution. The elliptic problem $\psi a^{ij}\partial^2_{ij}u + b^i\partial_i u = w$ was studied by Boley, Camus and Metevier [26]. Their result is related to our Theorem 4.7.5. Nevertheless there is a puzzling difference: they obtain Hölder estimates in the Euclidean metric. General degenerate elliptic problems have a long history. There is the survey of Levendorskij and Paneah [90] which contains many references and the monograph [51] by Drabek, Kufner and Nicolosi.

Basic for our treatment is an additional structure of the degenerate problem: there is a special case of equation (4.43) which provides a link to the large class of sub-elliptic equations: we set $\tilde{y}_n = \sqrt{y_n}$ and obtain, if $\sigma = -1/2$,

$$v_t - \tilde{y}_n^2 \Delta_{\mathbb{R}^{n-1}} v - \frac{1}{4}\partial^2_{\tilde{y}_n} v = w.$$

The vector fields $\tilde{y}_n\partial_i, \partial_n, [\tilde{y}_n\partial_i, \partial_{\tilde{y}_n}]$ span $\mathbb{R}^n$ at every point. Hence the spatial part of the differential operator satisfies the Hörmander condition. This class of equations is called sub-elliptic equation. This connection is fruitful beyond the particular value $\sigma = -1/2$. Our methods share with other work on general subelliptic problems much of the spirit, in particular the strong connection between geometry and fundamental solutions.

The study of sub-elliptic problems began with the fundamental contributions by Olejnik and Radkevitch [111], Hörmander [69] and [72] who introduced the so-called Hörmander condition, Rothschild and Stein [118] who introduced the lifting procedure which connects such operators to stratified nilpotent Lie groups. On the geometric side the sub-Riemannian geometry was considered and used in a much more general context by Nagel, Stein and Wainger [107]. Similar geometric situations have been studied by Carathéodory [33] and by Chow [34] in a different context.

Fefferman [57] announced results which strengthen the ties between fundamental solution, estimates and discretization. Jerison, Sanchez-Calle and Fefferman obtained pointwise estimates of the fundamental solution in a sequence of papers with various combinations of the authors [58], [120], [76], [75] and [119]. From a probabilistic point of view the study of pointwise estimates was continued by Kosuoka and Stroock [89], [88], [86] and [87].

The theory of sub-elliptic operators has its main motivation (besides its intrinsic beauty) by the study of strongly pseudo-convex domains in complex analysis and of nilpotent Lie groups (see Varopoulos, Saloff-Coste and Coulon [133]), where these operators occur naturally as Kohn-Laplacian resp. left invariant differential operators. The model operator (4.43) is connected to the group $x \to ax + b$, $a \in \mathbb{R}$, $b \in \mathbb{R}^{n-1}$, $x \in \mathbb{R}^{n-1}$ and the estimates should have some relation to the representation theory of that group. It is not clear however how the estimates of Cowling, Giulini, Hulanicki and Mauceri [41] resp. Damek, Hulanicki and Zienkiewicz [45] are related to the problem at hand.

Sometimes there is an explicit computation of the heat kernel. An elegant computation of it for the Heisenberg group has been found by Klingler [82]. I do not know whether such an explicit formula is available for the model operator here.

Regularity results can be obtained from kernel estimates in various ways. Our choice is motivated by the desire to stress the importance of the fundamental solution and to obtain local estimates. The proofs for the kernel estimates needed for the Schauder estimates are adapted from Grüter and Widman, [67]. The relation between fundamental solution and singular integrals is well-known, in particular if the coefficients are smooth. The most precise description can be obtained in terms of pseudo-differential operators, which at least sometimes can be well adapted to degenerate problems, see Geller [65]. Nevertheless the understanding is much poorer than in the Euclidean case even for the Heisenberg group (see Müller and Ricci [104] and Müller and Stein [105] for some interesting results in that direction).

Non-standard boundary value problems occured at several places in problems with free boundaries: for the Stefan problem in Koch [83], where the problem is simpler in the sense that one may decouple the problem into a problem at the boundary and a more standard parabolic problem, and more difficult in the sense that it is not clear that the operators are singular integral operators. The Navier-Stokes equation with a free boundary with surface tension leads again to a non-standard parabolic problem, see Beale [21], [22] and Solonnikov [122].

Chapter 5

The porous medium equation

5.1 Introduction

We study the motion of an ideal gas in a porous medium. The isothermic equation of state of an ideal gas is

$$p = c\rho^{\gamma}$$

where c and γ are positive constants, ρ is the density and p is the pressure. This is the constitutive law. The continuity equation states

$$\rho_t + \nabla \cdot (u\rho) = 0$$

where u is the velocity of the gas. Darcy's law gives a relation between pressure and velocity: in an isotropic homogeneous porous medium it is given by $u = -k\nabla p$ where k is a positive constant. We obtain $\rho_t - ck\Delta\rho^m = 0$ where $m = \gamma + 1$. We shall normalize $ck = 1$ which can be obtained by changing the x-scale. In the sequel we consider

$$\rho_t - \Delta\rho^m = 0 \tag{5.1}$$

for all $m > 1$. For physical reasons we only consider nonnegative solutions.

There are several explicit solutions. The most interesting of them are the Barenblatt solutions,

$$\rho(x,t) = t^{-\frac{1}{(m-1)+2/n}} \left(1 - \frac{m-1}{2m(n(m-1)+2)} \frac{|x|^2}{t^{\frac{2}{n(m-1)+2}}}\right)_+^{\frac{1}{m-1}},$$

where $f_+ = \max\{f, 0\}$. There is a number of insights which can be gained from these solutions: The Barenblatt solutions are not smooth. They are not even C^2 unless $m < 3/2$, they are classical solutions only if $m < 2$. They have compact support for all t. The pressure $v = \frac{m}{m-1}\rho^{m-1}$ is Lipschitz continuous. The radius of the support is

$$R(t) = \sqrt{2m}\sqrt{n + \frac{2}{m-1}}\; t^{\frac{1}{n(m-1)+2}}.$$

Some of these properties hold in a modified form for general solutions while others are misleading. In the sequel we shall prove large time regularity of solutions as discussed in the introduction to this work: for large t the pressure of solutions is analytic up to the boundary of its support under weak assumptions on the initial data.

5.2 Well-posedness

5.2.1 Existence and Uniqueness

We have seen that the Barenblatt solutions are not classical solutions in general. This raises the question what a good notion of a solution could be. A useful definition is the following.

Definition 5.2.1. *Let $\rho \in L^m_{loc}(\mathbb{R}^n \times [0,T]) \cap C(\mathbb{R}^n \times (0,T))$ be a nonnegative function which satisfies*

$$0 = \int_{\mathbb{R}^n \times (0,T)} \rho^m \Delta\phi + \rho\phi_t \, dx\, dt$$

for all $\phi \in C_0^\infty(\mathbb{R}^n \times (0,T))$. Then we say that ρ is a weak solution in the strip $\mathbb{R}^n \times (0,T)$ to the porous medium equation (5.1).

Definition 5.2.2. *Let μ be a nonnegative measure. We define*

$$\|\mu\|_{r,m} = \sup_{R \geq r} R^{-\frac{2}{m-1}-n}\mu(B(0,R))$$

with obvious modification if we evaluate this norm at a function instead of a measure.

These are good notions because of the following result.

Theorem 5.2.3. *1.* Existence of a trace: *Suppose that ρ is a weak solution in a strip. Then $\|\rho(t)\|_{1,m}$ is uniformly bounded for $0 < t \leq T/2$ and there exists μ such that $\int \rho(t)\phi \, dx \to \int \phi \, d\mu$ as $t \to 0$ for $\phi \in C_0^\infty(\mathbb{R}^n)$.*

2. Existence of a solution: *Suppose that $\|\mu\|_{1,m} < \infty$. Then there exist $T > 0$ and a unique weak solution with initial trace μ in the sense of 1.*

3. A-priori estimates: *Suppose that ρ is a weak solution. Then (in a distributional sense)*
$$\Delta \rho^{m-1} \geq -\frac{n}{(n(m-1)+2)t} \tag{5.2}$$
and
$$\frac{\partial \rho}{\partial t} \geq -\frac{n}{(n(m-1)+2)t}\rho. \tag{5.3}$$

Remark 5.2.4. The continuity requirement for weak solutions could be relaxed, see Dahlberg and Kenig [44]. This is however not important for our work here.

This is a highly nontrivial final result. Existence of a solution under the assumptions of Part 2 has been shown by Benilan, Crandall and Pierre [25]. Existence of the initial trace is due to Aronson and Caffarelli [11]. The proof has been simplified by Dahlberg and Kenig [43].

Pierre [114] showed uniqueness of the solution if $\mu \in L^1(\mathbb{R}^n)$. Dahlberg and Kenig [42] extended the uniqueness result to general initial data. The estimates (5.2) and (5.3) are from Aronson and Benilan [10].

Existence of a unique global weak solution for initial data in L^1 has been shown by Benilan and Crandall [24]. In the sequel we shall only consider initial data in $L^1(\mathbb{R}^n)$.

Lemma 5.2.5. *The following estimate holds:*
$$\|\rho(t)\|_\infty \leq c(n,m) t^{-\frac{n}{n(m-1)+2}} \|\rho_0\|_{L^1}^{\frac{2}{n(m-1)+2}}.$$

The worst case is the Barenblatt solution, see Vazquez [134], Theorem 3. It determines the optimal constant. There is a comparison principle of Aronson, Crandall and Peletier [6] and, in higher dimensions, of Dahlberg and Kenig [43],

Lemma 5.2.6. *Suppose that ρ and $\bar{\rho}$ are solutions to the porous medium equation* (5.1) *in the bounded domain $\Omega \times [0,T]$, continuous up to the boundary. If $\rho \leq \bar{\rho}$ at $t = 0$ and at $\partial\Omega \times [0,T]$ then $\rho \leq \bar{\rho}$ in $\Omega \times [0,T]$. The same statement holds for unbounded domains under weak assumptions on the behavior at ∞.*

5.2.2 Strong local solutions

We define a class of admissible initial data, for which we shall prove local existence of a weak solution with a smooth pressure in the support of the solution.

Definition 5.2.7. *We say that $\rho_0 \in C^{1/(m-1)}$ is admissible of class α, $0 < \alpha < 1$, if it is nonnegative, compactly supported, if $v_0 = \frac{m}{m-1}\rho^{m-1}$ is of class C^1 in the support Ω of ρ_0, if $v_0 + |\nabla v_0| \geq \delta > 0$ in Ω and if there exists $c_\alpha > 0$ with*

$$|\nabla v_0(x) - \nabla v_0(y)| \leq c_\alpha \left(\frac{|x-y|}{\sqrt{d(x,\partial\Omega) + d(y,\partial\Omega) + |x-y|}} \right)^\alpha$$

for all $x, y \in \Omega$. The set of α-admissible initial data ρ_0 is denoted by $A(\alpha)$.

Theorem 5.2.8. *Suppose that $\rho_0 \in A(\alpha)$. Let v be the pressure of the solution to (5.1). Then there exists $t^0 > 0$ and c depending only on δ, α, c_α and the measure of Ω, such that*

$$\sup_{(x,t):\rho(x,t)>0, 0\leq t\leq t^0} |\nabla v(x,t)| < \infty$$

and

$$\sup_{\substack{x,y\in\Omega, 0\leq s,\sigma\leq t^0\\ |x-y|+|t-s|>0}} \frac{|\nabla v(x,s) - \nabla v(y,\sigma)|}{\left(\sqrt{|s-\sigma|} + \frac{|x-y|}{\sqrt{d(x,\partial\Omega)+d(y,\partial\Omega)+|x-y|}}\right)^\alpha}$$

$$\leq c \sup_{\substack{x,y\in\Omega\\ x\neq y}} \frac{|\nabla v_0(x) - \nabla v_0(y)|}{\left(\frac{|x-y|}{\sqrt{d(x,\partial\Omega)+d(y,\partial\Omega)+|x-y|}}\right)^\alpha}.$$

This will be proven in Section 5.5.

5.3 Lipschitz continuity

Expansion of gas

Here we collect several useful and basic estimates. They have a physical meaning and they are essential for studying finer properties of solutions.

The first estimate has been proven by Caffarelli and Friedman [27]. It states that there is a relation between the speed of the gas (more precisely: the velocity of the boundary of the support) and the amount of gas near a point.

Proposition 5.3.1. *Let $L > 0$. There exists a positive constant c_0, depending only on n, m and L such that the following is true: If ρ satisfies* (5.1) *in a weak sense, $x^0 \in \mathbb{R}^n$, $t > 0$, $R > 0$, $0 < \sigma \leq Lt$, if*

$$\rho(y,t) = 0 \text{ for } |y - x| \leq R$$

and $\rho(z, t+\sigma) > 0$ on a set of positive measure in $\{z : |z - x| \leq R/6\}$ then

$$|B(x,R)|^{-1} \int_{B(x,R)} \rho^{m-1}(x, t+\sigma) dx \geq \frac{c_0 R^2}{\sigma}.$$

The second one is a Harnack type inequality (Aronson and Caffarelli [11], Theorem 4.1). It states that, if there is much gas in a ball, it has to cover a ball of a certain size after some time.

Proposition 5.3.2. *There exists $c_1 > 0$ depending on M and n such that*

$$\begin{aligned} R^{-\frac{2}{m-1}} |B(x,R)|^{-1} &\int_{B(x,R)} \rho(y,0)\, dy \\ &\leq c_1 \big(t^{-\frac{1}{m-1}} + t^{n/2} R^{-\frac{2}{m-1}} \rho(x,t)^{1+\frac{(m-1)n}{2}} \big) \end{aligned}$$

for all $x \in \mathbb{R}^n$ and $t > 0$.

In particular: if $\rho(y,0) \sim d(y, \{z : \rho(z,0) = 0\})^{1/(m-1)}$ then the boundary starts to move with uniform speed. There is a partial converse: if there is little gas then it will spread out slowly. We use the notation of Definition 5.2.2.

Proposition 5.3.3. *There exist c_2 and c_3 depending only on n and m such that for every weak solution ρ in the strip $\mathbb{R}^n \times [0,T]$ the following is true: if $0 < t < c_2 \|\rho_0\|_{r,m}^{1-m}$ and $0 \leq |x| \leq r$, then*

$$t^{\frac{n}{n(m-1)+2}} \rho(x,t) \leq c_3 r^2 \|\rho_0\|_{r,m}^{\frac{2}{n(m-1)+2}}.$$

This estimate is due to Benilan, Crandall and Pierre [25], Lemma 4.1. It provides an estimate from above of the support of the solution. If $\|\rho\|_{0,m}$ is finite then $\rho(t,0) = 0$ for $t < t^0 = c_2 \|\rho_0\|_{0,m}^{1-m}$.

The support of $\rho(t)$ is monotonically increasing in time. This follows from inequality (5.3). There is an estimate from below of the measure of the support which is a consequence of Theorem 2 of Vazquez [134], proved by symmetrization techniques:

Proposition 5.3.4. *There exists c_4 depending on n and m with*

$$|\ supp\ \rho(t)| \geq |B(0,1)|(c_4 M^{m-1} t)^{\frac{n}{n(m-1)+2}}$$

where $M = \int \rho_0\, dx$ and where $c_4 = (\frac{2m(n(m-1)+2)}{m-1})^{\frac{n(m-1)+2}{2n}}$ is the sharp constant for the Barenblatt solutions.

Weak solutions are continuous. The next two results have been proven by Caffarelli and Friedman [27], [28].

Theorem 5.3.5. *Let ρ be a bounded weak solution. Then ρ is Hölder continuous.*

Theorem 5.3.6. *Let $S(x) = \inf\{t : \rho(x,t) > 0\}$ and suppose that ρ_0 is compactly supported. Then the function $S(x)$ is Hölder continuous outside the support of ρ_0.*

In the remaining part of this section we follow, with modifications, the work of Caffarelli, Vazquez and Wolanski [30].

Lipschitz continuity of the support at fixed t

The comparison principle Lemma 5.2.6 has strong implications on the support of the solution for large t. This is mainly because of the following application of the comparison principle.

Proposition 5.3.7. *Let ρ be a continuous weak solution. If $K =$ $supp\, \rho(.,0)$ is compact and contained in $\{x : x_n > 0\}$ then $\rho(y, x_n, t) \geq \rho(y, -x_n, t)$ for all $y \in \mathbb{R}^{n-1}$ and $x_n \geq 0$, $0 \leq t < \infty$.*

Proof. Let $v(y, x_n, t) = \rho(y, -x_n, t)$. Then $v(y, 0, t) = \rho(y, 0, t)$ and $0 = v(y, x_n, 0) \leq \rho(y, x_n, 0)$ for all $x_n \geq 0$. The comparison principle, Lemma 5.2.6, implies the statement. □

Corollary 5.3.8. *Suppose that $K \subset B(0,R)$. If $R \leq \frac{|y|^2-|x|^2}{2|y-x|}$ then $\rho(y,t) \leq \rho(x,t)$.*

Proof. We want to apply Proposition 5.3.7. We have to check whether $B(0,R) \subset Z = \{z : (z - \frac{x+y}{2}) \cdot (y-x) \leq 0\}$. The supremum of the scalar product in $B(0,R)$ is assumed at $\frac{R(y-x)}{|x-y|}$. This leads to the condition $|y|^2 - |x|^2 \geq 2R|x-y|$. □

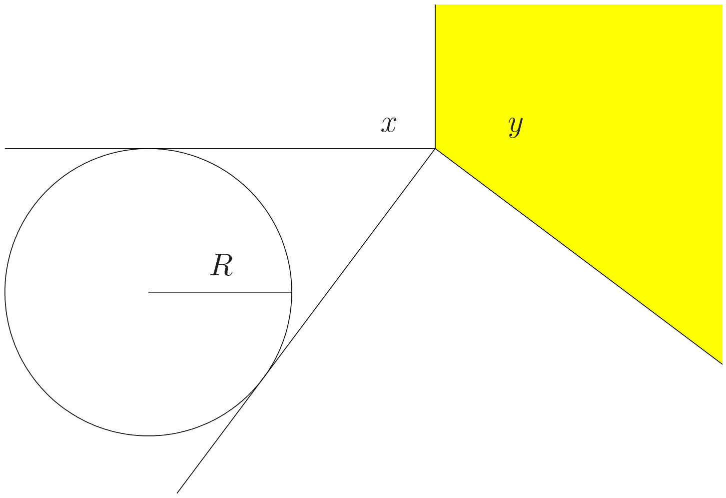

Figure 5.1: The Lipschitz condition in Corollary 5.3.8 and 5.3.9

Corollary 5.3.9. *Suppose that $K \subset B(0,R)$ and let $M = \int_{\mathbb{R}^n} \rho_0 \, dx$. The boundary of the support is a Lipschitz graph outside $B(0,R)$ over $\mathbb{S}^{n-1}$ for all $t > 0$ (this statement will become precise in the proof). Moreover $B(0,R)$ is contained in the support of $\rho(.,t)$ if*

$$t \geq t^0 = c_4^{-1} M^{1-m} (3\pi R)^{n(m-1)+2}.$$

Let $R(t)$ be the smallest number and $r(t)$ the largest number such that

$$B(0,r(t)) \subset \text{ supp } \rho(.,t) \subset B(0,R(t)).$$

Then, using again the constants from Proposition 5.3.3 and Proposition 5.3.4,

$$(c_4 M^{m-1} t)^{\frac{1}{n(m-1)+2}} \leq R(t) \leq R + (M^{m-1} t / c_2)^{\frac{1}{n(m-1)+2}}$$

and

$$R(t) \leq \max\left\{ r(t), \frac{2}{\sqrt{3}} R \right\} + 2\pi R.$$

Remark 5.3.10. In particular the support gets round fast. It is likely that the statement is true with $c_4 = c_2^{-1}$.

Proof. The upper bound for $R(t)$ is a consequence of Proposition 5.3.3: one simply checks, after translating x with $|x| = R(t)$ to the origin that

$$\|\rho_0\|_{0,m}^{1-m} \geq (R(t) - R)^{n(m-1)+2} M^{1-m}$$

hence $t \geq c_2 \|\rho_0\|_{0,m}^{1-m}$ which gives the estimate of $R(t)$ from above. The lower bound of $R(t)$ follows from Proposition 5.3.4:

$$|B(0,1)|R(t)^n \geq |\operatorname{supp} \rho(.,t)| \geq |B(0,1)|(c_4 M^{m-1} t)^{\frac{n}{n(m-1)+2}}.$$

Suppose that $\cos(< x, v >) \geq \frac{R}{|x|}$. Then the assumptions of Corollary 5.3.8 are satisfied with $y = x + v$ and hence $u(x+v) \leq u(x)$. Vice versa: if the assumptions of Corollary 5.3.8 are satisfied for all $y = x + \lambda v$, $\lambda > 0$ then $\cos(< x, v >) \geq \frac{R}{|x|}$.

This implies Lipschitz continuity when we write the boundary of the support as a graph over $\mathbb{S}^{n-1}$, i.e. by $S = \{x \in \mathbb{R}^n : |x| = h(x/|x|, t)\}$ for some h. Then the condition above implies a differential inequality

$$|\nabla_{\mathbb{S}^{n-1}} h| \leq \frac{Rh}{\sqrt{h^2 - R^2}} \Longrightarrow |\nabla_{\mathbb{S}^{n-1}} h| \leq 2R \quad \text{if} \quad h \geq \sqrt{4/3}R.$$

The ball $B(0,R)$ is contained in the support of $\rho(t)$ if $R(t) \geq 3\pi R$. The estimate on t^0 follows from the lower estimate of $R(t)$. More precisely $R < R(t) - 2\pi R \leq r(t)$ for $t > t^0$. □

Uniform estimates away from the boundary

We keep the notation of the previous part. There exists a minimal T such that $B(0,R) \subset \operatorname{supp} \rho(T)$. We fix this T in the sequel. Let $\varepsilon, \eta > 0$. We define the set A by

$$\Big\{(x,t) : t \geq T + \eta, R + \varepsilon(r(t) - R) \leq |x| \leq r(t) - \varepsilon(r(t) - R)\Big\}.$$

Proposition 5.3.11. *There exists $c_5 > 0$ such that*

$$c_5^{-1} \leq t^{\frac{n}{n(m-1)+2}} \rho(x,t) \leq c_5, \qquad t|\rho_t(x,t)| + |x||\nabla\rho| \leq c_5|\rho|,$$

and $x \cdot \nabla\rho \leq -c_5^{-1}\rho$ for all $(x,t) \in A$.

Remark 5.3.12. The bound from below on ρ holds for small x at some larger time. It follows from the Harnack inequality 5.3.2 that $\rho(x,t)$ satisfies the same lower bound for all x with $|x| \leq r(t) - \varepsilon(r(t) - R)$ if t is sufficiently large. It may happen that $\rho(0,t) = 0$ for some $t > T + \eta$. The estimate from above is true for all x and t (Proposition 5.2.5). The arguments below will imply that the estimates for the derivatives are true in the solid ball as above if t is sufficiently large, even when we replace $|x|$ by $R(t)$.

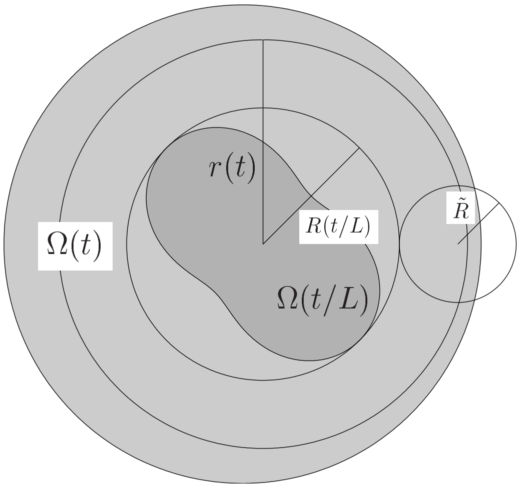

Figure 5.2: The balls: $\Omega(t) =$ supp $\rho(t)$

Proof. Let c_2, c_4 be as in Corollary 5.3.9. We prove the estimate from below first for large t and $|x| \geq R$. If $t \geq t^0$ then $R(t) - r(t) \leq 2\pi$. Hence

$$\begin{aligned} r(t) - R(s) &\geq R(t) - R(s) - 2\pi \\ &\geq (c_4 M^{m-1} t)^{\frac{1}{n(m-1)+2}} - (M^{m-1} s/c_2)^{\frac{1}{n(m-1)+2}} - (2\pi+1)R \\ &\geq \frac{1}{4}(c_4 M^{m-1} t)^{\frac{1}{n(m-1)+2}} \end{aligned}$$

if $s = t/L$ with $L = 4^{n(m-1)+2}/(c_2 c_4)$ and

$$t \geq t^1 = M^{1-m}[4(2\pi+1)]^{n(m-1)+2}/c_4.$$

We apply Proposition 5.3.1 with $\tilde{t} = t/L$, $\tilde{R} = \frac{c_4}{4}(M^{m-1}t)^{\frac{1}{n(m-1)+2}}$ and $|x| = r(t)$. Then, since $L \geq 2$, without loss of generality

$$|B(x^0, \tilde{R})|^{-1} \int_{B(x^0,R)} v(x,t)\, dx \geq \frac{c_0 L \tilde{R}^2}{(L-1)t} \geq \frac{c_0 c_4^2}{32} M^{\frac{2(m-1)}{n(m-1)+2}} t^{-\frac{n(m-1)}{n(m-1)+2}}.$$

There exists a point y in $B(x^0, \tilde{R})$ with

$$\rho(y,t) \geq \left(\frac{c_0 c_4^2}{32}\right)^{1/(m-1)} M^{\frac{2}{n(m-1)+2}} t^{-\frac{n}{n(m-1)+2}}.$$

The density ρ decreases in radial direction for $|x| > R$ by Corollary 5.3.8. Moreover the level set surfaces (outside $B(0, R)$) are Lipschitz surfaces over the sphere. This gives the lower bound for $\rho(x,t)$ for $R \leq |x| \leq r(t)-\varepsilon(r(t)-R)$ and $t > t_1$ (possibly after repeating the argument with some smaller $\tilde{R}$). The same argument works in whole A, but then we have to use the Hölder continuity of S (Theorem 5.3.6) to apply Proposition 5.3.1. The constants there are less explicit.

The upper bound of the size of derivatives follows from local considerations. We digress to:

Lemma 5.3.13. *Let* $Q(x,t,R) = B(x,R) \times (t - R^2, 0)$ *and suppose that* $\rho \in L^\infty(Q(x,t,2R))$ *satisfies* $\Lambda^{-1} \leq \rho \leq \Lambda$ *and* $\rho_t - \Delta\rho^m = 0$ *in* $Q(x,t,2R)$. *Then*

$$\|\rho\|_{C^k(Q(x,t,R))} \leq c(k,\Lambda).$$

Proof of Lemma 5.3.13: The equation can be written in the form

$$\rho_t - m\nabla(\rho^{m-1}\nabla\rho) = 0$$

and the coefficients are uniformly elliptic. Energy arguments show formally that $\nabla\rho \in L^2_{loc}$. They can be made rigorous by approximating the solution by uniformly positive solutions. Hence we can apply parabolic regularity theory (or, to be self contained, Theorem 4.5.5 with $\sigma = 0$), and ρ is Hölder continuous for all sub-domains. Now $\|\nabla\rho\|_{L^p(Q(x,t,r))} \leq c(r,R,p,c_0)$ for all $r < R$ and $1 < p < \infty$ by the device of frozen coefficients. Finite difference techniques give the same estimate for second order derivatives. Hence $\|\rho_t\|_{L^p(Q(x,t,R))} \leq c(n,\Lambda,p)$ for all $1 < p < \infty$ and $r < R$. Finite differences in time direction and Sobolev's imbedding imply boundedness and estimates of ρ_t and second derivatives of ρ. These arguments can be iterated to bound higher order derivatives. □

Lemma 5.3.14. *Suppose that* ρ *satisfies the assumptions of Lemma 5.3.13. If the nonnegative function* $w \in C^2$ *satisfies* $w_t - m\Delta(\rho^{m-1}w)$ *in* $Q(x,t,2R)$ *and* $\Lambda^{-1} \leq \rho \leq \Lambda$ *in* $Q(x,t,2R)$ *then*

$$\sup_{(y,s)\in Q(x,t,R)} w(y,s) \leq c(\Lambda) \inf_{(y,s)\in Q(x,t-2R^2,R)} w(y,s).$$

Proof. This is a version of Harnack's inequality for parabolic equations. It is crucial here that ρ is differentiable. □

Similar results hold in more general domains. This is seen by a Harnack chain, i.e. covering of a path connecting two points under consideration by a chain of cylinders for which Lemma 5.3.14 is applicable.

Proof of Proposition 5.3.11, continuation: We want to estimate $\nabla\rho(y,s)$ for $s \geq T+\eta$ with $|y| \leq r(t) - \varepsilon(r(s)-R)$. The estimates for derivatives of Lemma 5.3.11 are obtained by an application of Lemma 5.3.13 to

$$\tilde{\rho}(x,t) = s^{\frac{n}{n(m-1)+2}}\rho\Big(s^{\frac{1}{n(m-1)+2}}x, st\Big).$$

This is possible by the estimate from below and the Hölder regularity of ρ (Theorem 5.3.5). We turn to Lipschitz continuity of the pressure before we complete the proof.

Lipschitz continuity of the pressure

The pressure $v = \frac{m}{m-1}\rho^{m-1}$ satisfies (in the support of ρ)

$$v_t - (m-1)v\Delta v - |\nabla v|^2 = 0. \tag{5.4}$$

Proposition 5.3.15. *Let v be the pressure of the solution ρ to the porous medium equation* (5.1)*, let T be as in Proposition 5.3.11 and $\eta > 0$. Then there exists $c_6(n,m,M,\eta)$ such that*

$$t|v_t| + |x||\nabla v| \leq c_6 t^{-\frac{n(m-1)}{n(m-1)+2}}$$

if $t \geq T+\eta$ and $|x| \geq R + \frac{1}{4}(r(t)-R)$.

Proof. The estimate holds for $R+\frac{1}{4}(r(t)-R) \leq |x| \leq R+\frac{3}{4}(r(t)-R)$. This is part of the assertion of Proposition 5.3.11 which we have already proven. We define a family of solutions to the pressure equation (5.4) (and hence of pressures to solutions of the porous medium equation). Let $t^0 = T+\eta/2$ and

$$v_\varepsilon(x,t) = (1-K\varepsilon)v((1+\varepsilon)x, t^0 + (1+\varepsilon)^2(1-K\varepsilon)t).$$

We claim that $v_\varepsilon(x,t) \leq v_0(x,t)$ for $t \geq 0$ and $|x| \geq R(0) + \frac{r(t)-R(0)}{4}$ and $|x| \geq R$. We have

$$v_\varepsilon(x,0) - v_0(x,0) = -K\varepsilon v((1+\varepsilon)x,t^0) + v((1+\varepsilon)x,t^0) - v(x,t^0) \leq 0$$

if $|x| \geq R$ and $K > 0$.

In A we may differentiate with respect to ε and obtain

$$\begin{aligned}\frac{\partial v_\varepsilon}{\partial\varepsilon} &= -Kv_\varepsilon + (1+\varepsilon)(1-K\varepsilon)(2-3K\varepsilon-K)t\partial_t v_\varepsilon + (1-K\varepsilon)x\cdot\nabla v_\varepsilon \\ &\leq -Kv_\varepsilon + 3t\partial_t v_\varepsilon \leq 0\end{aligned}$$

if K is sufficiently large and ε is sufficiently small by Proposition 5.3.11. Thus by the comparison principle, Lemma 5.2.6, $v_\varepsilon \le v$ if K is large and ε is small.

The solution is smooth in the interior of its support. We differentiate with respect to ε at $\varepsilon = 0$ and get

$$2(t-t^0)v_t \le Kv - x\cdot\nabla v$$

and, by the pressure equation (5.4) and by (5.2), with $c_7(n,m,K)$, for all $t > T+\eta$ and x with $|x| \ge R + \varepsilon(r(t)-R)$

$$|\nabla v|^2 = v_t - (m-1)v\Delta v \le v_t + \frac{n(m-1)}{n(m-1)+2}\frac{v}{t} \le c_7\frac{v}{t-t^0} + \frac{|x|}{t-t^0}|\nabla v|$$

hence $|\nabla v|^2 \le 2c_7\frac{v}{t-t^0} + \frac{|x|^2}{|t-t^0|^2}$ and, using Lemma 5.2.5 and Corollary 5.3.9 for the last step,

$$\begin{aligned}
|x||\nabla v| + (t-t^0)v_t &\le c\Big(Kv + 3|x||\nabla v|\Big) \\
&\le c\Big(Kv + \frac{9|x|^2}{4(t-t^0)} + (t-t^0)|\nabla v|^2\Big) \\
&\le c\Big((K+2c_7)v + \frac{9|x|^2}{t-t^0}\Big) \\
&\le \tilde{c}_7 t^{-\frac{n(m-1)}{n(m-1)+2}} + \frac{cR^2}{t-t_0}
\end{aligned}$$

if (x,t) satisfies $t > T+\eta$, $|x| > r(t) + \frac{r(t)-R}{4}$ or if $t > t_1$ and x is an arbitrary point in the support of $\rho(t)$. The estimate follows since ρ_t is bounded from below by (5.3). $\square$

Remark 5.3.16. It is worth observing that the constants depend only on n, m, M and K. If $t \ge t^1$ then K can be explicitly described. The dependence on η is less transparent. Here again the Hölder continuity of S enters. In the next section we shall see that under additional assumptions one gets good bounds for derivatives of the function S.

Remark 5.3.17. It is a consequence that

$$0 \le h_t \le \lim_{\delta\to 0} \sup_{0<dist(x,\partial \mathrm{supp}\rho(t))<\delta} |\nabla v(x,t)|^2 |\partial_r v|^{-1}.$$

Indeed, because of estimate (5.2) and equation (5.4),

$$\liminf_{\delta\to 0} \inf_{d(x,\partial\ \mathrm{supp}\ \rho(t))\le\delta} (v_t - |\nabla v|^2) \le 0.$$

By Corollary 5.3.8 $-\partial_r \rho \geq c(\delta)|\nabla \rho|$ for $|x| > R + \delta$ and hence the speed in radial direction of level surfaces (which is given by $-v_t/\partial_r v$) is bounded by $c(\delta)^{-1}$ times the Lipschitz constant of v. Hence we can bound h_t using Proposition 5.3.15 by $c(\delta)^{-1} c_6 (M^{m-1}t)^{-\frac{2(m-1)}{n(m-1)+2}} h^{-1}$.

Proof of Proposition 5.3.11, completion: It remains to prove $x \cdot \nabla \rho(x,s) < -c_5^{-1}\rho$ in A for a possibly larger constant c_5. This is equivalent to $x \cdot \nabla \tilde{\rho} \leq -c_5^{-1}$ in the transformed domain $\tilde{A} = \{(x,t) : (s^{\frac{1}{n(m-1)+2}} x, st) \in A\}$. We want to apply Lemma 5.3.14 to $\frac{x_i^0}{|x^0|}\partial_i \tilde{\rho}$ in a suitable domain. The conclusion of Lemma 5.3.14 holds for more general geometric situations. The constants then depend on the number of application of Lemma 5.3.14 needed to relate two points.

We assume that s is sufficiently large in the sequel since the arguments are easier if s is small. In that case there is no need to scale the domain. We consider the sets

$$A(\delta) = \{(x,t) : 1/2 < t \leq 1, 2R < s^{\frac{1}{n(m-1)+2}}|x|, \tilde{\rho}(x,t) > \delta\}.$$

There exists δ_0 such that $\tilde{A} \cap \{(x,t) : t \geq 1/2\} \subset A(\delta)$ if $\delta \leq \delta_0$. Let $(re_n, 1) \in \tilde{A}$ with $r > 0$, let b be a small number, and $B = A(\delta_0/3) \cap \{x : |\hat{x}| < 2b, x_n > b\}$. Then

$$\partial_n v \leq 0 \text{ in } B$$

by Corollary 5.3.15 and there exists a point $(x, 3/4) \in \bar{B} = A(\delta_0/2) \cap \{x : |\hat{x}| \leq b, x_n > 2b\}$ with

$$|\partial_n \tilde{v}(x, 3/4)| \geq \delta_0/2 \sup_{x \in \sup \tilde{\rho}(.,1)} \{|x|\}^{-1}$$

if δ_0 is sufficiently small. We apply Harnack's inequality (using Hölder regularity of the solution to control the geometry, or more precisely the number of times we apply Harnack's inequality in the form of Lemma 5.3.14) to obtain a uniform lower bound for $-\partial_n \tilde{v}(re_n, 1)$ and hence for $-x \cdot \tilde{v}(x,1)$ for all $(x,1) \in \tilde{A}$. $\square$

Non-degeneracy

Here we want to derive estimates from above of the radial derivative. The arguments are similar as in the last section.

Proposition 5.3.18. *Suppose that ρ_0 is α-admissible (see Theorem 5.2.8). Given $t_0 > 0$ there exist c_8, $c_9 > 0$ such that for $t > t_0$*

$$c_8 v + x_i \partial_i v + ((c_8 + 2)t + c_9) v_t \geq 0.$$

Proof. Let $\varepsilon > 0$. We define the family of pressures

$$v_\varepsilon(x,t) = (1+c_8\varepsilon)v((1+\varepsilon)x,(1+\varepsilon)^2(1+c_8\varepsilon)t + c_9\varepsilon).$$

It suffices to show that $v_\varepsilon(x,t) \geq v(x,t)$ for ε sufficiently small. Then the assertion follows by differentiation with respect to ε at $\varepsilon = 0$. By the maximum principle it suffices to prove this for $t = t_0$ with t_0 small. There exists a smooth local solution in the sense of Proposition 5.2.8, hence v is smooth in its support for small t. We differentiate with respect to ε for t and ε small. Then

$$\begin{aligned}\left.\frac{\partial v_\varepsilon}{\partial \varepsilon}\right|_{\varepsilon=0} &= c_8 v + ((2+c_8)t + c_9)v_t + x\cdot\nabla v \\ &= c_8 v + (m-1)((2+c_8)t + c_9)v\Delta v + ((2+c_8)t + c_9)|\nabla v|^2 + x\cdot\nabla v.\end{aligned}$$

At the boundary $v = 0$ and $v_t = |\nabla v|^2$ is strictly positive there if t^0 is in the domain of existence of the local smooth solution. We choose c_9 large, $c_8 t^0$ small and then c_8 large to ensure $\frac{\partial v_\varepsilon}{\partial\varepsilon}|_{\varepsilon=0,t=t_0} > 0$. This implies $\frac{\partial v}{\partial \varepsilon}|_{\varepsilon=0} \geq 0$ and hence the assertion. ☐

Theorem 5.3.19. *Suppose that ρ_0 is α-admissible and supported in the ball $B(0,R)$. Then there exists $c_{10}(n,m,c_8,c_9,M)$ such that*

$$x\cdot\nabla v(x,t) \leq -c_{10}^{-1} t^{-\frac{n(m-1)}{n(m-1)+2}}$$

for $|x| \geq R(0) + \frac{r(t)-R(0)}{4}$ and $t \geq T + \eta$.

Proof. Let $h : \mathbb{S}^{n-1}\times I \to \mathbb{R}^+$ be such that the boundary of the support of $\rho(t)$ is given by $\{(x,t) : |x| = h(x/|x|,t)\}$. We consider it only at points where $h > R$. We claim that, in a distributional sense,

$$h_t \geq \frac{h}{(c_8+2)t + c_9} \geq (c_4 M^{m-1} t)^{\frac{1}{n(m-1)+2}}((c_8+2)t + c_9)^{-1}.$$

Formally $h_t = -h\frac{v_t}{x\cdot\nabla v}$ which is bounded from below by $h((c_8+2)t+c_9)^{-1}$ because of Proposition 5.3.18. This can be made rigorous in the following way. Let $w = e^{c_8 t/c_9}v$. Then $x_i\partial_i w + ((c_8+2)t + c_9)w_t \geq 0$ in the support of ρ. Now we apply the previous argument to level set surfaces of w: the argument above is true outside $B(0,R)$ without change for all level set surfaces. We obtain the claim for h if we approach the boundary after estimating R from below using Proposition 5.3.9 resp. the by the Hölder regularity of S, Theorem 5.3.6 for t small.

The proof of Proposition 5.3.11 shows that, with $\Sigma(t) = \partial \operatorname{supp} \rho(t)$,

$$|v(x,t)| \geq c_0 \frac{d(x,\Sigma(t))^2}{(t-s)}$$

as long as $t/2 \leq s \leq t$ and $h(x/|x|, s) \leq h(x/|x|, t) - 6d(x, \Sigma(t))$. We choose

$$s = t - 2c_{10} M^{\frac{m-1}{n(m-1)+2}} t^{\frac{n(m-1)+1}{n(m-1)+2}} d(x, \Sigma(t))$$

for x close to Σ. Then $s > T + \eta/2$ if $t > T + \eta$ if x is sufficiently close to the boundary of the support. Hence in that situation

$$|v(x,t)| \geq \frac{c_0}{2c_{10}} M^{-\frac{m-1}{n(m-1)+2}} d(x, \Sigma(t)) t^{-\frac{n(m-1)+1}{n(m-1)+2}}$$

which gives the desired conclusion by an application of Harnack's inequality in the same way as in the proof of Proposition 5.3.11. □

5.4 Coordinates

We will introduce changes of dependent and independent variables in a way going back to von Mises. The main geometric insight is that it is often equivalent to describe a function or its graph. The graph however can be parametrized in different ways. In particular we can parametrize it so that the problem becomes one with fixed domains.

In this section we assume that ρ is a solution to the porous medium equation (5.1) with pressure $v = \frac{m}{m-1}\rho^{m-1}$, which satisfies the pressure equation (5.4). We denote the initial trace of the pressure by v_0.

5.4.1 Local coordinates

We suppose that $v(x,t)$ is near (x^0, t^0) a C^1–function in its support with $v(x^0,t^0) = 0$ and $\nabla_x v(x^0,t^0) \neq 0$. After an affine isometry we may and do assume that $(x^0,t^0) = (0,0)$ and $\nabla v(0,0) = v_0 e_0$ with $v_0 < 0$. After a parabolic scaling we have $v = -x_n + t + o(|x| + |t|)$ in its support.

Locally the graph M of v has the form $\{x_n = u(x_1, \dots, x_{n-1}, v, t)\}$ for some function u. This suggests to introduce new coordinates $x_i = y_i$ if $1 \leq i \leq n-1$, $x_n = v$, $y_n = u$, $s = t$. The map $(y,s) \to (x,t)$ has the

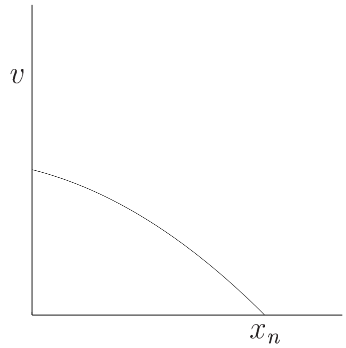

Figure 5.3: Original local coordinates

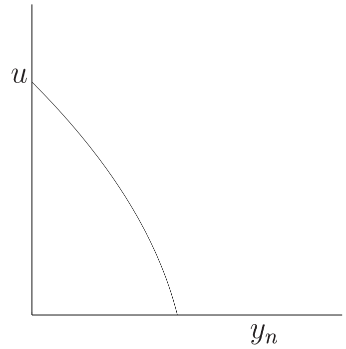

Figure 5.4: Changing dependent and independent variables

derivative

$$\frac{\partial(x,t)}{\partial(y,s)} = \begin{pmatrix} 1 & 0 & \cdots & \frac{\partial u}{\partial y_1} & 0 \\ 0 & 1 & \cdots & \frac{\partial u}{\partial y_2} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{\partial u}{\partial y_n} & 0 \\ 0 & 0 & \cdots & \frac{\partial u}{\partial s} & 1 \end{pmatrix}, \quad \frac{\partial(y,s)}{\partial(x,t)} = \begin{pmatrix} 1 & 0 & \cdots & -\frac{\partial u}{\partial y_1}\Big/\frac{\partial u}{\partial y_n} & 0 \\ 0 & 1 & \cdots & -\frac{\partial u}{\partial y_2}\Big/\frac{\partial u}{\partial y_n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & (\frac{\partial u}{\partial y_n})^{-1} & 0 \\ 0 & 0 & \dots & -\frac{\partial u}{\partial s}\Big/\frac{\partial u}{\partial y_n} & 1 \end{pmatrix},$$

hence

$$\partial_{x_i} v = -\frac{\partial_{y_i} u}{\partial_{y_n} u} \text{ if } 1 \le i \le n-1, \qquad \partial_{x_n} v \partial_{y_n} u = 1, \qquad \partial_t v = -\frac{\partial_s u}{\partial_{y_n} u}$$

and

$$\nabla_x (v \nabla_x v) = (\partial_{y_n} u)^{-1} \left(y_n \Delta_{\mathbb{R}^{n-1}} u - \partial_{y_n} \left(y_n \frac{1 + \sum_{i=1}^{n-1} |\partial_{y_i} u|^2}{\partial_{y_n} u} \right) \right).$$

Let $F(p) = -\frac{1+\sum_{i=1}^{n-1} p_i^2}{p_n}$. Using the previous calculations we rewrite problem (5.4) in the form

$$u_t - (m-1) y_n \Delta_{\mathbb{R}^{n-1}} u - (m-1) |y_n|^{\frac{m-2}{m-1}} \partial_{y_n} \left[|y_n|^{\frac{1}{m-1}} F(\nabla u) \right] = 0 \qquad (5.5)$$

in the set $\{y_n < 0\}$. The coefficients of the elliptic part of the linearization are

$$(a^{ij}(p)) = \begin{pmatrix} 1 & 0 & \cdots & -2p_1/p_n \\ 0 & 1 & \cdots & -2p_2/p_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (1+\sum_{i=1}^{n-1} p_i^2)/p_n^2 \end{pmatrix},$$

which satisfy

$$a^{ij}(p)\xi_i\xi_j \geq \min\{1, p_n^{-2}\}|\xi|^2 \quad \text{and} \quad |a^{ij}\xi_i\eta_j| \leq 2\frac{1+|p|^2}{p_n^2}|\xi|\,|\eta|.$$

We conclude with a remark. It suffices to assume that v is C^2 in the interior, Lipschitz up to the boundary and $\partial_n v < -\delta < 0$ for this construction.

5.4.2 Global coordinates I

Here we want to find coordinates so that the pressure equation (5.4) becomes a quasi-linear degenerate parabolic problem on a fixed domain with analytic nonlinearity. In contrast to the coordinates above we want to work globally in space but local in time. We suppose that $v_0 \in C^1(\mathrm{supp} v_0)$ and $v_0 + |\nabla v_0|^2 \geq \delta > 0$. We denote the graph of v_0 by M_0.

We extend v_0 to a function in $C_b^1(\mathbb{R}^n)$. Let $w(x,t)$ be the bounded solution to the heat equation with initial data w_0. If ε is sufficiently small then

$$|w(x,\varepsilon) - v_0(x)| + |\nabla w(x,\varepsilon) - \nabla v_0(x)||\nabla v_0| \leq \frac{1}{2}(|\nabla v_0(x)|^2 + v_0(x))$$

in a neighborhood of the support of v_0. We fix such an ε. Then the vector field (in $\mathbb{R}^{n+1}$)

$$X(x, x_{n+1}) := (-\nabla w(x,\varepsilon), w(x,\varepsilon))$$

is transversal to the hyper-surface M_0:

$$< X, (-\nabla_x v_0, 1) >= w(x,\varepsilon) + \nabla_x w(.,\varepsilon)\cdot \nabla v_0 \geq \frac{1}{2}(v_0(x) + |\nabla v_0|^2) \geq \frac{\delta}{2}.$$

The vector field X does not depend on x_{n+1} and we suppress its notation. The vector field X is transversal to the graph of $w(.,\delta)$ if δ is positive and sufficiently small. The map

$$\Phi : (y,u) \to (y, w(y,\delta)) + uX(y) \tag{5.6}$$

is a diffeomorphism in a neighborhood of the graph of $w(\delta)$ containing M_0. We fix ε and δ and define $\Psi(x) = w(x, \delta)$. The map Φ is an analytic diffeomorphism. Let Ω be the set which consists of the union of all path components of $\{y : \Psi(y) > 0\}$ with nonempty intersection with the support of $\rho(.,0)$. We may and do assume that there is only one such component. Points in Ω will be denoted by y in the sequel.

The graph M of a solution v can be described by a function $u : \Omega\times[0,t) \to \mathbb{R}^+$ such that

$$\{(x, x_{n+1}) : x_{n+1} = v(x,t)\} = \text{ Range } \{y \to \Phi(y, u(y,t))\}. \tag{5.7}$$

We want to describe M through u. Suppose that v be a sufficiently regular solution to the pressure equation (5.4) with initial data v_0. Then u is well defined for small t. We shall derive a quasi-linear problem for u and solve this problem. The solution u will be smooth and since all the steps are reversible this will give us a solution to equation (5.4).

We turn to the calculation and we denote the determinant of the derivative of Φ by $D(y,u)$. Let $\hat{\Phi}$ be the first n components of Φ and let $\hat{X}$ be the first n components of X. The derivative of the map $(y,s) \to (x,t) = (\hat{\Phi}(y, u(y,s), s))$ is

$$\frac{\partial(x,t)}{\partial(y,s)} = \begin{pmatrix} (\delta_{ij} + u\partial_{y_j}\hat{X}_i + (\partial_{y_j}u)\hat{X}_i) & u_s\hat{X} \\ 0 & 1 \end{pmatrix}.$$

We define $B_{ij} = \delta_{ij} + u\partial_j\hat{X}_i + (\partial_{y_j}u)\hat{X}_i$ and suppose that it is invertible (which holds for the functions v which we consider by construction). Then the inverse is given by $\frac{\partial(y,s)}{\partial(x,t)} = \begin{pmatrix} B^{-1} & -u_sB^{-1}\hat{X} \\ 0 & 1 \end{pmatrix}$ hence

$$v_t = u_s(X_{n+1} - < \nabla\Phi_{n+1}(y,u), B^{-1}\hat{X} >). \tag{5.8}$$

Now $\partial_{x_i}v = \frac{\partial y_j}{\partial x_i}\partial_{y_j}\Phi_{n+1}(y,u)$ and $\nabla_x v = B^{-t}\nabla_y\Phi_{n+1}(y,u)$. Moreover

$$D(y,u) = |B|(X_{n+1}(y,u) - < \nabla_y\Phi_{n+1}(y,u), B^{-1}\hat{X}(y,u) >) \tag{5.9}$$

where the second factor is $X\cdot(-\nabla_x v, 1)$. We rewrite the equation (5.8) using this identity

$$v_t = D(y,u)|B|^{-1}u_s. \tag{5.10}$$

Let E and F be $n \times n$ matrices and let E be invertible. We denote the determinant by $|.|$. Then $|E+F| = |E|(|1+E^{-1}F|) = |E|\ \mathrm{tr}\ E^{-1}F + o(E^{-1}F)$. The derivative

$$\partial_{y_i}(|B|(B^{-1})^{ij}) = |B|\left[(B^{-1})^{lm}(B^{-1})^{ij}\partial_{y_i}B^{ml} - (B^{-1})^{il}(B^{-1})^{mj}\partial_{y_i}B^{lm}\right]$$

vanishes because $\partial_{y_i}B^{ml} = \partial_{y_i}\partial_{y_l}\Phi_m(y,u) = \partial_{y_l}B^{mi}$. Hence

$$\Delta_x v = \sum_{i=1}^n \partial_{x_i}(\partial_{x_i}v) = B^{-1}\nabla_y \cdot (B^{-t}\nabla_y\Phi_{n+1}) = |B|^{-1}\nabla_y \cdot (|B|B^{-1}B^{-t}\nabla_y\Phi_{n+1})$$

and $|\nabla_x v|^2 = |B^{-t}\nabla_y\Phi_{n+1}|^2$. We combine this to a rewritten form of equation (5.4),

$$\begin{aligned} u_s - (m-1)\Phi_{n+1}^{\frac{2-m}{m-1}}\nabla_y \cdot \left(\Phi_{n+1}^{\frac{1}{m-1}}\frac{|B|}{D}B^{-1}B^{-t}\nabla_y\Phi_{n+1}\right) & \\ = \Phi_{n+1} < \nabla_y D^{-1}, |B|B^{-1}B^{-t}\nabla_y\Phi_{n+1} > . & \end{aligned} \tag{5.11}$$

Let

$$F(y,u,\nabla_y u) := (m-1)\frac{\Phi_{n+1}}{\psi}\frac{|B|}{D}B^{-1}B^{-t}\nabla_y\Phi_{n+1}$$

and

$$f(y,u,\nabla_y u) := < \frac{\Phi_{n+1}}{\psi}\nabla_y D^{-1} + D^{-1}\frac{m+2}{m-1}\nabla\left(\frac{\Phi_{n+1}}{\psi}\right), |B|B^{-1}B^{-t}\nabla_y\Phi_{n+1} > .$$

Then equation (5.11) takes the form

$$u_s - \psi^{\frac{2-m}{m-1}}\nabla_y\left(\psi^{\frac{1}{m-1}}F(y,u,\nabla u)\right) = \psi f. \tag{5.12}$$

Proposition 5.4.1. *Let $a^{ij}(y,u,p) = \frac{\partial F^i}{\partial p_j}$. Then $a^{ij} + a^{ji}$ is uniformly positive definite. More precisely*

$$a^{ij}(y,u,p) + a^{ji}(y,u,p) = 2(m-1)\frac{\Phi_{n+1}}{\psi}(B^{-1})^{ik}(B^{-1})^{jk}.$$

Proof. This follows from

$$\begin{aligned} \nabla_y \cdot \left(\frac{\Phi_{n+1}|B|}{\psi D}B^{-1}B^{-t}\nabla_y\Phi_{n+1}\right) &= |B| < B^{-t}\nabla_y\left(\frac{\Phi_{n+1}}{\psi D}\right), B^{-t}\nabla\Phi_{n+1} > \\ &\quad + \frac{\Phi_{n+1}|B|}{\psi D}B^{-1}\nabla \cdot (B^{-t}\nabla\Phi_{n+1}). \end{aligned}$$

Only the second summand contains second order derivatives. Now $\partial_{y_i} B^{-t} = -B^{-t}(\partial_{y_i} B^t)B^{-t}$. The terms containing second derivatives are

$$\begin{aligned}\frac{\Phi_{n+1}|B|}{D\psi}\big(X_{n+1} - < \tilde{X}, B^{-t}\nabla_y\Phi_{n+1} > \big)(B^{-1})^{ik}(B^{-1})^{jk}\partial_{ij}u \\ = \frac{\Phi_{n+1}}{\psi}(B^{-1})^{ik}(B^{-1})^{jk}\partial_{ij}u.\end{aligned}$$

We used equation (5.9). The matrix B is nonsingular, hence the uniform ellipticity. □

We shall use these coordinates to construct local in time solutions to (5.1). Suppose that $u \in C^2$ satisfies (5.12). Then (5.7) defines a solution v to problem (5.4).

5.4.3 Global coordinates II

The coordinates introduced above are suitable for local existence. Here we want to introduce coordinates which are suitable for studying the large time behavior. In particular Barenblatt solutions centered at $x = 0$ starting at $t = 0$ correspond to constants in the coordinates we introduce here. This will allow the conclusion that every nondegenerate solution approaches the Barenblatt solution polynomially in time in a very precise sense. We assume that all functions below are sufficiently regular for the manipulations.

We define $w = e^{\frac{tn}{n(m-1)+2}}\rho(e^t, e^{\frac{t}{n(m-1)+2}}x)$. It satisfies

$$w_t - \Delta w^m - \frac{n}{n(m-1)+2}w - \frac{1}{n(m-1)+2}x \cdot \nabla w = 0.$$

Now let $v = \frac{m}{m-1}w^{m-1}$. It solves

$$v_t - (m-1)v\Delta v - |\nabla v|^2 - \frac{n(m-1)}{n(m-1)+2}v - \frac{1}{n(m-1)+2}x \cdot \nabla v = 0.$$

There are explicit stationary solutions: the Barenblatt solutions

$$\left(\lambda - \frac{1}{2(n(m-1)+2)}|x|^2\right)_+, \qquad \text{with } \lambda > 0.$$

We want to simplify the notation and set $y = \frac{x}{\sqrt{n(m-1)+2}}$ and $s = \frac{t}{n(m-1)+2}$ which leads to

$$v_s - (m-1)v\Delta_y v - |\nabla_y v|^2 - n(m-1)v - y \cdot \nabla_y v = 0. \tag{5.13}$$

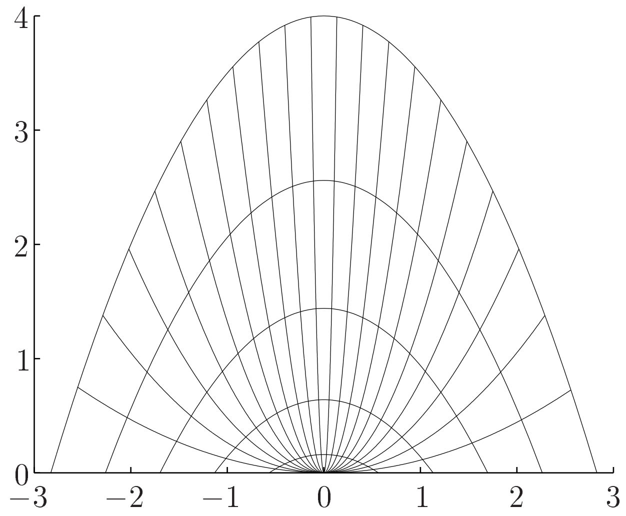

Figure 5.5: Global coordinates II

We denote again s by t and y by x. The Barenblatt solutions above are transformed to $w_\lambda(x) = (\lambda - |x|^2/2)_+$.

We define the map

$$\Phi : B(0, \sqrt{2}) \times \mathbb{R}^+ \ni (y, u) \to (yu, u^2(1 - |y|^2/2)_+).$$

This map is a diffeomorphism. It is visualized in Figure 5.5. The surfaces $u = const$ correspond to Barenblatt solutions. If the graph of u is the same as the graph of v (with different coordinates) and if

$$F^i(y, u, \nabla_y u) = \partial_{y_i} u - \frac{y_i |\nabla_y u|^2}{u + y_k \partial_k u}$$

$$\begin{aligned} u_t - (m-1) \sum_i \partial_i \Big[\Big(1 - \frac{|y|^2}{2}\Big) F^i \Big] - (m-2) y_i F^i \\ = \Big(1 - \frac{|y|^2}{2}\Big) |\nabla_y u|^2 \Big(\frac{2}{w + y_k \partial_{y_k} u} - (n-1)(m-1) \Big). \end{aligned} \tag{5.14}$$

This is the main result of this section. We shall i) show that this equation holds provided the graph can be described in that form and ii) show that this representation holds for large t. We begin with the calculation, which is similar to the one above.

$$\nabla_{y,u}\Phi = \begin{pmatrix} u1_{\mathbb{R}^n} & y \\ -u^2y^t & 2u(1-|y|^2/2) \end{pmatrix}$$

with determinant $D(y,u) = 2u^{n+1}$. Let $B_{ij} = u\delta_{ij} + y_i\partial_j u$. Then

$$|B| = u^n|\delta_{ij} + u^{-1}y_i\partial_j u| = u^{n-1}(u + y_k\partial_y w)$$

and

$$B_{ij}^{-1} = u^{-1}\left(\delta_{ij} - \frac{y_i\partial_j u}{u + y_k\partial_k u}\right)$$

because

$$\sum_j (u\delta_{ij} + y_j\partial_i u)\left(\delta_{jl} - \frac{y_l\partial_j u}{u + y_k\partial_k w}\right) = u\delta_{il}.$$

The time derivative is given by

$$v_t = \frac{2u^2}{u + y_k\partial_k u}u_t. \tag{5.15}$$

We compute

$$B^{-1}B^{-t} = u^{-2}\left(\delta_{ij} - \frac{y_j\partial_i u + y_i\partial_j u}{u + y_k\partial_k u} + \frac{|\nabla u|^2 y_i y_j}{(u + y_k\partial_k u)^2}\right),$$

$$\begin{aligned}
\partial_{x_i} v &= u^{-1}\sum_j \frac{\partial y_i}{\partial x_j}\partial_{y_i}\Phi_{n+1} \\
&= \sum_j \left(\delta_{ij} - \frac{y_j\partial_i u}{u + y_k\partial_k u}\right)\left(2(1-|y|^2/2)\partial_j u - uy_j\right) \\
&= \frac{2u\partial_i u}{u + y_k\partial_k u} - uy_i,
\end{aligned}$$

hence $x_k\partial_{x_k} v = u^2\left[\frac{2y_k\partial_k u}{u+y_k\partial_k u} - |y|^2\right]$ and

$$|\nabla_x v|^2 = u^2\left[\frac{4|\nabla u|^2}{(u + y_k\partial_k u)^2} - \frac{4y_k\partial_k u}{u + y_k\partial_k u} + |y|^2\right],$$

thus

$$\begin{aligned}
\frac{|B|}{D}\left(|\nabla_x v|^2 + x_k\partial_k v\right) &= \frac{|B|u^2}{D}\left[\frac{4|\nabla u|^2}{(u+y_k\partial_k u)^2} - \frac{2y_k\partial_k u}{u+y_k\partial_k u}\right] \\
&= \frac{2|\nabla u|^2}{u+y_k\partial_k u} - y_k\partial_k u.
\end{aligned} \tag{5.16}$$

We compute the Laplacian:

$$\begin{aligned}
|B|\Delta_x v &= \partial_{y_i}\left(\left(|B|B^{-1}B^{-t}\right)^{ij}\partial_{y_j}\left(u^2\left(1-\frac{|y|^2}{2}\right)\right)\right) \\
&= \partial_i\left[u^{n-3}\left(\delta_{ij}(u+y_k\partial_k u) - y_j\partial_i u - y_i\partial_j u + \frac{|\nabla u|^2 y_i y_j}{u+y_k\partial_k u}\right)\right. \\
&\qquad\qquad \left.\times\partial_j\left(u^2\left(1-\frac{|y|^2}{2}\right)\right)\right] \\
&= 2\partial_i\left[u^{n-2}\left(1-\frac{|y|^2}{2}\right)\left(y\partial_i y - y_i|\nabla y|^2 + \frac{|\nabla y|^2 y_i y_k\partial_k u}{u+y_k\partial_k u}\right)\right] \\
&\quad -\partial_i\left[u^{n-1}\left(uy_i - |y|^2\partial_i u + \frac{|\nabla u|^2|y|^2 y_i}{u+y_k\partial_k u}\right)\right] \\
&= 2\partial_i\left[u^{n-1}\partial_i u - u^{n-2}|\nabla u|^2 y_i\left(1-\frac{y_k\partial_k u}{u+y_k\partial_k u}\right) - \frac{u^n y_i}{2}\right] \\
&= 2\partial_i\left[u^{n-1}\left(\partial_i u - y_i\left(\frac{|\nabla u|^2}{u+y_k\partial_k u} + \frac{u}{2}\right)\right)\right]
\end{aligned}$$

hence

$$\begin{aligned}
\frac{v|B|}{D}(\Delta_x v + n) &= u^{1-n}(1-|y|^2/2)\partial_i\left[u^{n-1}\left(\partial_i u - \frac{y_i|\nabla u|^2}{u+y_k\partial_k u}\right)\right] \\
&= \partial_i\left[(1-|y|^2/2)\left(\partial_i u - \frac{y_i|\nabla u|^2}{u+y_k\partial_k u}\right)\right] \\
&\quad -(n-1)(1-|y|^2/2)u^{-1}\left(|\nabla u|^2 - \frac{|\nabla u|^2 y_k\partial_k u}{u+y_k\partial_k u}\right) \\
&\quad +y_k\partial_k u - \frac{|y|^2|\nabla u|^2}{u+y_k\partial_k u} \\
&= \partial_i\left[\left(1-\frac{|y|^2}{2}\right)\left(\partial_i u - \frac{y_i|\nabla u|^2}{u+y_k\partial_k u}\right)\right] \\
&\quad -(n-1)\left(1-\frac{|y|^2}{2}\right)|\nabla u|^2 + y_k\partial_k u - \frac{|y|^2|\nabla u|^2}{u+y_k\partial_k u}.
\end{aligned} \tag{5.17}$$

We combine (5.15), (5.16) and (5.17) to

$$
\begin{aligned}
u_t &= (m-1)\partial_i\left[\left(1-\frac{|y|^2}{2}\right)\left(\partial_i u - \frac{y_i|\nabla u|^2}{u+y_k\partial_k u}\right)\right] \\
&\quad -(n-1)(m-1)(1-|y|^2/2)|\nabla u|^2 + (m-1)\left(y_k\partial_k u - \frac{|y|^2|\nabla u|^2}{u+y_k\partial_k u}\right) \\
&\quad +\frac{2|\nabla u|^2}{u+y_k\partial_k u} - y_k\partial_k u \\
&= (m-1)\partial_i\left[\left(1-\frac{|y|^2}{2}\right)\left(\partial_i u - \frac{y_i|\nabla u|^2}{u+y_k\partial_k u}\right)\right] \\
&\quad +(m-2)\left(y_k\partial_k u - \frac{|y|^2|\nabla u|^2}{u+y_k\partial_k u}\right) \\
&\quad +\left(1-\frac{|y|^2}{2}\right)|\nabla u|^2\left(\frac{2}{u+y_k\partial_k u} - (n-1)(m-1)\right).
\end{aligned}
\tag{5.18}
$$

We conclude this section with an examination when the graph of a solution can be parametrized in this way. This is true for equation (5.13) whenever $< (y, 2u(1-|y|^2/2), (-\nabla v, 1) >> \delta > 0$ if $x = uy$, $v = u^2(1-|y|^2/2)$ or $u = \sqrt{v+|x|^2/2}$, $y = x/\sqrt{v+|x|^2}$ hence this inequality becomes

$$2v - x\cdot\nabla v > \delta\sqrt{2(v+|x|^2/2)}.$$

This is satisfied where v decays in radial direction. For $x=0$ it reduces to $v>0$. Proposition 5.3.15 and Theorem 5.3.19 imply that these inequalities are satisfied for large t and that $|v|$ and $|\nabla v|$ are uniformly bounded. The crucial quantity

$$u + y_k\partial_k u = (v+|x|^2/2)^{-3/2}(2v - x\cdot\nabla v)$$

is bounded from below and above by the same arguments. The coefficients of the second order terms are

$$a_{ij} = \delta_{ij} - 2\frac{y_i\partial_j u}{u+y_k\partial_k u} + \frac{y_i y_k|\nabla u|^2}{(u+y_k\partial_k u)^2}.$$

They satisfy

$$\left(\frac{u}{u+x_k\partial_k u}\right)^2|\xi|^2 \le a^{ij}\xi_i\xi_j \le \left(1+\frac{\sqrt{2}|\nabla u|}{u}\right)^2.$$

5.5 Local existence of classical solutions

In this section we prove Theorem 5.2.8. We use the coordinates of Subsection 5.4.2 and the notation from there and from Subsection 4.7.2. In particular the Hölder spaces are understood with respect to the Carnot-Carathéodory metric. The assertion of Theorem 5.2.8 follows from well-posedness (with $\sigma > -1$) for (see equation (5.12))

$$u_s - \psi^{-\sigma} \sum_{i=1}^n \partial_i (\psi^{1+\sigma} F^i(y, u, \nabla u)) - \psi f(y, u, \nabla u) = 0 \tag{5.19}$$

in $\Omega \times [0, t_0)$ with initial value u_0 which satisfies $\nabla u_0 \in C^\alpha(\Omega)$.

Theorem 5.5.1. *Suppose that $\nabla u_0 \in C^\alpha(\Omega)$ and that the assertions of Proposition 5.4.1 hold. Then there exist $t_0 > 0$, a unique weak local solution u with $\nabla u \in C^\alpha(\Omega \times [0, t^0])$. Moreover the maps $t \to u(t), \nabla u(t, .) \in C^\alpha(\Omega)$ are analytic in $0 < t \le t_0$.*

Remark 5.5.2. It follows that u restricted to the boundary is in $C^{1,\alpha}(\Gamma_T)$. In particular the speed of the boundary is bounded from below and from above for small times in terms of the norms of the initial data. Theorem 5.6.1 below strengthens the result: u is actually C^∞ for $0 < t < \delta$. The same is then true for v in its support. This gives the C^∞ part of Theorem 4.

Proof. Every classical solution to (5.19) defines a solution to the porous medium equation (5.1). The latter is unique. Hence solutions to (5.19) are unique.

We construct a solution by the Banach fixed point principle. Equation (5.19) can be rewritten as

$$u_s - \psi^{-\sigma} \partial_i (\psi^{1+\sigma} a_0^{ij} \partial_j u) = \psi^{-\sigma} \partial_i \left(\psi^{1+\sigma} G^i(y, u, \nabla u) \right) + \psi f(y, u, \nabla u)$$

with $a_0^{ij}(y) = \frac{dF^i}{dp_j}\Big|_{(y, u_0(y), \nabla u_0(y))}$ and

$$G^i(y, u, \nabla u) = F^i(y, u, \nabla u) - a_0^{ij} \partial_j u.$$

Let $X_T = \{u \in C^\alpha(\Omega \times [0, T]) : \nabla u \in C^\alpha(\Omega \times [0, T]), u(., 0) = u_0\}$ with norm

$$\max\{\|u\|_{C^\alpha(\Omega_T)}, \|\nabla u\|_{C^\alpha(\Omega_T)}\}.$$

We define the map $X_T \to X_T$, $\tilde{u} \to u$ where u is the solution to

$$\begin{aligned} u_s - \psi^{-\sigma} \partial_i (\psi^{1+\sigma} a_0^{ij} \partial_j u) &= \psi^{-\sigma} \partial_i (\psi^{1+\sigma} G^i(y, \tilde{u}, \nabla \tilde{u})) + \psi f(y, \tilde{u}, \nabla \tilde{u}) \\ u(., 0) &= u_0 \end{aligned}$$

and obtain by Theorem 4.7.5 that there exists a unique solution which satisfies

$$\|u\|_{X_T} \le c_1 \left(\|G\|_{C^\alpha(\Omega\times[0,T])} + \|f\|_{L^\infty(\Omega\times[0,T])} + \|u_0\|_\infty + \|\nabla u_0\|_{C^\alpha(\Omega)} \right),$$

and, if u^i is the solution with right hand side $G^i = G(\tilde u^i)$, $f^i = f(\tilde u^i)$, $i = 1, 2$,

$$\|u^2 - u^1\|_{X_T} \le c_1 \left(\|G^2 - G^1\|_{C^\alpha(\Omega_T)} + \|f^2 - f^1\|_{L^\infty(\Omega_T)} \right).$$

Since $|G^i(y, u, q) - F^i(y, u, q)| \le c_2|q|^2$ we obtain by a simple computation

$$\|G\|_{C^\alpha(\Omega_T)} + \|f\|_{L^\infty(\Omega_T)} \le c_3(\|\nabla u_0\|_{C^\alpha(\Omega_T)}, \|\nabla \tilde u\|_{L^\infty(\Omega_T)}) \left(1 + T^{\alpha/2}\|\tilde u\|_{X_T}\right)$$

and, with $|.|_\alpha = |.|_{C^\alpha(\Omega_T)}$ and $\|.\|_\infty = \|.\|_{L^\infty(\Omega_T)}$,

$$\|G^2 - G^1\|_\alpha + \|f^2 - f^1\|_\infty \le c_4(\|\nabla \tilde u^1\|_\alpha + \|\nabla \tilde u^2\|_\alpha) T^{\alpha/2} \|\nabla(\tilde u^2 - \tilde u^1)\|_\alpha.$$

We choose $R = 2c_1c_3(\|\nabla u_0\|_{C^\alpha}, 2\|u_0\|_{C^1})$. If $T \le R^{-2/\alpha}$ and $\|\tilde u\|_{X_T} \le R$ then $\|u\|_{X_T} \le R$ since $\|\nabla \tilde u(t)\|_\infty \le \|\nabla u_0\|_\infty + t^{\alpha/2}R$. Hence the map $\tilde u \to u$ maps a ball of radius R to itself. By the same argument it is a contraction (after decreasing T if necessary). There exists a unique fixed point by the contraction mapping principle, which is the unique solution. The construction however gives much more than that.

It is clear that

$$X_T \ni u \to (F(y, u, \nabla u), f(y, u, \nabla u)) \in C^\alpha(\Omega_T)$$

is differentiable. Moreover F and f are both analytic. Hence they can be extended (for fixed y) into a neighborhood of $\mathbb{R}^{n+1} \subset \mathbb{C}^{n+1}$.

We introduce the complex parameter z and consider

$$u_s - (1+z)\psi^{-\sigma} \sum_{i=1}^n \partial_i \left[\psi^{1+\sigma} F^i(y, u, \nabla u)\right] - (1+z)\psi f(y, u, \nabla u) = 0$$

with initial values u_0 and complex valued solutions. The map constructed above is a contraction near $z = 0$. Hence the derivative is invertible. The implicit function theorem gives existence of a solution for small $z \in \mathbb{C}$, or, equivalently, we obtain a solution for t in the intersection of a complex sector $|\arg t| \le \delta$, $|t| \le \delta$ for some nonnegative δ. The solution is holomorphic in t. Hence it depends analytically on t.

Lipschitz dependence, and more generally smooth dependence on the initial data requires a simple check. □

Remark 5.5.3. We could derive estimates for higher time derivatives by applying the integral formula of Cauchy (with values in $C^\alpha(\Omega)$).

5.6 Regularity

Let $\sigma > -1$, $V \subset \mathbb{R}^n$ open and $F \in C^\infty(V)$. We assume that there exists $\kappa > 0$ such that

$$\sum_{i=1}^{n-1} \xi_i^2 + \left.\frac{\partial F}{\partial p_j}\right|_p \xi_j \xi_n \geq \kappa |\xi|^2 \quad \text{for } \xi \in \mathbb{R}^n \text{ and } p \in V.$$

We study weak solutions to the local problem (with $\sigma > -1$)

$$u_s - y_n \Delta_{\mathbb{R}^{n-1}} u - y_n^{-\sigma} \partial_n (y_n^{1+\sigma} F(\nabla u)) = 0 \tag{5.20}$$

in the set $U \subset \mathbb{R}^{n-1} \times [0, \infty) \times \mathbb{R}$ which we assume to be open in the relative topology of the half space. Our main result is

Theorem 5.6.1. *Suppose that u is Lipschitz continuous in U and $\nabla u \in V$. Then $u \in C^\infty(U)$.*

Remark 5.6.2. This implies that the solutions of Theorem 5.5.1 are smooth if $t > 0$. Here we have to change coordinates twice: once from global coordinates to x and v, and then to the local coordinates of Section 5.4.1.

It is even more significant that the assumptions are locally satisfied if

1. we start with a solution to the porous medium equation with α admissible initial data.

2. if we consider a point outside the ball $B(0, R)$

3. if we introduce the local coordinates of Section 5.4.1.

As consequence the pressure v is C^∞ in its support outside the ball of radius R. Moreover the function S is C^∞ at all x with $|x| > R$ in this situation.

Proof. Let v be a difference quotient in a direction tangential to the boundary. It satisfies an equation

$$v_s - v_n \Delta_{\mathbb{R}^{n-1}} v - y_n^{-\sigma} \partial_n (y_n^{1+\sigma} a^j \partial_j v) = 0$$

in the energy sense. Hence, by Theorem 4.5.5 v is uniformly Hölder continuous for some Hölder exponent α.

The vertical derivative $v = \partial_n u$ satisfies

$$v_s - y_n \Delta_{\mathbb{R}^{n-1}} v - y_n^{-\sigma} \partial_{y_n} (y_n^{1+\sigma} a^j \partial_j v) = \Delta_{\mathbb{R}^{n-1}} \partial_j v \tag{5.21}$$

and hence, by Theorem 4.5.6, if Q is a cube with closure contained in U

$$\|\partial_n v\|_{C^\alpha(Q)} \leq c\left(\|\partial_n v\|_{L^1(x_n^\sigma)(U)} + \|\nabla_{\mathbb{R}^{n-1}} v\|_{C^\alpha(U)}\right)$$

provided we can show that $\nabla_y \partial_n v \in L^2_{loc}(U, y_n^{\sigma+2})$.

To see this we assume that $(0,0) \in U$ and prove the assertion in a neighborhood of 0. Let R be sufficiently small, $\tilde{\psi} \in C_0^\infty$ and $\psi_0 \in C_0^\infty(\mathbb{R}^+)$ with $\sum \psi_0(2^i \tau) = 1$ for all $\tau > 0$. We define $\psi_i(y) = \psi_0(2^i|y_n|)\tilde{\psi}((|y|^2 + |t|^2)/R^2)$. Then, if $2^{-i} \leq R$, using energy estimates

$$\|\psi_i \nabla v\|_{L^2(y_n^{\sigma+2})} \leq c(2^{i/2} + R^{-1})2^{-(\sigma+1+1/2)/i},$$

which is summable with respect to i.

This implies that the coefficients of the linearization are Hölder continuous. We claim that $\partial^\alpha u \in L^p_{loc}(U)$ for all α and $p > \max\{1, \frac{1}{1+\sigma}\}$. It clearly suffices to do this for spatial derivatives. Time derivatives can be iteratively expressed in terms of spatial derivatives.

We prove the claim by double induction on $|\alpha|$ and on α_n. Suppose that the statement is true for $|\alpha| \leq N$ and $\alpha_n \leq M-1$. Let α be a multi-index with $|\alpha| = N$ and $0 < \alpha_n = M \leq N$.

Let $\beta = \alpha^{(n)}$, the multi-index which has n^{th} component lowered by 1. Then, with $w = \partial^\beta v$,

$$\begin{aligned} w_s - y_n \Delta_{\mathbb{R}^{n-1}} w - \partial_n(y_n a^j \partial_j w) - (\sigma + M - 1)a^{nj}\partial_j w& \\ = (M-1)\Delta_{\mathbb{R}^{n-1}} \partial^{\beta^{(n)}} v + l.o.t.& \end{aligned}$$

where the lower order terms are of the form $g + \sum \partial_i(y_n f^i)$ with $g, f^i \in L^p$ by the induction hypothesis and $\beta^{(n)}$ is zero if $\beta_n = 0$ and has n-component lowered by 1 otherwise.

Hence by Theorem 4.7.2

$$\|\partial_n w\|_{L^p(Q(x,t,R))} \leq c\|\Delta_{\mathbb{R}^{n-1}} \partial^{\beta^{(n)}} v\|_{L^p(Q(x,t,2R))} + l.o.t.$$

The right hand side is bounded by the induction hypothesis.

There is a second case to consider, $M = N$. In that case we have to prove the estimate for $\alpha_n \leq 1$ and $|\alpha| = N+1$. Let $\beta = \alpha^{(n)}$ if $\alpha_n = 1$ and $\beta = \alpha^{(1)}$ without loss of generality otherwise. The equation of $w = \partial^\beta v$ has the same form as above, but with the term $\Delta_{\mathbb{R}^{n-1}}$ on the right hand side.

This argument is formal: we have to justify the arguments since we do not know apriori that the solution is sufficiently regular for this argument. In the case $\beta_n = 0$ this is done by finite differences in the same way as above in one direction. If $\beta_n > 0$ we have to check that the highest order vertical derivative exists and lies in $L^2(x_n^{\sigma+\alpha_n})$. This is done in the same way as above. The proof is complete. $\square$

5.7 Notes

In the one-dimensional case existence and uniqueness have been proven by Oleinik, Kalashnikov and Yui-Lin [110]. The regularity has been studied by Aronson [8] and Aronson and Caffarelli [12], who showed that the pressure is Lipschitz continuous under suitable assumptions as indicated by the Barenblatt solutions.

Non-degeneracy is crucial for regularity: without non-degeneracy there are solutions with waiting time. This has been shown in higher space dimensions by Vespri [137]. In one space dimension the waiting time, the behavior at the transition time and the regularity after the waiting have been studied by Aronson [15], Aronson, Caffarelli and Vazquez [16], Aronson and Vazquez [14] and Vazquez [135]. It is even analytic when it is moving (Angenent [2], [1]). It has been shown by Angenent [3] that the boundary determines the solution, again either in one space dimensions or for radially symmetric solutions. The corresponding results in higher dimensions follows from this chapter and Proposition 6.2.1 and Proposition 6.3.1 below.

In higher space dimensions all these issues are more complicated. Hölder continuity (Theorem 5.3.5) and Hölder continuity away of the initial support of $S(x) := \inf\{t : \rho(x,t) > 0\}$ have been shown by Caffarelli and Friedman [27], [28]. Lipschitz continuity of S away from the initial support follows from Proposition 5.3.18 under some assumptions on the initial data. This has been shown by Caffarelli, Vazquez and Wolanski [30]. One may hope that $S(x)$ is Lipschitz continuous away from the initial support provided the initial data are compactly supported. This question seems to be open.

If S is Lipschitz continuous and $S(x)$ is large it has been shown by Caffarelli and Wolanski [31] that $S(x)$ is of class $C^{1,\alpha}$ near x. This has been improved in this chapter: S is then C^∞.

The Graveleau solutions (see Aronson and Graveleau [13]) and similar solutions (see Aronson and Angenent, [4]) show that the pressure is in general not Lipschitz continuous. The optimal Hölder exponent of the density is unknown.

The transformation of coordinates goes back in the simplest case to von Mises. The more general transformation used here appears in Meirmanov, Pukhnachov, Shmarev [96], who proved local existence of classical solutions under suitable assumptions on the initial data. This has been improved by Daskalopoulos and Hamilton [47], [48]: they require less regularity of the initial data (but more than is known from [30] and [31] for weak solutions) and construct local smooth solutions. Their main tools are Schauder estimates for the linear problem.

The global coordinates of Section 5.4.3 can be used to study the asymptotic behavior. This reduces to a study of eigenvalues and eigenfunctions for the linearized operator. The rotationally symmetric case has been studied by Angenent and Aronson [5].

There are excellent surveys on the porous medium equation. Two of them are Aronson [9] and Vazquez [136].

The porous medium equation should be seen in the context of other free boundary problems. The simplest ones are the obstacle problems. Here non-degeneracy is well-understood. Lipschitz free boundaries are non-degenerate and the whole problem reduces to nonlinear elliptic or parabolic problems. The situation is similar (but with additional twists) for the one phase Stefan problem, the quasistationary one phase Stefan problem, the Hele-Shaw problem and certain filtration problems which can be related to obstacle problems via a Baiochi transform. For the two-phase Stefan problem C^1 regularity has been studied by Athanasopoulos, Caffarelli and Salsa [17], [18] and [19]. Classical smooth solutions to the one-phase Stefan problem have a long history: see Hanzawa [68], Meirmanov [94] and, even earlier Kinderlehrer and Nirenberg [81], [80]. There is a strong relation to the two phase Stefan problem. The linearized problem has a different character here. In Koch [83] the gap between known regularity of some weak solutions ([17], [18], [19]) and smooth classical solutions is bridged. Global regularity (in fact analyticity) for large time for the one phase Stefan problem with compactly supported initial data has been obtained by Matano [93].

It is clear that the model problem (5.1) is not sufficient for more complex physical problems. A survey of this is beyond the scope of this work. It is however likely that the insights gained for (5.1) are useful for these more complex problems.

Chapter 6

Analyticity

There are two main results in this section: if the solution to the porous medium equation is locally non-degenerate, if its pressure is locally Lipschitz continuous and if the boundary of the support is locally Lipschitz then the pressure is locally analytic in the space variables up to the boundary of its support, and of Gevrey type 2 in time. If the pressure $p(t_0) \in C^{1,\alpha}(\text{ supp } \rho_0)$ and $|\nabla p(.,t_0)| + p(.,t_0) \geq \delta$ for all points x in its support then p is analytic in all variables for $t_0 < t < t_1$ for some $t_1 > t_0$. Conversely: if we have a candidate for the interface of that regularity then there exists a unique solution in a neighborhood of that surface. This is a Cauchy-Kowalevski-type theorem.

There should be a third related result: the interface should determine the solution even when the constitutive functions are not analytic. However I was not able to obtain neither the corresponding Carleman estimates nor the doubling property. This would be useful for inverse problems because Carleman estimates would imply logarithmic convexity estimates which quantify the dependence of the solution on the boundary of the support.

We shall work in Banach spaces of analytic functions, with the Banach fixed point theorem as main tool for constructing solutions. There is a prize to pay compared to approaches using hard implicit function theorems: we have to be much more careful in the choice of the norms. The advantage is that the structure of the arguments is fairly simple.

Spaces of analytic functions or Gevrey regular functions are usually characterized as follows: a function f is analytic in K if there exists R such that

$$\sup_{x \in K} \sup_{\alpha} \alpha!^{-1} R^{|\alpha|} |\partial^\alpha f| < \infty.$$

In particular the Taylor expansion converges in every poly-disc $\{x \in \mathbb{R}^n : \sup_i |x_i - x_i^0| < r\}$ if $r < R$ and $x^0 \in K$. It turns out to be simpler to work with norms of the type

$$\|u\|_R = \sum_\alpha \alpha!^{-1} R^{|\alpha|} |\partial^\alpha f|.$$

Finiteness of this norm is equivalent to convergence of the power series in $C(Q_R)$ where Q_R denotes the (complex or real) polycylinder $|z_i| \leq R$.

The main advantage of these semi-norms lies in natural estimates of composition operators: the simplest such estimate is certainly

$$\|f \circ u\|_{L^\infty(K)} \leq \|f\|_{L^\infty(B(0,\|u\|_{L^\infty(K)}))}.$$

A formula of similar simplicity holds for analytic functions:

$$\|f \circ u\|_R \leq \|f\|_{\|u\|_R}$$

where the norm of f at the right hand side is evaluated only at $x = 0$. This, together with modifications, will be the first object of study after preliminary considerations and basic definitions.

The application to the porous medium equation is natural and surprising at the same time. It is natural, since the flow of arguments is straight forward once the weights and the approach have been properly chosen. It is surprising because we obtain a different equation for every vertical derivative. One might expect that at least vertical derivatives grow as fast as time derivatives. Since it is clear from the beginning that one cannot hope for more than Gevrey-2 in time by local considerations this would imply non-analyticity in the vertical direction. This heuristic consideration fails.

The mechanism leading to analyticity relies on the particular structure of the problem. There are two requirements on the norms we shall use: there have to be good estimates for solutions to the linear equation. Moreover we need good estimates for nonlinear functions of unknowns. These two points cause the need for good estimates for the Gamma function. We first prove the simpler Cauchy-Kowalevski-type theorem before we turn to the question of the regularity of solutions to the pressure equation. The results needed from other sections are: 1. The coordinate systems of Section 5.4. 2. Regularity (Theorem 5.6.1) and analyticity in time (Proposition 5.5.1). 3. The estimate of singular integral operators in Theorem 4.7.2.

6.1 Preliminaries

Multi-indices

A multi-index α is an ordered (n+1) tuple of nonnegative integers $(\alpha_0, \dots \alpha_n)$. Its length is $|\alpha| = \sum_{i=0}^n \alpha_i$. We define $\alpha! = \prod_{i=0}^n \alpha_i!$. Multi-indices are a semi-subgroup of $\mathbb{Z}^{n+1}$ from which they heir their additive structure. We denote multi-indices of length 1 by e_i. If α is any multi-index we define $\alpha^{(j)}$ as zero if $\alpha_j = 0$ and as $(\alpha_0, \dots \alpha_j - 1, \dots)$ otherwise. Similarly $\alpha_{(j)} = \alpha + e_j$. The notation ∂^α is used in the standard way.

The Gamma function and powers

We define the Gamma function by Euler's identity:

$$\Gamma(s) = \int_0^\infty t^{s-1} e^{-t}\, dt \text{ for } s > 0.$$

Induction by n and integration by parts shows that $\Gamma(n+1) = n!$. It satisfies the formula of Stirling (see Remmert [115], p. 52)

$$\Gamma(s) = \sqrt{2\pi} s^{s-1/2} e^{-s} e^{\mu(s)} \tag{6.1}$$

with $|\mu(s)| \leq \frac{1}{8s}$ for $s > 0$. Moreover $\ln \Gamma(s)$ is convex for $s > 0$ (this follows from formula $1'$, p. 42 of Remmert [115]). Let α be a multi-index. Then

$$|\alpha|^{|\alpha|} \leq \left(\frac{e^{1/8} n}{2\pi|\alpha|}\right)^{n/2} (en)^{|\alpha|} \alpha! \tag{6.2}$$

This can be seen as follows: by the convexity of the logarithm of the Gamma function

$$\alpha! = \prod \alpha_i! = \prod \Gamma(\alpha_i + 1) \geq \Gamma((|\alpha|/n) + 1)^n$$

and, with $z = |\alpha|/n$, $|\alpha|^{|\alpha|} = n^{|\alpha|}(z^z)^n$. Stirling's formula (6.1) now gives for all α

$$|\alpha|^{|\alpha|} \leq \left(\frac{e^{1/4}|\alpha|}{2\pi n}\right)^{n/2} (en)^{|\alpha|} \Gamma(z)^n \leq \left(\frac{e^{1/8} n}{2\pi|\alpha|}\right)^{n/2} (en)^{|\alpha|} \alpha!.$$

There is an immediate well-known consequence: if $|\partial^\alpha f| \leq c_0 R^{-|\alpha|} |\alpha|^{|\alpha|}$ then f is analytic – more precisely it satisfies

$$|\partial^\alpha f| \leq c_0 \left(\frac{e^{1/8} n}{2\pi}\right)^{n/2} (R/ne)^{-|\alpha|} \alpha!$$

and hence the radius of convergence is at least R/ne. This argument could be partly reversed in the sense that a bounded analytic function which is given by normally converging power series on the disc of radius R satisfies

$$\sum_\alpha |\alpha|^{-|\alpha|} R^{|\alpha|} |\partial^\alpha f| \leq \sum_\alpha \alpha!^{-1} R^{|\alpha|} |\partial^\alpha f| < \infty.$$

It will be useful to have a sample of super-multiplicative functions on multi-indices.

Lemma 6.1.1. *The functions* $l(\alpha) = \alpha!$, $l(\alpha) = \alpha_i!$, $(0 \le i \le n)$, $l(\alpha) = |\alpha|^{|\alpha|}$ *and* $l(\alpha) = \frac{|\alpha|^{|\alpha|}}{\alpha!}$ *satisfy* $l(\alpha+\beta) \ge l(\alpha)l(\beta)$.

Proof. The first assertion for $\alpha!$ follows from the assertion for $l(\alpha) = \alpha_i!$. This is equivalent to the fact that $\binom{m}{n} \ge 1$ if $m \ge n$. The claim for $l(\alpha) = |\alpha|^{|\alpha|}$ follows from $(x+y)\ln(x+y) \ge x\ln x + y\ln y$ for $x, y \ge 1$, which is a consequence of the monotonicity of ln.

The assertion for $l(\alpha) = \frac{|\alpha|^{|\alpha|}}{\alpha!}$ is the most difficult one. We observe that, given $|\alpha|$ and $|\beta|$, $\prod \binom{\alpha_i+\beta_i}{\alpha_i}$ is largest for $\alpha = |\alpha| e_0$ and $\beta = |\beta| e_0$ (by induction it suffices to consider two indices. The number of factors of nominator and denominator depends only on $|\alpha|$ and $|\beta|$. The factors can be compared one by one in a suitable order). Thus we have reduced the problem to proving

$$\frac{(m+n)^{m+n}}{m^m n^n} \Big/ \frac{(n+m)!}{n!m!} \ge 1.$$

We may assume that $n > m$ since this is clearly correct if $n = m$ and symmetric in n and m. We take the logarithm. Since $(x\ln x - x)' = \ln(x)$ we have to verify

$$I = 1 + \left(\int_n^{n+m} \ln x \, dx - \sum_{i=n+1}^{m+n} \ln i \right) + \left(\sum_{i=1}^{m} \ln i - \int_1^m \ln x \, dx \right) \ge 0.$$

Now

$$\begin{aligned} J = \int_k^{k+1} \ln x \, dx &= k(\ln(k+1) - \ln(k)) + \ln(k+1) - 1 \\ &= \frac{\ln(k+1) + \ln k}{2} + \left(k + \frac{1}{2}\right) \int_k^{k+1} x^{-1} dx - 1 \end{aligned}$$

hence by a Taylor expansion at $k + 1/2$,

$$0 \le J - \frac{\ln(k+1) + \ln k}{2} = \sum_{j=1}^{\infty} \frac{(2k+1)^{-2j}}{2j+1} \le (12k(k+1))^{-1}.$$

Thus $I \ge 1 + (\ln m + \ln n - \ln(n+m))/2 - \frac{1}{12} \ge 1 - \frac{\ln 2}{2} - \frac{1}{12} \ge 0$. $\square$

Analytic and Gevrey regular functions

Definition 6.1.2. *We say that* u *is in the Gevrey class* a, $a \ge 1$ *in the open set* U, *if for all compact subsets* $K \subset U$ *there exists* R *such that*

$$\sum_\alpha |\alpha|^{-a|\alpha|} R^\alpha |\partial^\alpha u| < \infty.$$

By the previous subsection f is of Gevrey class 1 if and only if f is analytic. It is not difficult to see that $f(t) = 0$ if $t \leq 0$ and $f(t) = e^{-\frac{1}{t^s}}$ otherwise is in Gevrey class $\frac{s}{s-1}$. We shall see that the product of Gevrey regular functions is in the same Gevrey class. Hence there are nontrivial compactly supported functions in the Gevrey class a if and only if $a > 1$.

In the sequel we will work with a particular space G_2 which we define by the norms

$$\|u\| = \sup_{x\in K} \sum R^{|\alpha|} |\alpha|^{-|\alpha|-|\alpha_0|} |\partial^\alpha_{x,t} f| \tag{6.3}$$

where K is a compact set. For functions u with values in $\mathbb{R}^n$ we denote by $|.|$ the maximum of the size of the components. The space G_2 consists of functions which are analytic in x and of class Gevrey class 2 with respect to t.

Composition Operators

Lemma 6.1.3. *Suppose that f is an analytic function in a polycylinder around zero of radius R_0. Then*

$$\|f(u)\|_R \leq \|f\|_{\|u\|_R}$$

if $\|u\|_R < R_0$. Here $K = \{0\}$ for the norm of f.

Proof. We have to estimate $\sum_\alpha R^{|\alpha|}\alpha!^{-1}|\partial^\alpha f(u)|$ which we compare to

$$\sum_\beta \beta!^{-1} \left(\sum_\alpha R^{|\alpha|}\alpha!^{-1}|\partial^\alpha u| \right)^{|\beta|} |\partial^\beta f|.$$

The summands are in one to one correspondence (after an expansion of the power $|\beta|$) and the summands of the second are the modulus of those of the first one. ☐

There are many possible modifications. First there is no need to sum over all terms. If we restrict ourselves to certain α on the left hand side (like $|\alpha| \leq N$ or $\alpha_i = 0$ for some i) than we may omit all terms on the right hand side which do not contribute to those terms. Secondly, if $l_0(\alpha) = \frac{l(\alpha)}{\alpha!}$ satisfies $l_0(\alpha+\beta) \geq l_0(\alpha)l_0(\beta)$ then, with

$$\|u\|_{R,l} = \sup_{x\in K} \sum l(\alpha)^{-1} R^{|\alpha|} |\partial^\alpha u| \tag{6.4}$$

the following proposition holds. Examples are given in Lemma 6.1.1.

Proposition 6.1.4. *The estimate*

$$\|f(u)\|_{R,l} \leq \|f\|_{\|u\|_{R,l}}$$

holds for all u such that the right hand side is finite.

Proof. Repeat the previous argument. The condition on l ensures that the inequality holds for corresponding summands. $\square$

There is a useful consequence: functions in Gevrey class a are an algebra even if one fixes R.

Normalizations

In this chapter we study analyticity of solutions to the porous medium equation. We assume that $p = \frac{m}{m-1}\rho^{m-1}$ is the pressure of a solution to the porous medium equation. (x_0, t_0) with $t_0 > 0$ is a point in the boundary of the support. We assume that $p \in C^\infty$ in the intersection of its support and a small ball around (x_0, t_0) and that p is non-degenerate near that point, i.e. $|\nabla p(x_0, t_0)| \neq 0$.

An affine change of coordinates allows us to assume that $x_0 = 0$, $t_0 = 0$ and $\nabla p(0,0) = \delta e_n$. A further parabolic scaling $t \to \lambda^2 t$, $x \to \lambda x$ reduces the problem to local solutions to

$$p_t = (m-1)p\Delta p - |\nabla p|^2$$

in a neighborhood of zero with $p(0,0) = 0$ and $\nabla p(0,0) = e_n$.

Changing dependent and independent variables as in Subsection 5.4.1 (and scaling the time derivative) we arrive at

$$u_t - x_n \Delta_{\mathbb{R}^{n-1}} u + x_n^{-\sigma} \partial_n \left(x_n^{1+\sigma} \frac{1 + \sum_{i=1}^{n-1} |\partial_i u|^2}{\partial_n u} \right) = 0$$

in $U \subset \mathbb{R}^{n-1} \times \mathbb{R}^+ \times \mathbb{R}$ with $\sigma = \frac{2-m}{m-1}$ and $u = x_n - \frac{1}{m-1} t + O(|x|^2 + t^2)$. The questions we study are the following:

1. What can be said about the regularity of u beyond smoothness?

2. Suppose $h : V \subset \mathbb{R}^{n-1} \times \mathbb{R}$ is given with $h_t(0,0) = -\frac{1}{m-1}$. Is there a local solution with $u(\hat{x}, 0, t) = h(\hat{x}, t)$ and $\nabla u(0,0) = e_n$?

The answers are contained in Proposition 6.2.1 and Proposition 6.3.1. They immediately imply Theorem 5 and Theorem 4 of the introduction.

We do a further reduction: Let $v = u - x_n + \frac{1}{m-1}t$. It satisfies in U

$$v_t - \sum_{i=1}^{n} x_n^{-\sigma}\partial_i(x_n^{1+\sigma}\partial_i v) = -x_n^{-\sigma}\partial_n\left(x_n^{1+\sigma}\frac{|\nabla v|^2}{1+\partial_n v}\right). \tag{6.5}$$

Let $g(\nabla v) := -\frac{|\nabla v|^2}{1+v_n}$. The function g is analytic. It satisfies $g(0) = \frac{\partial g}{\partial p_i}(0) = 0$ for $1 \le i \le n$. Moreover, with $K = \{0\}$,

$$\|g\|_R = G(R) = \frac{nR^2}{1-R} \le \frac{nR^2}{2} \quad \text{if} \quad R \le 1/2. \tag{6.6}$$

The regularity and the extension problem for v and that for p are equivalent.

The weights

We define the weighted spaces which we shall use and apply the previous results to them. We define the weight function

$$l(\alpha) = \frac{(|\alpha| + \alpha_n + \alpha_0)!}{R^{|\alpha|}\alpha_n!} \tag{6.7}$$

and

$$\|f\|_{R,l} = \sup_{(\hat{x},t)\in B} \sum_\alpha l(\alpha)^{-1}|\partial^\alpha_{x,t} f|.$$

Lemma 6.1.5. *The following estimate holds provided the right hand side is finite*

$$\|g(\nabla v)\|_{R,l} \le G(\|v\|_{R,l}) \le n\|\nabla v\|^2_{R,l}/2$$

where the last inequality holds if additionally $\|\nabla v\|_{R,l} \le 1/2$.

Proof. In view of Proposition 6.1.4 and equation (6.6) we have to show that $l(\alpha+\beta)/(\alpha+\beta)! \ge (l(\alpha)/\alpha!)(l(\beta)/\beta!)$, or, equivalently

$$\frac{(|\alpha+\beta| + \alpha_n + \beta_n + \alpha_0 + \beta_0)!}{(|\alpha| + \alpha_n + \alpha_0)!(|\beta| + \beta_n + \beta_0)!} \ge \frac{(\alpha+\beta)!}{\alpha!\beta!}\frac{(\alpha_n+\beta_n)!}{\alpha_n!\beta_n!}.$$

A stronger estimate is

$$\binom{|\alpha+\beta|}{|\alpha|} \ge \prod \binom{\alpha_i+\beta_i}{\alpha_i}$$

because $\binom{n+m+k}{n} \ge \binom{n+m}{n}$ allows to omit α_0 and β_0 on the left hand side. The estimate can be iteratively reduced to $\binom{n_1+n_2}{m_1+m_2} \ge \binom{n_1}{m_1}\binom{n_2}{m_2}$ for $0 \le m_i \le n_i$, which itself is a consequence of the functional equation $\sum_{i=0}^k \binom{n}{i}\binom{m}{k-i} = \binom{n+m}{k}$. □

6.2 The extension problem

Let

$$l(\alpha) = \frac{(2n)^{\alpha_n}(|\alpha| + \alpha_n + \alpha_0)!}{R^{|\alpha|}\alpha_n!}.$$

Proposition 6.2.1. *Let $B(0, R_0) \subset \mathbb{R}^{n-1} \times \mathbb{R}$ be an open ball and $f \in C^\infty(B)$ with $f(0) = \nabla f(0) = 0$ such that*

$$\sup_{(x,t)\in B_0} \sum_{\{\alpha:\alpha_n=0\}} l(\alpha)^{-1}|\partial^\alpha \nabla_{\hat{x},t} f| \leq \frac{5}{64n} \tag{6.8}$$

where $B_0 = B(0, R_0) \cap (\mathbb{R}^{n-1} \times \{0\} \times \mathbb{R})$. Then there exists a unique solution v to (6.5) in $\{(x, t) : x_n \leq R_0/(2n), (t, \hat{x}) \in B(x, R_0)\}$ with $v(., 0) = f$, analytic in x_n, with

$$\sup_{(x,t)\in B_0} \sum_\alpha l(\alpha)^{-1}|\partial^\alpha \partial_n v| \leq 4 \sup_{(x,t)\in B_0} \sum_{\{\alpha:\alpha_n=0\}} l(\alpha)^{-1}|\partial^\alpha \nabla_{(\hat{x},t)} f|.$$

Remark 6.2.2. The estimate of Theorem 6.2.1 implies that the formal Taylor series in x_n converges. Hence a local solution exists. It will be proven in the next chapter that every solution is indeed analytic in x_n. Hence the local solution is unique in the class of $C^{1,\alpha}$ solutions.

Remark 6.2.3. Assumption (6.8) can easily be achieved provided the expression is bounded by simply decreasing R_0 and R. The norm is $O(R + R_0)$ near zero.

Proof. We prove estimates assuming the solution exists. In the end we argue that these estimates actually imply existence of a solution. The problem can be rewritten in the form

$$x_n v_{nn} + (\sigma + 1)v_n = v_t - x_n \Delta_{\mathbb{R}^{n-1}} v - (\sigma + 1)g(\nabla\tilde{v}) - x_n \partial_n g(\nabla\tilde{v})$$

where v is a solution if $\tilde{v} = v$. Otherwise we obtain a map $\tilde{v} \to v$.

We apply ∂^α to both sides and evaluate at $x_n = 0$. To shorten the notation we set $\sigma_\alpha = \sigma + \alpha_n + 1$. Then

$$\partial_n \partial^\alpha v = \frac{\partial^\alpha \partial_t v}{\sigma_\alpha} - \frac{\alpha_n}{\sigma_\alpha}\partial^{\alpha^{(n)}} \Delta_{\mathbb{R}^{n-1}} v - \partial^\alpha g(\nabla v)$$

hence, if $\alpha_n \geq 1$,

$$\begin{aligned} l(\alpha)^{-1}|\partial^\alpha_{x,t}\partial_n v| \leq \; & (2n)^{-1}\frac{\alpha_n}{\sigma_\alpha} l(\alpha^{(n)}_{(0)})^{-1}|\partial^{\alpha^{(n)}_{(0)}}\partial_n v| \\ & + (2n)^{-1}\frac{\alpha_n}{|\alpha| + \alpha_n + \alpha_0}\sum_{i=1}^{n-1} l(\alpha^{(n)}_{(i)})^{-1}|\partial^{\alpha^{(n)}_{(i)}}\partial_i v| \\ & + l(\alpha)^{-1}|\partial^\alpha g(\nabla\tilde{v})|. \end{aligned}$$

The estimate is somewhat different but simpler if $\alpha_n = 0$. There we get $\partial^\alpha \partial_t v$ on the right hand side. For a given multi-index β there are exactly n multi-indices α such that β occurs on the right hand side. We sum over all α:

$$\|\partial_n v\|_{R,l} \le \frac{1}{2}\|\partial_n v\|_{R,l} + \|g(\nabla \tilde v)\|_{R,l} + \frac{1}{2}\|f\|_{l_0,R}$$

where l_0 refers to tangential derivatives. Thus $\|\partial_n v\|_{R,l} \le 2\|g\|_{R,l} + \|f\|_{l_0,R}$ and, since $\partial_n\partial_i v = \partial_i \partial_n v$,

$$\|\nabla v\|_{R,l} \le 3\|g\|_{R,l} + 2\|f\|_{l_0,R}.$$

We apply Lemma 6.1.5 and obtain

$$\|\nabla v\|_{R,l} \le 3n/2\|\nabla \tilde v\|^2_{R,l} + 2\|f\|_{l_0,R}$$

provided that $\|\nabla \tilde v\|_{R,l} \le 1/2$. In particular $\|\nabla v\|_{R,l} \le \frac{1}{4n}$ if this is true for $\|\tilde v\|_{R,l}$. This completes the apriori estimates.

There is a number of observations to be made. Given h and $\tilde v$ we can construct v as a formal power series. The estimate can be applied to a finite number of derivatives. Hence it holds for v constructed in that way. We obtain a map of the ball of radius $(4n)^{-1}$ to itself. The same is true for similar estimates for a finite number of derivatives at one point. This space is finite dimensional, hence there is a fixed point. This fixed point is unique since we can express the normal derivatives in term of tangential derivatives. We can express every derivative in terms of f. The argument above shows that the estimate holds if we restrict the summation to $|\alpha| \le N$ uniformly in N. The limit $N \to \infty$ gives the result. $\square$

6.3 Gevrey regularity and analyticity

Let $U \subset \mathbb{R}^{n-1} \times \mathbb{R}^+_0 \times \mathbb{R}$ be a relatively open neighborhood of the origin and $v \in C^2(U)$ with $v(0,0) = \nabla v(0,0) = 0$. We suppose that $\sigma > -1$ and

$$v_t - x_n \Delta v - (\sigma+1)\partial_n v = -x_n^{-\sigma}\partial_n \left(x_n^{1+\sigma} \frac{|\nabla v|^2}{1+\partial_n v} \right) =: w. \tag{6.9}$$

It follows from the previous construction that this defines a local solution to the porous medium equation which is, together with v, of class C^∞ in its support by Theorem 5.6.1.

For $R > 0$ (which we shall choose later) we set

$$l(\alpha) = \frac{(|\alpha| + \alpha_0)!}{R^{|\alpha|}}$$

and $B(x^0, t^0, R) := \{(x,t) : |x - x^0|^2 + (t-t^0)^2 < R^2, x_n \ge 0\}$.

Proposition 6.3.1. *There exists $R > 0$ such that, if $v \in C^2(B(0,0,R_0))$ satisfies* (6.9)*, $v(0) = 0$, $\nabla v(0) = 0$, and $\|v\|_{C^2(B(0,0,R_0))} \le \delta(n,\sigma)$, then*

$$\sup_{(x,t)\in B(0,0,R_0/2)} \sum_\alpha l(\alpha)^{-1} |\partial^\alpha \nabla_{x,t} v| \le c(n,\sigma)\|v\|_{C^2}.$$

Remark 6.3.2. The smallness assumption is not a serious restriction. It can be achieved by scaling $v \to \lambda v(./\lambda, ./\lambda)$ if necessary. The equation is invariant with respect to this scaling.

Proof. We shall prove these estimates in $B(0,0,R_0/2) \cap \{t \ge R_0/4\}$. This suffices because the problem is translation invariant. The estimates away from the boundary could be proved along the same lines. They are much simpler and we omit the proof here. The estimates which we shall prove are much more precise than the statement of the theorem.

Let $B_+ = \{(x,t) : |x| \le R_0, 0 < t < R_0, x_n > 0\}$. We suppose that all the assumptions hold true in B_+ instead of $B(0,0,R_0)$. Let $\phi(x,t) = t/R_0(1-(|x|/R_0)^2)$. It is nonnegative on B_+. We define the operator $L_\sigma v = v_t - x_n\Delta v - (\sigma+1)\partial_n v$. For later reference we state

$$L_{\sigma+\alpha_n}\partial^\alpha v = \alpha_n \partial^{\alpha^{(n)}} \Delta_{R^{n-1}} v + \partial^\alpha w \tag{6.10}$$

if $L_\sigma v = w$, and, since $[L_\sigma, \psi] = L_\sigma\psi - 2x_n(\partial_i\psi)\partial_i$,

$$\begin{aligned} L_{\sigma+\alpha_n}(\Phi^k\partial^\alpha v) \;&= \phi^k\partial^\alpha w + \phi^k\alpha_n\partial^{\alpha^{(n)}}\Delta_{\mathbb{R}^{n-1}} v + k(L_{\sigma+\alpha_n}\phi)\phi^{k-1}\partial^\alpha v \\ &\quad -2kx_n(\partial_i\phi)\phi^{k-1}\partial_i\partial^\alpha v - k(k-1)x_n|\nabla\phi|^2\phi^{k-2}\partial^\alpha v. \end{aligned} \tag{6.11}$$

We extend the data and $\phi^k\partial^\alpha v$ by zero outside B_+, suppose that $p > \min\{1, \frac{1}{1+\sigma}\}$, apply Theorem 2 in $t < R_0$, commute ϕ^k with a finite number of derivatives and use the abbreviation $\|.\|$ for $\|.\|_{L^p(B_+)}$ to get

$$\begin{aligned} \|\phi^k\partial^\alpha\partial_t v\| + \|\Phi^k x_n D_x^2\partial^\alpha v\| + \|\phi^k\partial^\alpha\nabla v\| \;&\le c_1(n,\sigma+\alpha_n)\Big(\|\phi^k\partial^\alpha w\| \\ +\|\phi^k\partial^{\alpha^{(n)}}\Delta v\| + k\|x_n\phi^{k-1}\partial^\alpha\nabla v\| &+ k\|\phi^{k-1}\partial^\alpha v\| + k(k-1)\|x_n|\phi|^{k-2}\partial^\alpha v\|\Big) \end{aligned} \tag{6.12}$$

The question is: how can we use all these estimates simultaneously?

Step 1, Estimates of tangential derivatives: Let $R > 0$, $p > \min\{1, \frac{1}{1+\sigma}\}$, $k(\alpha) = |\alpha| + \alpha_0$ and define

$$I_j = \sum_{\{\alpha:\alpha_n\le j\}} l(\alpha)^{-1}\left(\|\Phi^{k(\alpha)}x_n\partial^\alpha D_x^2 D_{x,t} v\| + \|\Phi^{k(\alpha)}\partial^\alpha D_{x,t}^2 v\|\right)$$

and

$$K_j = \sum_{\{\alpha:\alpha_n\le j\}} l(\alpha)^{-1}\|\Phi^{k(\alpha)}\partial^\alpha D_{x,t}w\|.$$

We claim that

$$I_j \le c_5 K_j + c_6\|v\|_{C^2} \tag{6.13}$$

provided that R is sufficiently small and $j \le n+1$.

We solve equation (6.10) for $\partial_t^{k+2}v$ and estimate time derivatives in terms of spatial derivatives, which we combine with (6.12) to

$$\begin{aligned}\|\Phi^{2k}\partial_t^{k+2}v\| &\le \|\Phi^{2k}\partial_t^{k+1}w\| + \|\Phi^{2k}x_nD_x^2\partial_t^{k+1}v\| + (\sigma+1)\|\Phi^{2k}\partial_t^{k+1}\partial_nv\| \\ &\le c_2\left(\|\Phi^{2k}\partial_t^{k+1}w\| + k\|\Phi^{2k-1}\partial_t^{k+1}\nabla v\| + k^2\|\Phi^{2(k-1)}\partial_t^{k+1}v\|\right).\end{aligned}$$

Similarly

$$\begin{aligned}\|\Phi^{2k+1}\partial_t^{k+2}\nabla v\| \ &\le c_3\Big(\|\Phi^{2k+1}\partial_t^{k+1}\nabla w\| + k\|\Phi^{2k}\partial_t^{k+1}w\| \\ &\qquad + k^2\|\Phi^{2k-1}\partial_t^{k+1}\nabla v\| + k^3\|\Phi^{2(k-1)}\partial_t^{k+1}v\|\Big).\end{aligned}$$

We define

$$I_{-1} = \sum_{i=0}^\infty\left(\frac{R^k}{(2k)!}\|\Phi^{2k}\partial_t^{k+2}v\| + \frac{R^{k+1}}{(2k+1)!}\|\Phi^{2k+1}\partial_t^{k+2}\nabla v\|\right)$$

and sum over k. Using the previous two estimates and shifting the summation by one we get

$$c_4^{-1}I_{-1} \le \sum_{k=0}^\infty\Big(\frac{R^k}{(2k)!}\|\Phi^{2k}\nabla\partial_t^{k+1}w\| + \frac{R^{k+1}}{(2k+1)!}\|\Phi^{2k+1}\partial_t^{k+1}\nabla w\|\Big) + RI_{-1} + \|v\|_{C^2}.$$

Here $C^2 = C^2(B_+)$. Hence, if $R \le 1/(2c_4)$,

$$I_{-1} \le 2c_4\Big[\sum_{k=0}^\infty\left(\frac{R^k}{(2k)!}\|\Phi^{2k}\nabla\partial_t^{k+1}w\| + \frac{R^{k+1}}{(2k+1)!}\|\Phi^{2k+1}\partial_t^{k+1}\nabla w\|\right) + \|v\|_{C^2}\Big].$$

By the same arguments $I_j \le c_5K_j + c_5RI_j + c_5I_{j-1}$ which implies for $R < \frac{1}{2c_5}$ the inequality (6.13) by induction on j. The argument here requires boundedness of the norm. We assume that all sums converge and deal with this question later.

Step 2, Estimates of vertical derivatives: Here we shall derive L^2 estimates of higher order vertical derivatives. This requires a new idea since the constants in the previous step deteriorate if α_n gets large. We set $p = 2$ in this step and use the same notation as in the previous step.

We redefine

$$l(\alpha) = \frac{(|\alpha + \alpha_0| + \alpha_n + \alpha_0)!}{R^{|\alpha|} r^{\alpha_n} \alpha_n!}$$

where r is a small number to be determined in the proof. Let $k(\alpha) = |\alpha| + \alpha_0 + \alpha_n$ and

$$J = \sum_\alpha l(\alpha)^{-1} \left[(2 + \alpha_n)^{-1} \|\Phi^{k(\alpha)} x_n D_x^2 \partial^\alpha D^{n+1} v\| + \|\Phi^{k(\alpha)} \partial^\alpha \nabla D^{n+1} v\| \right].$$

We claim

$$J \leq c \sum_\alpha l(\alpha)^{-1} (2 + \alpha_n)^{-1} \|\Phi^{k(\alpha)} \partial^\alpha D^{n+1} w\|_2 + c I_{n+1}. \tag{6.14}$$

The crucial point is the following substitute for estimate (6.12). Consider $L_\sigma u = w$ with $\sigma > 0$. Then $L_\sigma u(u_t - \sigma \partial_n u) = w(u_t - \sigma \partial_n u)$ and integration over B_+ gives

$$\|u_t - \sigma \partial_n u\|^2 + \sigma \| \, |\nabla u| \, \|^2 \leq \|w\|^2,$$

hence,

$$\sigma \|\partial_n u\| \leq \|u_t\| + \|w\|.$$

This inequality, applied to (6.11) for $\alpha_n > 0$, gives (with $k \geq 2$ and otherwise unrestricted)

$$\begin{aligned} (\sigma + \alpha_n) \|\phi^k \partial_n \partial^\alpha v\| \; &\leq \|\phi^k \partial^\alpha w\| + \|\phi^k \partial^\alpha v_t\| + \alpha_n \|\Phi^k \partial^{\alpha^{(n)}} \Delta_{\mathbb{R}^{n-1}} u\| \\ &\quad + ck(\sigma + \alpha_n) \|\phi^{k-1} \partial^\alpha v\| + ck \|x_n \phi^{k-1} \partial^\alpha \nabla v\|_2 + ck^2 \|\phi^{k-2} x_n \partial^\alpha v\|. \end{aligned}$$

We use this inequality for $\partial^\alpha D$ with D denoting the set of all derivatives of order $\leq n + 1$, and $k = |\alpha| + \alpha_0 + \alpha_n$.

$$\begin{aligned} l(\alpha)^{-1} \|\phi^k \partial_n \partial^\alpha D v\| \; &\leq \frac{1}{l(\alpha)(\alpha_n + \sigma)} \|\phi^k \partial^\alpha D w\| \\ &+ 2r \Big(l(\alpha^{(n)}_{(0)})^{-1} \|\phi^k \partial^{\alpha^{(n)}_{(0)}} \partial_n D v\| + \sum_{i=1}^{n-1} l(\alpha^{(n)}_{(i)})^{-1} \|\phi^k \partial^{\alpha^{(n)}_{(i)}} \partial_i D v\| \Big) \\ &+ crR \Big(l(\alpha^{(n)})^{-1} \|\phi^{k-1} \partial^{\alpha^{(n)}} \partial_n D v\| + l(\alpha^{(n)})^{-1} \|\phi^{k-1} x_n \partial^{\alpha^{(n)}} \nabla_x \partial_n D v\| \\ &\quad + l(\alpha^{(nn)})^{-1} \|\phi^{k-2} x_n \partial^{\alpha^{(nn)}} \partial_{nn} D v\| \Big). \end{aligned}$$

We set $k = |\alpha| + \alpha_0 + \alpha_n$. We sum with respect to α. If r and R are sufficiently small we obtain (6.14), using

$$\sum_{ij} \|x_n \partial_i \partial_j u\|^2 + \sum_{i=1}^{n-1} \|\partial_i u\|^2 = \|\sum_i \partial_i (x_n \partial_i u)\|^2 = \|u_t - \sigma \partial_n u - w\|^2$$

to estimate $x_n D_x^2 Dv$ in the last step.

Step 3, Estimate of w in terms of ∇v: Here we consider $1 \le p \le \infty$. We fix a constant M with $\frac{1}{p} < \frac{M}{n+1}$. We recall that $w = x_n^{-\sigma} \partial_n (x_n^{1+\sigma} g(\nabla v))$.

Lemma 6.3.3. *Let $k(\alpha)$ and $l(\alpha)$ be as in Step 1 or as in Step 2. Then the following estimate holds:*

$$\begin{aligned} &\sum_\alpha l(\alpha)^{-1} (2+\alpha_n)^{-1} \|\phi^{k(\alpha)} \partial^\alpha D^M w\| \\ &\qquad \le c \Big(\sum_\alpha l(\alpha)^{-1} \big[\|\phi^{k(\alpha)} \partial^\alpha D^M \nabla v\| + (2+\alpha_n)^{-1} \|x_n \phi^{k(\alpha)} \partial^\alpha D^M D_x^2 v\| \big] \Big)^2 \end{aligned}$$

if the right hand side is smaller than a constant depending only on n and σ.

Proof. To clarify the idea we only consider the cases $p = \infty$, $M = 0$ and $p > n+1$, $M = 1$. We begin with $p = \infty$ and $M = 0$. We have

$$\partial^\alpha [x_n^{-\sigma} \partial_n (x_n^{1+\sigma} g(\nabla u))] = x_n \partial^\alpha \partial_n g(\nabla u) + (\alpha_n + \sigma + 1) \partial^\alpha g(\nabla u).$$

The functions $l(\alpha)/\alpha!$ and $(2+\alpha_n)^{-1}$ are super-multiplicative and the sum over all α, (after division by $(2+\alpha_n)^{-1}$) is estimated in exactly the same way as in Proposition 6.1.4.

Now let $p > n+1$ and $M = 1$. Then we use the equivalent norm

$$\|D^1 v\|_p = \max\{\|v\|_{L^\infty(B_+)}, \|\nabla_{x,t} v\|_{L^p(B_+)}\}.$$

Let $G(R) = \|p \cdot \nabla g(p)\|_R$. The L^∞ term of the norm is covered by the case $p = \infty$, $M = 0$. Now

$$\Phi^{k(\alpha)} \partial^\alpha [g'(\nabla u)(\nabla_x \nabla_{x,t} u)]$$

is a sum of products. We distribute $\Phi^{k(\alpha)}$ over the factors according to the number of derivatives and take the L^p norm of the product. The first factor is estimated in L^∞, the second one in L^p. The combinatorial argument of Proposition 6.1.4 covers this situation. The modifications for the term $(2+\alpha_n)^{-1} x_n \partial^\alpha \nabla \partial_n g(\nabla u)$ are straight forward. $\square$

Step 4, Completion: We combine estimates (6.13), (6.14), Lemma 6.3.3 and Sobolev's inequality to get the estimate of Proposition 6.3.1 under the assumption that the norm on the left-hand side is sufficiently small. It is clear from the proof that these estimates also hold for finite sums. We have to show that the smallness condition of Lemma 6.3.3 is satisfied for a finite number of derivatives. This is again done by scaling: $v_\lambda(x,t) = \lambda v(\lambda^{-1}x, \lambda^{-1}t)$ is again a solution. We choose λ large. Then k^{th} order derivatives decay like λ^{1-k}. We apply the estimate to v_λ for λ sufficiently large. Then we decrease λ and see that the estimate holds as long as $\|v_\alpha\|_{C^2} \le \delta$. This completes the proof of Proposition 6.3.1. □

6.4 Notes

The use of von Mises variables for higher regularity of solutions to free boundary problems goes back at least to Kinderlehrer and Nirenberg [81], [80], who proved analyticity for the Stefan problem and related problems. A related transformation combined with a different technique was used by Angenent ([1] and [2]) to prove analyticity of the interface in the 1-D porous medium equation. The work of Escher and Simonett [52] applies a modification of the ideas of Angenent to a dynamic filtration problem.

There has been a recent interest in Gevrey regularity (more precisely: estimates in Banach spaces of analytic functions of a related but somewhat different type than ours) for parabolic equations (Ferrari and Titi [59]). It is not clear what their approach gives in the presence of boundaries. Closer to the Cauchy-Kowalevski theorem, Proposition 6.2.1, is the paper of Friedman [60].

The related regularity question in pseudo-convex domains is very interesting and deep. The positive analyticity results in the strongly pseudo-convex situation are due to Tartakoff [128], [129], Treves [131], Metevier [97], [98] and Geller [65]. The weakly pseudo-convex situation is currently intensively studied: see the survey of Christ [36] and Christ and Li [37].

Cauchy-Kowalevski (see Kowalevski [85]) results are fairly classical. Nevertheless our type of result in this sub-elliptic setting seems to be new. Finally we remark that one obtains analyticity in x and t in the setting of Theorem 5.5.1. The reason is that the proof of Theorem 6.3.1 gives joint analyticity in x and t once the solution is analytic in time (with values in suitable function spaces). Theorem 5.5.1 does not imply analyticity in time for coordinates used in here. It is nevertheless clear that the proof of Theorem 6.3.1 applies to more general nonlinearities, in particular to those occuring in Theorem 5.5.1.

Chapter 7

Notation

$\overline{A}$	The closure of the set A
A^0	The complement of the set A
$\overset{o}{A}$	The interior of the set A
$\mathbb{R}$	The real numbers
$\mathbb{C}$	The complex numbers
$\mathbb{R}^n$	The vector space
H	The half space $\{x \in \mathbb{R}^n : x_n > 0\}$
Ω	A bounded domain in $\mathbb{R}^n$, defined by ϕ by $\Omega = \{x : \phi(x) > 0\}$.
$d(x, y)$	distance, often a Riemannian distance with respect to the Riemannian quadratic form $\|v\|_x = \phi(x)\|v\|^2$
$B(x, R)$	open ball centered at x of radius R, often with respect to a Riemannian metric
$Q_R(x, t)$	The cylinder $B(x, R) \times (t - R^2, 0]$
μ_σ	The measure $\phi^\sigma dx$ in bounded domains defined by ϕ resp. $x_n^\sigma dx$
μ_σ^0	$\mu_\sigma dt$
$\|A\|$	measure of the set A with implicit understood measure, or the deteminant of the matrix A
$\|A\|_\sigma$	$\mu_\sigma(A)$ for $A \in \Omega$ or $A \in H$
$\|A\|_{\sigma,0}$	$\mu_\sigma^0(A)$ for $A \in \Omega \times \mathbb{R}$ or $A \in H \times \mathbb{R}$

$\lvert B_R(x)\rvert_\sigma$	is comparable to $R^n\phi(x)^{\sigma+n} + R^{2n+2\sigma}$, Chapter 4
f_+	the positive part of f, defined by $f_+ = \max\{0, f\}$
$M, \tilde{M}$	the Hardy-Littlewood maximal function
$\hat{f}$	the Fourier transform of f, $(2\pi)^{-n/2}\int e^{-ix\xi} f(x)\,dx$
$g(x,t,y,s)$	generalized heat kernel
L^p	equivalence classes of Lebesgue measurable functions with $\lvert u\rvert^p$ integrable
$L^p(\mu)$	L^p space with respect to the measure μ
$L^p(\omega)$	the L^p-space on $\omega \subset \mathbb{R}^n$
$L^{pq}(\mu)$	the Lorentz space with norm $\lVert f\rVert_{pq}^q = \int_0^\infty \lambda^q \mu(\{\lvert f\rvert > \lambda\})^{q/p} \frac{d\lambda}{\lambda}$ with the obvious modification if $q = \infty$
L^p_w	the same as $L^{p,\infty}$
A_p	the Muckenhoupt class
A_∞	the union of all Muckenhoupt weights
$\lvert A\rvert$	determinant of the matrix A
$\Gamma(s)$	the Gamma function
K_ν	the Macdonald function

Bibliography

[1] S. Angenent. Analyticity of the interface of the porous media equation after the waiting time. *Proc. Am. Math. Soc.*, 102(2):329–336, 1988.

[2] S. Angenent. Nonlinear analytic semiflows. *Proc. Roy. Soc. Edinburgh Sect. A*, 115(1-2):91–107, 1990.

[3] S. Angenent. Solutions of the one-dimensional porous medium equation are determined by their free boundary. *J. Lond. Math. Soc., II. Ser.*, 42(2):339–353, 1990.

[4] S. Angenent and D. G. Aronson. The focusing problem for the radially symmetric porous medium equation. *Comm. Partial Differential Equations*, 20(7-8):1217–1240, 1995.

[5] S.B. Angenent and D.G. Aronson. Optimal asymptotics for solutions to the initial value problem for the porous medium equation. In T. S. Angell, editor, *Nonlinear problems in applied mathematics*, pages 10–19. SIAM, 1996.

[6] D. Aronson, M. G. Crandall, and L.A. Peletier. Stabilization of solutions of a degenerate nonlinear diffusion problem. *Nonlinear Anal., Theory Methods Appl*, 6:1001–1022, 1982.

[7] D. G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc. (N.S.)*, 73:890–896, 1967.

[8] D. G. Aronson. Regularity properties of flows through porous media. *SIAM J. Appl. Math.*, 17:461–467, 1969.

[9] D. G. Aronson. The porous medium equation. In *Nonlinear diffusion problems*, volume LNM 1224, pages 1–46. Springer, 1986.

[10] D. G. Aronson and Ph. Benilan. Regularité des solutions de l'équation des milieux poreux dans $\mathbb{R}^n$. *C. R. Acad. Sci. Paris Sér. I Math.*, 288:103–105, 1979.

[11] D. G. Aronson and L. A. Caffarelli. The initial trace of a solution of the porous medium equation. *Trans. Amer. Math. Soc.*, 280:351–366, 1983.

[12] D. G. Aronson and L. A. Caffarelli. Optimal regularity for one-dimensional porous medium flow. *Rev. Mat. Iberoam.*, 2:357–366, 1986.

[13] D. G. Aronson and J. Graveleau. A selfsimilar solution to the focusing problem for the porous medium equation. *Eur. J. Appl. Math.*, 4(1):65–81, 1993.

[14] D. G. Aronson and J. L. Vazquez. Eventual C^∞-regularity and concavity for flows in one-dimensional porous media. *Arch. Rational Mech. Anal.*, 99:329–348, 1987.

[15] D.G. Aronson. Regularity properties of flows through porous media: The interface. *Arch. Rational Mech. Anal.*, 37:1–10, 1970.

[16] D.G. Aronson, L.A. Caffarelli, and J.L. Vazquez. Interfaces with a corner point in one-dimensional porous medium flow. *Commun. Pure Appl. Math.*, 38:375–404, 1985.

[17] I. Athanasopoulos, L. Caffarelli, and S. Salsa. Caloric functions in Lipschitz domains and the regularity of solutions to phase transition problems. *Ann. of Math. (2)*, 143(3):413–434, 1996.

[18] I. Athanasopoulos, L. Caffarelli, and S. Salsa. Regularity of the free boundary in parabolic phase-transition problems. *Acta Math.*, 176(2):245–282, 1996.

[19] I. Athanasopoulos, L. Caffarelli, and S. Salsa. Phase transition problems of parabolic type: Flat free boundaries are smooth. *Comm. Pure Appl. Math.*, 11(1):77–112, 1998.

[20] M. Avellaneda and F. Lin. L^p bounds on singular integrals in homogenization. *Comm. Pure Appl. Math.*, 44(8/9):897–910, 1991.

[21] J.T. Beale. The initial value problem for the Navier-Stokes equations with a free surface. *Comm. Pure Appl. Math.*, 34:359–392, 1984.

[22] J.T. Beale. Large-time regularity of viscous surface waves. *Arch. Rational Mech. Anal.*, 84:307–352, 1984.

[23] J. Bemelmans. Eine Außenraumaufgabe für die instationären Navier-Stokes-Gleichungen. *Math. Z.*, 162:145–173, 1978.

[24] Ph. Benilan and M. G. Crandall. The continuous dependence on ϕ of solutions of $u_t - \Delta\phi(u) = 0$. *Indiana Univ. Math. J.*, 30:161–177, 1981.

[25] Ph. Benilan, M. G. Crandall, and M. Pierre. Solutions of the porous medium equation in $\mathbb{R}^N$ under optimal conditions on the initial values. *Indiana Univ. Math. J.*, 33:51–87, 1984.

[26] P. Bolley, J. Camus, and G. Metivier. Estimations de Schauder et regularité Hoelderienne pour une classe de problèmes aux limites singuliers. (Schauder estimates and Hoelder regularity for a class of singular boundary value problems). *Comm. Partial Differential Equations*, 11:1135–1203, 1986.

[27] L. A. Caffarelli and A. Friedman. Continuity of the density of a gas flow in a porous medium. *Trans. Amer. Math. Soc*, 252:99–113, 1979.

[28] L. A. Caffarelli and A. Friedman. Regularity of the free boundary of a gas flow in an n-dimensional porous medium. *Indiana Univ. Math. J.*, 29:361–391, 1980.

[29] L. A. Caffarelli and I. Peral. On $W^{1,p}$ estimates for elliptic equations in divergence form. *Comm. Pure Appl. Math.*, 51(1):1–22, 1998.

[30] L. A. Caffarelli, J. L. Vazquez, and N. I. Wolanski. Lipschitz continuity of solutions and interfaces of the n-dimensional porous medium equation. *Indiana Univ. Math. J.*, 36:373–401, 1987.

[31] L. A. Caffarelli and N. I. Wolanski. $C^{1,\alpha}$ regularity of the free boundary for the n-dimensional porous media equation. *Comm. Pure Appl. Math.*, 43(7):885–902, 1990.

[32] A. Calderón and A. Zygmund. On the existence of certain singular integrals. *Acta Math.*, 88:85–139, 1952.

[33] C. Carathéodory. Untersuchungen über die Grundlagen der Thermodynamik. *Math. Ann.*, 67:355–386, 1909.

[34] W. Chow. Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung. *Math. Ann.*, 117:98–105, 1939.

[35] M. Christ. *Lectures on singular integral operators*, volume 77 of *Regional Conference Series in Mathematics*. American Mathematical Society, Providence, RI, 1990.

[36] M. Christ. Remarks on analytic hypoellipticity of $\overline{\partial}_b$. In T. Bloom, D. Catlin, J. d'Angelo, and Y.-T. Siu, editors, *Modern methods in complex analysis. The Princeton conference in honor of Robert C. Gunning and Joseph J. Kohn, Princeton University, Princeton, NJ, USA, Mar. 16-20, 1992.*, volume 137 of *Ann. Math. Stud.*, pages 41–62. Princeton University Press, 1995.

[37] M. Christ and S.-Y. Li. On analytic solvability and hypoellipticity for $\overline{\partial}$ and $\overline{\partial}_b$. *Math. Z.*, 226(4):599–606, 1997.

[38] R. Coifman and C. Fefferman. Weighted norm inequalities for maximal functions and singular integrals. *Studia Math.*, 51:241–250, 1974.

[39] R. Coifman, P. Jones, and J. L. Rubio de Francia. Constructive decomposition of BMO functions and factorization of A_p-weights. *Proc. Amer. Math. Soc.*, 87:675–676, 1983.

[40] R. Coifman and R. Rochberg. Another characterization of BMO. *Proc. Amer. Math. Soc.*, 79:249–254, 1980.

[41] M. Cowling, S. Giulini, A. Hulanicki, and G. Mauceri. Spectral multipliers for a distinguished Laplacian on certain groups of exponential growth. *Stud. Math.*, 111(2):103–121, 1994.

[42] B. E. J. Dahlberg and C. E. Kenig. Nonnegative solutions of the porous medium equation. *Comm. Partial Differential Equations*, 9:409–437, 1984.

[43] B. E. J. Dahlberg and C.E. Kenig. Nonnegative solutions of the initial-Dirichlet problem for generalized porous medium equations in cylinders. *J. Am. Math. Soc.*, 1(2):401–412, 1988.

[44] B. E.J. Dahlberg and C. E. Kenig. Weak solutions of the porous medium equation in a cylinder. *Trans. Am. Math. Soc.*, 336(2):701–709, 1993.

[45] E. Damek, A. Hulanicki, and J. Zienkiewicz. Estimates for the Poisson kernels and their derivatives on rank one *na* groups. *Stud. Math.*, 126(2):115–148, 1997.

[46] J. P. D'Angelo. Finite type conditions and subelliptic estimates. In *Bloom, Thomas (ed.) et al., Modern methods in complex analysis. The Princeton conference in honor of Robert C. Gunning and Joseph J. Kohn, Princeton University, Princeton, NJ, USA, Mar. 16-20, 1992.*

Princeton, NJ, volume 137 of *Ann. Math. Stud.*, pages 63–78. Princeton University Press, 1995.

[47] P. Daskalopoulos and R. Hamilton. The free boundary for the N-dimensional porous medium equation. *Int. Math. Res. Not.*, 17:817–831, 1997.

[48] P. Daskalopoulos and R. Hamilton. Regularity of the free boundary for the porous medium equation. *J. Amer. Math. Soc.*, 11(4):899–965, 1998.

[49] G. David and J.-L. Journé. A boundedness criterion for generalized Calderón-Zygmund operators. *Ann. of Math. (2)*, 120:371–397, 1984.

[50] De Giorgi, E. Sulla differenziabilita e l'analiticita d'elle estremali degli integrali mutlipli regolari. *Mem. Acc. Sci. Torino, Cl. Sc. Fis: Mat. Nat.(3)*, pages 25–43, 1957.

[51] P. Drabek, A. Kufner, and F. Nicolosi. *Quasilinear elliptic equations with degenerations and singularities.* de Gruyter Series in Nonlinear Analysis and Applications. 5. Walter de Gruyter, 1997.

[52] J. Escher and G. Simonett. Analyticity of the interface in a free boundary problem. *Math. Ann.*, 305:439–459, 1996.

[53] E. B. Fabes and D. W. Stroock. A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash. *Arch. Rational Mech. Anal.*, 96(4):327–338, 1986.

[54] R. Farwig. A variational approach in weighted Sobolev spaces to the operator $-\Delta + \partial/\partial x_1$ in exterior domains of $\mathbb{R}^3$. *Math. Z.*, 210(3):449–464, 1992.

[55] R. Farwig. The stationary exterior 3D-problem of Oseen and Navier-Stokes equations in anisotropically weighted spaces. *Math. Z.*, 211(3):409–448, 1992.

[56] R. Faxén. Fredholm'sche Integralgleichungen zu der Hydrodynamik zäher Flüssigkeiten. *I.Ark. Mat. Astr. Fys.*, 21 A(14), 1928/29.

[57] C. Fefferman. The uncertainty principle. *Bull. Amer. Math. Soc. (N.S.)*, 9:129–206, 1983.

[58] C. Fefferman and A. Sanchez-Calle. Fundamental solutions for second order subelliptic operators. *Ann. of Math. (2)*, 124:247–272, 1986.

[59] A. B. Ferrari and E. S. Titi. Gevrey regularity for nonlinear analytic parabolic equations. *Comm. Partial Differential Equations*, 23:1–16, 1998.

[60] A. Friedman. A new proof and generalizations of the Cauchy-Kowalewski theorem to nonanalytic and to non-normal systems. *Proc. Sympos. Pure Math.*, 4:115–119, 1961.

[61] A. Friedman. *Partial differential equations.* Holt, Rinehart and Winston, New York, 1969.

[62] G.P. Galdi, J.G. Heywood, and Y. Shibata. On the global existence and convergence to steady state of Navier-Stokes flow past an obstacle that is started from rest. *Arch. Rational Mech. Anal.*, 138(4):307–318, 1997.

[63] J. Garcia-Cuerva and J. L. Rubio de Francia. *Weighted norm inequalities and related topics.* North-Holland Mathematics Studies, 116 Notas de Matematica (104). North-Holland, Amsterdam - New York - Oxford, 1985.

[64] F. Gehring. The L^p integrability of the partial derivatives of a quasiconformal mapping. *Acta Math.*, 130:265–277, 1973.

[65] D. Geller. *Analytic pseudo-differential operators for the Heisenberg group and local solvability*, volume 37 of *Math. Notes.* Princeton Univ.Press, 1990.

[66] M. Giaquinta. *Multiple integrals in the calculus of variations.* Annals of Math. Studies. Princeton Univ. Press, 1983.

[67] M. Grueter and K.-O. Widman. The Green function for uniformly elliptic equations. *Manuscr. Math.*, 37:303–342, 1982.

[68] E. I. Hanzawa. Classical solutions of the Stefan problem. *Tôhoku Math. J. (2)*, 33:297–335, 1981.

[69] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.

[70] L. Hörmander. *The analysis of linear partial differential operators I*, volume 256 of *Grundlehren der mathematischen Wissenschaften.* Springer Verlag, 1983.

[71] L. Hörmander. *The analysis of linear partial differential operators II*, volume 257 of *Grundlehren der mathematischen Wissenschaften*. Springer Verlag, 1983.

[72] L. Hörmander. *The analysis of linear partial differential operators III*, volume 274 of *Grundlehren der mathematischen Wissenschaften*. Springer Verlag, 1985.

[73] R.A. Hunt, B. Muckenhoupt, and R.L. Wheeden. Weighted norm inequalities for the conjugate function and Hilbert transform. *Trans. Amer. Math. Soc.*, 176:227–251, 1973.

[74] D. Jerison and C. E. Kenig. The inhomogeneous Dirichlet problem in Lipschitz domains. *J. Funct. Anal.*, 130(1):161–219, 1995.

[75] D. Jerison and A. Sanchez-Calle. Subelliptic second order differential operators. In *Complex analysis III, Proc. Spec. Year, College Park/Md. 1985-86*, volume 1277 of *Lect. Notes Math.*, pages 46–77. Springer, 1987.

[76] D. S. Jerison and A. Sanchez-Calle. Estimates for the heat kernel for a sum of squares of vector fields. *Indiana Univ. Math. J.*, 35:835–854, 1986.

[77] P. W. Jones. Factorization of A_p weights. *Ann. of Math. (2)*, 111:511–530, 1980.

[78] B. Kawohl. *Rearrangements and convexity of level sets in PDE.*, volume 1150 of *Lecture Notes in Mathematics*. Springer-Verlag., Berlin etc., 1985.

[79] C. E. Kenig. *Harmonic analysis techniques for second order elliptic boundary value problems*, volume 83 of *CBMS*. AMS, 1994.

[80] D. Kinderlehrer and L. Nirenberg. Analyticity at the boundary of solutions of nonlinear second-order parabolic equations. *Comm. Pure Appl. Math.*, 31:283–338, 1978.

[81] D. Kinderlehrer and L. Nirenberg. The smoothness of the free boundary in the one phase Stefan problem. *Comm. Pure Appl. Math.*, 31:257–282, 1978.

[82] A. Klingler. New derivation of the Heisenberg kernel. *Comm. Partial Differential Equations*, 22:2051–2060, 1997.

[83] H. Koch. Classical solutions to phase transition problems are smooth. *Comm. Partial Differential Equations*, 23:389–438, 1998.

[84] A. N. Kolmogorv. Zufällige Bewegungen. *Ann. of Math. (2)*, 35:116–117, 1934.

[85] S. Kowalevski. Zur Theorie der partiellen Differentialgleichungen. *J. Reine Angew. Math.*, 80:1–32, 1857.

[86] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. II. *J. Fac. Sci., Univ. Tokyo, Sect. I A*, 32:1–76, 1985.

[87] S. Kusuoka and D. Stroock. Some boundedness properties of certain stationary diffusion semigroups. *J. Funct. Anal*, 60:243–264, 1985.

[88] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. III. *J. Fac. Sci., Univ. Tokyo, Sect. I A*, 34:391–442, 1987.

[89] S. Kusuoka and D. Stroock. Long time estimates for the heat kernel associated with a uniformly subelliptic symmetric second order operator. *Ann. Math., II. Ser.*, 127(1):165–189, 1988.

[90] S. Z. Levendorskij and B. Paneah. Degenerate elliptic problems and boundary problems. In Y. V. Egorov and M. A. Shubin, editors, *Partial differential equations 6*, pages 131–201. Springer, 1994. Translation from Itogi nauki i tekhniki, Sovremennye problemy matematiki, Fundamental'nye napravlemiya, Vol 63, Differentsial'nye uravneniya s chastnymi proizvodnymi 6, VINITI, Moscow 1990.

[91] P. I. Lizorkin. On multipliers of Fourier integrals in the space L_p. *Trudy Mat. Inst. Steklov*, 89:269–290, 1967. Engl. trans. in Proc. Steklov Inst. Math. 89, 269-290, 1967.

[92] J. Marcinkiewicz. Sur les séries de Fourier. *Fund. Math.*, 27:38–69, 1936.

[93] H. Matano. The global existence of classical solutions of the one-phase Stefan problem. *Sugaku*, 36:247–254, 1984.

[94] A. M. Meirmanov. The classical solution of the multidimensional Stefan problem for quasilinear parabolic equations. *Math. USSR-Sb.*, 112(2):170–192, 1980.

[95] A. M. Meirmanov. *The Stefan Problem.* De Gruyter Berlin, New York, 1992.

[96] A.M. Meirmanov, V.V. Pukhnachov, and S.I. Shmarev. *Evolution equations and Lagrangian coordinates.*, volume 24 of *De Gruyter Expositions in Mathematics.* de Gruyter, Berlin, 1997.

[97] G. Metivier. Hypoellipticité analytique sur des groupes nilpotents de rang 2. *Duke Math. J.*, 47:195–221, 1980.

[98] G. Metivier. Analytic hypoellipticity for operators with multiple characteristics. *Comm. Partial Differential Equations*, 6:1–90, 1981.

[99] J. Moser. On Harnack's theorem for elliptic differential equations. *Comm. Pure Appl. Math.*, 14:577–591, 1961.

[100] J. Moser. A Harnack inequality for parabolic differential equations. *Comm. Pure Appl. Math.*, 17:101–134, 1964. Err.: Vol. 20, 231-236, 1967.

[101] B. Muckenhoupt. Weighted norm inequalities for the Hardy maximal function. *Trans. Amer. Math. Soc.*, 165:207–226, 1972.

[102] B. Muckenhoupt. The equivalence of two conditions for weight functions. *Studia math.*, 49:101–106, 1974.

[103] D. Müller. Functional calculus on Lie groups and wave propagation. *Documenta Mathematica*, Extra volume ICM 98 2:679–690, 1998.

[104] D. Müller and F. Ricci. Analysis of second-order differential operators on the Heisenberg group 1. *Invent. Math.*, 101:545–582, 1990.

[105] D. Müller and E. M. Stein. L^p–estimates for the wave equation on the Heisenberg group. Technical report, Berichtsreihe des Mathematischen Seminars Kiel, 1997.

[106] A. I. Nachman. The wave equation on the Heisenberg group. *Comm. Partial Differential Equations*, 7:675–714, 1982.

[107] A. Nagel, E. M. Stein, and S. Wainger. Balls and metrics defined by vector fields. I: Basic properties. *Acta Math.*, 155:103–147, 1985.

[108] J. Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, 80:931–954, 1958.

[109] L. Nirenberg. On elliptic partial differential equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 13:115–162, 1959.

[110] O. A. Olejnik, A.S. Kalashnikov, and C. Yui-Lin. The Cauchy problem and boundary problems for equations of the type of nonlinear filtration. *Izv. Akad. Nauk SSSR, Ser. Mat.*, 22:667–704, 1958.

[111] O. A. Olejnik and E.V. Radkevich. *Second-order equations with non-negative characteristic form.* Itogi Nauki Tekh., Ser. Mat. Anal., English trans.: AMS, 1971.

[112] B. Opic and A. Kufner. *Hardy-type inequalities.* Pitman Research Notes in Mathematical Series. Longman House, Harlow, 1990.

[113] C. W. Oseen. *Neuere Methoden und Ergebnisse in der Hydrodynamik.* Akademische Verlagsgesellschaft, Leipzig, 1927.

[114] M. Pierre. Uniqueness of the solutions of $u_t - \Delta(\phi(u)) = 0$ with initial datum a measure. *Nonlinear Anal., Theory Methods Appl.* , 6:175–187, 1982.

[115] R. Remmert. *Funktionentheorie 2.* 2^{nd} ed. Springer, 1995.

[116] I.M. Rokhman and V.A. Solonnikov. Estimates in L_p weight norms for singular integrals with anisotropic kernels. *Zap. Nauchn. Semin. Leningr. Otd. Mat. Inst. Steklova*, 147:124–137, 1985. Engl. Trans.: J. Soviet Math. 37(1):869-876, 1987.

[117] M. Rosenblum. Summability of Fourier series in $L^p(d\mu)$. *Trans. Amer. Math. Soc.*, 105:32–42, 1962.

[118] L. P. Rothschild and E. Stein. Hypoelliptic differential operators and nilpotent groups. *Acta Math.*, 137:247–320, 1976.

[119] A. Sanchez-Calle. Fundamental solutions and geometry of the sum of squares of vector fields. *Invent. Math.*, 78:143–160, 1984.

[120] A. Sanchez-Calle. L^p estimates for degenerate elliptic equations. *Rev. Mat. Iberoam.*, 4(1):177–185, 1988.

[121] N. Shimakura. Les fonctions de Green pour certains opérateurs paraboliqeues dégénérés dans le demi-espace. (the Green functions for certain degenerate parabolic operators in the half-space). *Proc. Japan Acad.*, 47:699–704, 1971.

[122] V.A. Solonnikov. Unsteady motion of a finite mass of fluid, bounded by a free surface. *J. Sov. Math.*, 40(5):672–686, 1988.

[123] V.A. Solonnikov. Estimates of Oseen's potentials in weighted Hoelder spaces. *Math. Nachr.*, 177:307–321, 1996.

[124] E. M. Stein. *Singular integrals and differentiability properties of functions*, volume 30 of *Princeton Mathematical Series*. Princeton University Press, 1970.

[125] E. M. Stein. *Harmonic analysis*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, 1993.

[126] E. M. Stein and G. Weiss. *Introduction to Fourier analysis in Euclidean spaces*, volume 32 of *Princeton Mathematical Series*. Princeton University Press, 1971.

[127] J.-O. Stroemberg and A. Torchinsky. *Weighted Hardy spaces*, volume 1381 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1989.

[128] D. S. Tartakoff. Local analytic hypoellipticity for $\Box_b$ on nondegenerate Cauchy-Riemann manifolds. *Proc. Natl. Acad. Sci. USA*, 75:3027–3028, 1978.

[129] D. S. Tartakoff. The local real analyticity of solutions to $\Box_b$ and the $\bar{\partial}$–Neumann problem. *Acta Math.*, 145:177–204, 1980.

[130] E. C. Titchmarsh. *Introduction to the theory of Fourier integrals*. Oxford, 1937.

[131] F. Treves. Analytic hypo-ellipticity of a class of pseudodifferential operators with double characteristics and applications to the $\overline{\partial}$–Neumann problem. *Comm. Partial Differential Equations*, 3:475–642, 1978.

[132] H. Triebel. *Interpolation theory, function spaces, differential operators*. North-Holland, 1978.

[133] N. Th. Varopoulos, L. Saloff-Coste, and T. Coulhon. *Analysis on Lie groups*, volume 100 of *Cambridge Tracts in Mathematics. Cambridge etc.: Cambridge University*. Cambridge University Press, 1992.

[134] J. L. Vazquez. Symétrisation pour $u_t = \Delta\phi(u)$ et applications. *C. R. Acad. Sci. Paris Sér. I Math.*, 295:71–74, 1982.

[135] J. L. Vazquez. Regularity of solutions and interfaces of the porous medium equation via local estimates. *Proc. R. Soc. Edinb., Sect. A*, 112(1/2):1–13, 1989.

[136] J. L. Vazquez. An introduction to the mathematical theory of the porous medium equation. In *Shape optimization and free boundaries, Proc. NATO ASI, Semin. Math. Super., Montreal/Can. 1990*, volume 380 of *NATO ASI Ser., Ser. C*, pages 347–389, 1992.

[137] V. Vespri. Analytic semigroups, degenerate elliptic operators and applications to nonlinear Cauchy problems. *Ann. Mat. Pura Appl. (4)*, 155:353–388, 1989.

[138] G. N. Watson. *A treatise on the theory of Bessel functions.* Cambridge University Press, 2nd edition, 1944.

[139] A. Zygmund. *Trigonometric series.* Warsaw, 1935.

[140] A. Zygmund. *Trigonometric series, (2nd edition), I and II.* Cambridge, 1959.