PSEUDO DIFFERENTIAL OPERATORS GENERATING MARKOV PROCESSES

Der Fakultät für Mathematik
der Universität Bielefeld

als Habilitationsschrift

vorgelegt von

Walter Hoh

aus Nürnberg

Contents

1 Introduction **3**

Notations **11**

2 Negative definite symbols **13**
2.1 Negative definite functions 13
2.2 The Lévy–Khinchin formula and the positive maximum principle 17

3 The martingale problem: Existence of solutions **28**
3.1 The martingale problem for jump processes 28
3.2 The solution of the martingale problem for pseudo differential operators 36

4 Generators of Feller semigroups **46**
4.1 Technical preliminaries 46
4.2 The construction of Feller semigroups 52

5 The martingale problem: Uniqueness of solutions **61**
5.1 Localization 61
5.2 A general uniqueness criterion 64
5.3 Well-posedness of the martingale problem for a class of pseudo differential operators 68
5.4 The Feller property 77

6 A symbolic calculus **83**
6.1 General remarks 83
6.2 The symbol classes $S_\varrho^{m,\lambda}$ and $S_0^{m,\lambda}$ 86
6.3 A calculus for $S_\varrho^{m,\lambda}$ and $S_0^{m,\lambda}$ 91
6.4 Friedrichs symmetrization 100
6.5 Application to generators of Feller semigroups 106
6.6 Perturbation results 109

7 Operators of variable order **116**
7.1 Statement of results 116
7.2 Existence of solutions 118
7.3 Regularity of solutions 123
7.4 Localization by the martingale problem 124

8 Associated Dirichlet forms, hyper-contractivity estimates, and the strong Feller property **126**

9 A non-explosion result **137**

Bibliography **147**

Chapter 1

Introduction

The present work is concerned with a class of pseudo differential operators which arise as generators of Markov processes and Feller semigroups.

Pseudo differential operators are a well known tool from the theory of partial differential equations, but at first sight there seems to be no closer relation to probability theory. To understand how the notion of pseudo differential opertors enters into the field of stochastic processes, it is most convenient to consider first the well-studied case of Lévy-processes, since the investigation of Lévy-processes by Fourier analytical methods has a long tradition in probability theory going back to S. Bochner.
By definition a Lévy-process $(X_t)_{t\geq 0}$ is a process on $\mathbb{R}^n$ with independent and stationary increments which is continuous in probability. The independence of the stationary increments immediately implies that the distributions $\mu_t = \mathcal{L}(X_{s+t} - X_s) = \mathcal{L}(X_t - X_0)$ of the increments form a convolution semigroup $(\mu_t)_{t\geq 0}$ of probability measures. Consequently, if we turn to the Fourier transforms of the measures μ_t, the semigroup property yields that there is a continuous function $\psi : \mathbb{R}^n \to \mathbb{C}$ such that the Fourier transforms are given by

$$\hat{\mu}_t(\xi) = e^{-t\psi(\xi)}.$$

For every initial distribution this function ψ completely determines the distribution of the Lévy-process (X_t) and is called the characteristic exponent of the Lévy-process.

Now a function $\psi : \mathbb{R}^n \to \mathbb{C}$ such that $\psi(0) \geq 0$ and which has the property that $e^{-t\psi}$ is a positive definite function in the sense of Bochner for all $t \geq 0$ is called a negative definite function (see Theorem 2.3 for equivalent definitions). Thus by the theorem of Bochner, [4], Theo.3.2, the characteristic exponent of a Lévy-process is a continuous negative definite function and it turns out that there is a one-to-one correspondence between Lévy-processes, the corresponding convolution semigroups and continuous negative definite functions.

Note that moreover properties of the Lévy-process like transience and recurrence, conservativeness and path properties can easily be deduced from properties of the characteristic exponent ψ, see for example [73], [25] and [11]. Furthermore every continuous negative definite function admits a unique representation by the Lévy-Khinchin formula, see Theorem 2.13, which describes diffusion, drift, killing and jump part of the Lévy-process in the Lévy decomposition, see [74].

In addition it is well-known that the convolution semigroup $(\mu_t)_{t\geq 0}$ induces a semigroup $(T_t)_{t\geq 0}$ on the space of continuous functions vanishing at infinity given by $T_t u = u * \mu_t$, which is a translation invariant Feller semigroup and all translation invariant Feller semigroups are obtained in this way, see [4]. In this situation the convolution theorem for the Fourier transform then yields for testfunctions $u \in C_0^\infty(\mathbb{R}^n)$

$$T_t u(x) = \int_{\mathbb{R}^n} e^{i(x,\xi)} \hat{\mu}_t(\xi) \cdot \hat{u}(\xi)\, đ\xi = \int_{\mathbb{R}^n} e^{i(x,\xi)} e^{-t\psi(\xi)} \cdot \hat{u}(\xi)\, đ\xi, \qquad đ\xi = (2\pi)^{-n}\, d\xi,$$

and therefore the generator A of this semigroup is given by

$$Au(x) = \lim_{t\to 0} \frac{1}{t}(T_t u(x) - u(x)) = \lim_{t\to 0} \int_{\mathbb{R}^n} e^{i(x,\xi)} \frac{e^{-t\psi(\xi)} - 1}{t} \cdot \hat{u}(\xi)\, đ\xi = -\int_{\mathbb{R}^n} e^{i(x,\xi)} \psi(\xi) \cdot \hat{u}(\xi)\, đ\xi.$$

Thus in the translation invariant case of Lévy-processes it turns out that the situation is completely described by the charactristic exponent, i.e. the continuous negative definite function ψ. In particular the semigroup operators T_t and the generator A are types of Fourier multiplier operators and the multiplier of the generator is (up to the sign) just given by the continuous negative definite function ψ.

Therefore it is reasonable to conjecture that a similar representation of the generator holds also in the non-translation invariant case, i.e. the generator is given by an operator

$$Au(x) = -p(x,D)u(x) = -\int_{\mathbb{R}^n} e^{i(x,\xi)} p(x,\xi) \cdot \hat{u}(\xi)\, đ\xi,$$

where $\xi \mapsto p(x,\xi)$ is a continuous negative definite function for each fixed $x \in \mathbb{R}^n$.
Operators of this structure are called pseudo differential operators, the function $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ is the symbol of the operator. Symbols which are continuous negative definite functions with respect to the second variable we call negative definite symbols.

The link between operators of this type and generators of Markov processes now is given by a theorem of Ph. Courrège [13] concerning the positive maximum principle. Note that a generator of a Markov process (X_t) in $\mathbb{R}^n$ satisfies the positive maximum principle, that is for each function φ in the domain of the generator A which attains its nonnegative maximum in a point $x_0 \in \mathbb{R}^n$ we have $A\varphi(x_0) \leq 0$, since obviously $\mathbb{E}^{x_0}[\varphi(X_t) - \varphi(x_0)] \leq 0$, where $\mathbb{E}^{x_0}$ denotes the expectation of the process started at x_0. In this way the positive maximum principle for the generator describes the fact that the associated semigroup is the transition semigroup of a process, i.e. is positivity preserving. Moreover the positive maximum principle also characterizes the submarkovian property of Feller semigroups. More precisely in the case of strongly continuous semigroups in $C_\infty(\mathbb{R}^n)$, the space of continuous functions vanishing at infinity, the positive maximum principle for the generator is equivalent to the submarkovian property of the semigroup, i.e. that the semigroup is Feller.

Thus the positive maximum principle is a natural property of generators of Markov processes and the affirmative answer to the above conjecture is given by the theorem of Ph. Courrège (see Theorem 2.16), which states that a linear operator defined on $C_0^\infty(\mathbb{R}^n)$ and which satisfies the positive maximum principle is a pseudo differential operator $-p(x,D)$ with a negative definite

symbol $p(x,\xi)$.
Courrège also gave another equivalent representation of this class of operators as so-called Lévy-type operators. These are operators consisting of a second order diffusion operator plus an additional non-local integro-differential part, which is characterized by a kernel of measures, the Lévy-kernel of the operator, see Theorem 2.12. This second representation is widely used in probability theory, since it does not involve Fourier transform and gives an immediate interpretation of the coefficients of the operator in terms of the characteristics of the associated process. On the other hand it turns out that using the representation as a pseudo differential operator has certain advantages. It clearly provides a rich L^2-theory and L^2-estimates obtained in this way yield additional informations for an associated process and give a link to the theory of Dirichlet forms. Moreover note that as in the case of the characteristic exponent for a Lévy-process also the negative definite symbol has a natural probabilistic interpretation in terms of the associated process, see [45].

Therefore our starting point will be a negative definite symbol $p(x,\xi)$. Our primary interest is the question whether there exists a $\mathbb{R}^n$-valued Markov process or a semigroup which is associated to the corresponding pseudo differential operator $-p(x,D)$. Moreover, since all information on the process must be contained in the symbol, another complex of problems is related to the question how these information can be read from properties of the symbol. We will consider questions of this type in the last two chapters. Let us also mention in this context related results of R.L. Schilling concerning the path behaviour of the processs, see [77], [78], and in particular the monograph [44] by N. Jacob, which presents a good summary of all topic in this subject.

Let us emphasize from the very beginning that we are working with in general non-local operators. Therefore an associated process is a jump process and we have to replace the continuous path space of a diffusion process by the space of all càdlàg paths, i.e. paths which are right continuous and have left limits. Note also that we are in particular interested in the case that the generator has no leading second order term. This means that the non-local part is the dominating part itself and it cannot be treat as a perturbation of a diffusion part. Moreover it is important to remark that in general continuous negative definite functions are not differentiable, as for example the characteristic exponent $\xi \mapsto |\xi|^\alpha$ of the symmetric α-stable process, $0 < \alpha \leq 2$, shows. Thus a negative definite symbol $p(x,\xi)$ in general is not differentiable with respect to ξ. But standard symbol classes of pseudo differential operators like the Hörmander class $S^m_{\varrho,\delta}$ require a C^∞-behaviour of the symbol. Therefore the corresponding calculus for pseudo differential operators is not applicable and we have to use other techniques.

Since the translation invariant case of Lévy-processes is well-understood and described by the characteristic exponent, i.e. a continuous negative definite function, our general philosophy will be the following. We fix a continuous negative definite function ψ as a reference function and consider negative definite symbols $p(x,\xi)$ which satisfy estimates in terms of the reference function. Then $-p(x,D)$ is regarded as an operator with "variable" coefficents which is comparable to the generator $-\psi(D)$ of the associated Lévy process, just as a diffusion operator is compared to the Laplace operator. The Lévy-process then can be regarded as a reference process for a process associated to $-p(x,D)$. An essential idea for the following therefore will be to model our analysis in terms of this reference function ψ.

In this connection we not necessarily aim at the most general behaviour of the symbol $p(x, \xi)$ with respect to x, since this is simply a question how regular the "coefficients" are. But we want to be as general as possible concerning the behaviour of the symbol with respect to ξ, since this behaviour determines the type of operator under consideration. In particular we make no homogeneity assumptions.

We next give a summary of the subsequent chapters with a particular emphasis on the results obtained by the author:

In Chapter 2 first some properties of negative definite functions are provided. In particular we give a generalized Peetre-type inequality, Lemma 2.6, in terms of a negative definite function, which will be very important for our analysis in the following.

Section 2.2 is devoted to the result of Courrège. We give a new proof of the Lévy-Khinchin formula (Theorem 2.13). This proof is based on a characterization of functionals that satisfy the positive maximum principle (Proposition 2.10), which is taken from an unpublished manuscript of F. Hirsch. We then easily deduce both representations of an operator satisfying the positive maximum principle as a Lévy-type operator as well as a pseudo differential operator (Theorem 2.16).

After this preparatory chapter we turn in Chapter 3 to the first construction of a process associated to a pseudo differential operator using the martingale problem. Our main result (Theorem 3.15) gives the existence of a solution of the martingale problem for a pseudo differential operator with a continuous negative definite symbol $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$, $p(x, 0) = 0$, such that

$$|p(x, \xi)| \leq c(1 + |\xi|^2).$$

A natural interpretation of this condition is that the operator has bounded "coefficients". This result improves the result obtained in the author's paper [32], where only real-valued symbols are treated. It includes the existence results of Komatsu [51] and Stroock [81], which are formulated for Lévy-type operators. It is more general since it is also applicable in cases of variable order where the coefficients of the Lévy-type operator may become discontinuous, but the symbol remains continuous. This will be important in Chapter 7, where operators of this type are considered. For the proof the symbol is first decomposed by a convolution-like operation (Theorem 3.12) into a small part which is treated as a perturbation and a part for which the corresponding operator can be extended to functions on the one-point compactification of $\mathbb{R}^n$. For this part the martingale problem is solved by an approximation argument and a tightness result for solutions of martingale problems, which is derived in Section 3.1 and which in the case of jump processes relies on a result of Th.G Kurtz [56].

From Chapter 4 on we restrict to the case of real-valued symbols. As already mentioned we now fix a continuous negative definite reference function $\psi : \mathbb{R}^n \to \mathbb{R}$. We define the function $\lambda(x) = (1 + \psi(\xi))^{1/2}$, which is more convenient to state assumptions on a symbol. We want to apply modified Hilbert space methods to investigate an operator $-p(x, D)$. For this purpose we introduce in Section 4.1 an appropriate scale of anisotropic Sobolev spaces $H^{s,\lambda}(\mathbb{R}^n)$, $s \in \mathbb{R}$, which are defined in terms of the reference function. The main result of this section are certain commutator estimates (Theorems 4.3, 4.4) for pseudo differential operators with negative definite symbol. These results, which are improvements of the commutator estimates obtained in [29], show that also in the case of negative definite symbols the commutator has an order reducing property. This is very useful, because it allows to treat the effect of the x-dependence of a symbol as a lower order perturbation.

Using these results we decribe in Section 4.2 the approach of N. Jacob [43] to generators of Feller semigroups. Here it is assumed that the continuous negative definite symbol has a decomposition

$$p(x, \xi) = p_1(\xi) + p_2(x, \xi),$$

where $p_1(\xi)$ is bounded from above and below in terms of the reference function and $p_2(x, \xi)$ is sufficiently smooth with respect to x and also satisfies upper bounds with respect to the reference function. If these bounds are sufficiently small, the operator $-p_2(x, D)$ is a small perturbation of the generator $-p_1(D)$ of a Lévy-process and the equation

$$(p(x, D) + \tau)u = f$$

is solvable by modified Hilbert space methods for $\tau \geq 0$ sufficiently large. This, a Sobolev type embedding and the positive maximum principle then show via the Hille-Yosida thorem that $-p(x, D)$ generates a Feller semigroup and therefore a Markov process.

In Chapter 5 the assumptions on the symbol are reduced using the martingale problem. For this purpose we prove well-posedness of the martingale problem, i.e. there is a unique solution of the martingale problem for any initial distribution. These results are first proven in the author's article [32]. We obtain in Theorem 5.7 that the martingale problem is well-posed for a continuous negative definite symbol $p(x, \xi)$ which is sufficiently smooth with respect to x, which together with the derivatives satisfies an upper bound with respect to the reference function and which also admits a lower bound in terms of the reference function. In contrast to the results of Section 4.2 here no smallness assumptions of the upper bound is assumed and also merely a local lower ellipticity estimate is supposed.

The improvement is mainly due to the localization procedure for solutions of the martingale problem, which is discussed in Section 5.1. In particular we give in Therorem 5.3 a reformulation of this technique in terms of the symbols of the operators. Therefore it is enough to prove well-posedness for a localized symbol, i.e. a symbol which coincides on a small ball with a given continuous negative definite symbol and which is independent of x outside some neighbourhood of this ball. By a refined decomposition of this symbol using a Taylor expansion we can prove in Section 5.3 that the corresponding pseudo differential operator generates a strongly continuous semigroup in an anisotropic Sobolev space $H^{s,\lambda}(\mathbb{R}^n)$. This yields well-posedness of the martingale problem, since in Section 5.2 it is shown that well-posedness is implied by the solvability of the corresponding Cauchy problem in $H^{s,\lambda}(\mathbb{R}^n)$, which can be done using the semigroup.

Note that there are a lot of recent results concerning well-posedness of the martingale problem for Lévy-type operators. But often they either assume a dominating second order elliptic term, see Komatsu [51], Stroock [81] or a particular structure of the operators as for example perturbations of generators of α-stable processes in Komatsu [52],[53]. More general examples

are included in the paper [70] of Negoro and Tsuchiya, but they need a strong integrability assumption for the Lévy kernel, which is only satisfied by operators of order less than one. In the one-dimensional situation Bass [1] proves well-posedness for the martingale problem for the generator of stable-like processes, i.e. operators of variable order. The advantage of Theorem 5.7 however lies in the fact that in gives well-posedness for a general (real-valued) continuous negative definite reference function without assuming a particular structure.
Let us also mention that a solution of the martingale problem also can be obtained as a solution of a stochastic differential equation of jump type as introduced by Skorohod [80], see also Lepeltier, Marchal [61], but the relation between the coefficients of the stochastic differential equation and a Lévy-type operator or a pseudo differential operator is merely measurable and therefore only useful if a priori additional informations on the structure of the operator are known, see Tsuchiya [86].

Once well-posedness of the martingale problem is established it is well-known that the family of processes given by the unique solutions of the martingale problem starting at the points $x \in \mathbb{R}^n$ define a strong Markov process. Moreover it is easy to see, that the solution of the martingale problem depends on the initial distribution in a continuous way. But it is important to note that in the case of jump processes this does not immediately imply that the transition semigroup maps continuous functions to continuous functions, since the one-dimensional projections are not continuous in contrary to the continuous paths case. Nevertheless under a mild additional assumption on the symbol we prove in Section 5.4 that the semigroup defined by the Markov process is even a Feller semigroup. In particular it leaves the space of continuous functions vanishing at infinity invariant (see Theorem 5.23). These results are taken from the article [33].

In Chapter 6 we establish a symbolic calculus which is suitable for pseudo differential operators with negative definite symbol as it is developed in the articles [34] and [36] of the author. A symbolic calculus is very useful since it allows to carry out intuitive ideas in a justified framework. Moreover it supplies a rich L^2-theory and therefore yields L^2-estimates as in the approach of Chapter 4, but witout the restriction to consider only small perturbations of operators with x-independent symbol. Now continuous negative definite symbols $p(x, \xi)$ in general are not differentiable with respect to ξ and also in the differentiable case the derivatives do not satisfy the requirements of standard symbol classes for pseudo differential operators. The idea to overcome the differentiability problems is to restrict to symbols having Lévy-kernels with bounded support. The symbols then turn out to be infinitely differentiable with respect to ξ and the derivatives satisfy estimates that remind of the behaviour of standard symbol classes. Basing on these considerations in Section 6.2 we define symbol classes $\mathrm{S}_{\varrho}^{m,\lambda}$ and $\mathrm{S}_{0}^{m,\lambda}$ appropriate for continuous negative definite symbols in terms of a continuous negative definite reference function and we discuss some typical examples. In Section 6.3 a corresponding symbolic calculus is established. We show that the associated operators form an algebra similar to the standard case, i.e. the composition of two operators and the formally adjoint are again pseudo differential operators (Corollaries 6.12, 6.13) and the corresponding symbols have expansion formulas, which in highest order are the product and the complex conjugate of the given symbols (Corollary 6.18). Moreover the order $m \in \mathbb{R}$ of a symbol in $\mathrm{S}_{\varrho}^{m,\lambda}$ has a natural interpretation in terms of mapping properties of the corresponding operators between anisotropic Sobolev spaces $H^{s,\lambda}(\mathbb{R}^n)$ (Theorem 6.14).

Furthermore as a more refined technique for pseudo differential operators also the Friedrichs

symmetrization is introduced for the symbol class $\mathrm{S}_\varrho^{m,\lambda}$, which assigns to every operator with real-valued symbol a symmetric operator which coincides with the original one up to an lower order perturbation, see Section 6.4.
In this way an appropriate symbolic calculus is associated to every given continuous negative definite reference function. In Section 6.5 the results are applied to a situation which is elliptic in the sense that the symbol satisfies a lower estimate in terms of the reference function. In particular for continuous negative definite symbols in class $\mathrm{S}_\varrho^{2,\lambda}$ we obtain examples of generators of Feller semigroups (Theorem 6.29).

Starting with an arbitrary continuous negative definite symbol the idea is to decompose the corresponding Lévy-kernel into a part supported in a bounded neighbourhood of the origin, which is treated by the calculus, and a remainder part consisting only of finite measures. The remainder part therefore typically can be considered as a perturbation. In Section 6.6 the results of [37] are presented, where in particular the question is investigated whether the remainder part is a bounded operator in $L^2(\mathbb{R}^n)$ and in the space $C_\infty(\mathbb{R}^n)$ of continuous functions vanishing at infinity. Note that the class of generators of strongly continuous L^2-semigroups and Feller semigroups is stable under this type of perturbation. It turn out that the equi-continuity of the symbol $p(x,\xi)$ at $\xi = 0$ with respect to x is equivalent to a tightness property of the Lévy-kernel and in particular implies that the remainder part is a bounded perturbation in $C_\infty(\mathbb{R}^n)$ (Theorems 6.31 and 6.33). We also show that in typical examples the remainder part also defines a small perturbation in an L^2-sense.

The symbolic calculus is not restricted to elliptic situations. In Chapter 7 we study the explicitly non-elliptic situation of symbols of variable order. These results were obtained in the article [35]. For a continuous negative definite symbol $s \in \mathrm{S}_\varrho^{2,\lambda}$, which also satisfies a lower ellipticity bound in terms of the reference function we consider the continuous negative definite symbol

$$p(x,\xi) = s(x,\xi)^{m(x)}$$

of variable order with a function $m : \mathbb{R}^n \to (0,1]$. We assume that m has bounded derivatives, satisfies an oscillation bound and is strictly bounded away from 0. We now apply typical techniques for pseudo differential operators in the framework of the calculus of $\mathrm{S}_\varrho^{m,\lambda}$ to study the corresponding operator. The Friedrichs symmetrization yields a sharp Garding inequality, that is a lower bound for the associated bilinear form in terms of a lower order norm (Theorem 7.7), which reflects the non-ellipticity of the operator. From this we deduce that the bilinear form is a closed coercive form on a domain which is continuously embedded between anisotropic Sobolev spaces $H^{s,\lambda}(\mathbb{R}^n)$. This implies weak solvability of the corresponding equation (Theorem 7.2). To get more regularity of the solutions we use in Section 7.3 the symbolic calculus to construct a parametrix, that is an inverse modulo a smoothing operator. These results finally imply by the Hille-Yosida theorem that $-p(x,D)$ extends to the generator of a Feller semigroup (Theorem 7.1). In Section 7.4 we show that the localzation technique for the martingale problem enables weaker assumptions for the function $m(x)$ and we get an associated Feller semigroup for smooth m without oscillation bound and without strict lower bound (Thereom 7.10.)

The results obtained so far yield not only examples of generators of Feller semigroups, but the representation as a pseudo differential operator also implies L^2-estimates in terms of the norms in anisotropic Sobolev spaces for the operator as well as for the associated bilinear form.

In Chapter 8 we use these estimates as a starting point and show that they have important implications for the corresponding process and semigroup. First we see that the operator also generates a strongly continuous semigroup in $L^2(\mathbb{R}^n)$ and moreover the bilinear form defines a regular semi-Dirichlet form, see [66], [65]. But moreover we have an explicit knowledge on the domain of the form in terms of spaces $H^{s,\lambda}(\mathbb{R}^n)$. This implies a Sobolev inequality for the semi-Dirichlet form. It is well-known that Sobolev inequalities for Dirichlet forms can be used to get hypercontractivity bounds for the norms $\|T_t\|_{L^p \to L^q}$ of the associated semigroup (T_t). But these results are not applicable, since the operator is nor symmetric neither the adjoint operator generates a submarkovian semigroup in general. We therefore use an approach via the dual semigroup, show that it defines a strongly continuous contraction semigroup on $L^1(\mathbb{R}^n)$ and prove a hypercontractivity estimate using the original approach of Nash (Theorem 8.7). Moreover, this result is used to extend a result, which was obtained jointly with N. Jacob in [40], and to prove that the semigroups even have the strong Feller property.

In Chapter 9 we consider symbols which do not satisfy an upper bound $p(x,\xi) \leq c(1+|\xi|^2)$ uniformly with respect to x. We discuss the question under what conditions on the symbol in the “unbounded coefficient” case nevertheless there exists an associated non-exploding process in $\mathbb{R}^n$. We use an approach via the martingale problem. It turns out that the growth of the symbol with respect to x as $|x| \to \infty$ is admissible if it is compensated by the decay of the reference function as $|\xi| \to 0$, see Theorem 9.4. In particular for the reference function $|\xi|^\alpha$ of the symmetric α-stable process the maximal growth with respect to x is given by $|x|^\alpha$, that is

$$p(x,\xi) \leq c\,|x|^\alpha \cdot |\xi|^\alpha \qquad \text{for all } \ |x| \geq 1$$

implies non-explosion. For $\alpha = 2$ we obtain the well-known quadratic growth condition for the second order coefficients of a diffusion operator.
The result is combined with the uniqueness results for the martingale problem of Chapter 5 and well-posedness is proven also in the case of not uniformly bounded symbols (Theorem 9.5).

I thank the German Science Foundation DFG for the financial support I obtained during the last years as a holder of a Habilitanden-scholarship Ho 1617/2-x as well as in the DFG-project Ja 522/3-x. Moreover I am grateful for the financial support by SFB 343 “Diskrete Strukturen in der Mathematik” and for the assistance of the Faculty of Mathematics at the University of Bielefeld.
Among many others I want to thank Dr. R.L. Schilling for many stimulating discussions and useful remarks.
I am indebted to Prof. N. Jacob who over the past years was a constant source of encouragement and motivation for my work.
I wish to thank Prof. M. Röckner for the support I received during the last years, for many valuable suggestions and advice and the fruitful working conditions I found in his group.

Notations

The notations we use are mostly standard. For the readers' convenience here we summerize some important definitions. We usually work on the n-dimensional euclidian space $\mathbb{R}^n$, its one-point-compactification is denoted by $\overline{\mathbb{R}^n} = \mathbb{R}^n \cup \{\Delta\}$.

Function spaces

$C(E)$	continuous functions on a metric space E
$C_b(E)$	bounded continuous functions
$C_0(E)$	continuous functions with compact support
$B(E)$	bounded measurable functions
$M(E)$	real-valued maps on E
$C_\infty(\mathbb{R}^n)$	continuous functions on $\mathbb{R}^n$vanishing at infinity
$C^k(\Omega)$	k times continuously differentiable functions on Ω
$C^\infty(\Omega)$	arbitrarily often differentiable functions on Ω
$C_0^\infty(\Omega)$	testfunctions on Ω
$\mathcal{S}(\mathbb{R}^n)$	Schwartz space of rapidly decreasing testfunctions
$\mathcal{D}'(\Omega)$	distributions on Ω
$\mathcal{D}^{(k)'}(\Omega)$	distributions of order k on Ω
$\mathcal{S}'(\mathbb{R}^n)$	tempered distributions
$L^p(\mathbb{R}^n)$	L^p-space with respect to the Lebesgue mesure
$H^s(\mathbb{R}^n)$	L^2-Sobolev space of order s
$H^{s,\lambda}(\mathbb{R}^n)$	anisotropic Sobolev space, see Section 4.1

The function spaces are always assumed to consist of real-valued functions, except for Chapter 6, where for complex-valued function spaces this is remarked separately using the notation like $C(\mathbb{R}^n, \mathbb{C})$.
By $\wedge$ and $\vee$ we denote minimum and maximum, u^+ is the positive part of a real-valued function.

Norms

$\|\|\cdot\|\|_{L^p}$	$1 \leq p \leq \infty$, L^P–norm with respect to the Lebesgue measue
$\|\|\cdot\|\|_0$	L^2–norm,
$\|\|\cdot\|\|_\infty$	L^∞–norm or sup–norm,
$\|\|\cdot\|\|_s$	$s \in \mathbb{R}$, standard Sobolev space norm,
$(\cdot,\cdot)_0$	scalar product in L^2,
$\|\cdot\|_{s,\lambda}$	norm in $H^{s,\lambda}(\mathbb{R}^n)$
$(\cdot,\cdot)_{s,\lambda}$	inner product in $H^{s,\lambda}(\mathbb{R}^n)$

$\mathcal{M}_1(\cdot)$ and $\mathcal{M}_b(\cdot)$ denote spaces of probability measures and signed measures of bounded variation, respectively.
The Fourier transform is defined by the convention

$$\hat{u}(\xi) = \int_{\mathbb{R}^n} e^{-i(x,\xi)} u(x)\, dx,$$

then the inverse Fourier transform is given by

$$u(x) = \int_{\mathbb{R}^n} e^{i(x,\xi)} \hat{u}(x)\, đ\xi,$$

where we use the notation $đ\xi = (2\pi)^{-n}\, d\xi$.
In connection with the Fourier transform is helpful to use for partial derivatives the notation

$$D = (D_1, \ldots, D_n) = (-i\partial_{x_1}, \ldots, -i\partial_{x_n}).$$

We sometimes use the notation $\langle \xi \rangle = (1 + |\xi|^2)^{1/2}$.
For a multiindex $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$, let $|\alpha| = \alpha_1 + \ldots + \alpha_n$, $\alpha! = \alpha_1! \cdot \ldots \cdot \alpha_n!$ and $\xi^\alpha = \xi_1^{\alpha_1} \cdot \ldots \cdot \xi_n^{\alpha_n}$ for $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$.
Moreover we use the notation $(A, D(A))$ for a linear operator A with domain $D(A)$.

Finally by c we denote a nonnegative constant, the values of which may be different at every occurrence.

Chapter 2

Negative definite symbols

2.1 Negative definite functions

In this section we summerize properties of negative definite functions. Negative definite functions are investigated in great detail in [4] by Ch. Berg and G. Forst. For all proofs concerning this subject we refer to this monograph.

Definition 2.1. *A function* $\psi : \mathbb{R}^n \to \mathbb{C}$ *is called a* **negative definite** *function if for all* $m \in \mathbb{N}$ *and for any choice of* $\xi^j \in \mathbb{R}^n,\ 1 \leq j \leq m$*, the matrix*

$$(\psi(\xi^i) + \overline{\psi(\xi^j)} - \psi(\xi^i - \xi^j))_{i,j=1,\dots,m}$$

is non-negative Hermitian, i.e. for all $c_1, \dots, c_m \in \mathbb{C}$

$$\sum_{i,j=1}^{m} (\psi(\xi^i) + \overline{\psi(\xi^j)} - \psi(\xi^i - \xi^j))\ c_i\overline{c_j} \geq 0. \tag{2.1}$$

We first give some elementary properties of negative definite functions which follow easily from the definition.

Proposition 2.2.

(i) The set of negative definite functions is a convex cone containing the non-negative constants.

(ii) For a negative definite function ψ *we have*

$$\psi(-\xi) = \overline{\psi(\xi)}$$

and

$$Re\,\psi(\xi) \geq \psi(0) \geq 0.$$

In particular, a real-valued negative definite function is non-negative and even.

(iii) For a negative definite function ψ *also* $Re\,\psi$ *and* $\overline{\psi}$ *are negative definite.*

(iv) The continuous negative definite functions form a convex cone which is closed under locally uniform convergence.

There are some other equivalent definitions of negative definite functions (see [4], 7.4, 7.8):

Theorem 2.3. *For a function $\psi : \mathbb{R}^n \to \mathbb{C}$ the following are equivalent:*

(i) ψ is negative definite.

(ii) $\psi(0) \geq 0$, $\psi(-\xi) = \overline{\psi(\xi)}$ and for all $m \in \mathbb{N}$ and $\xi_1, \ldots, \xi_m \in \mathbb{R}^n$ we have

$$\sum_{i,j=1}^{m} \psi(\xi^i - \xi^j)) c_i \overline{c_j} \leq 0$$

for all $c_1, \ldots, c_m \in \mathbb{C}$ satisfying $\sum_{j=1}^m c_j = 0$.

(iii) $\psi(0) \geq 0$ and the function $e^{-t\psi}$ is a positive definite function for all $t \geq 0$.

Recall that a function $\varphi : \mathbb{R}^n \to \mathbb{C}$ is called positive definite if for all $m \in \mathbb{N}$ and $\xi_1, \ldots, \xi_m \in \mathbb{R}^n$ the matrix

$$(\varphi(\xi^i - \xi^j))_{i,j=1,\ldots,m}$$

is non-negative Hermitian.

The equivalence of (i) and (iii) in Theorem 2.3 is known as the theorem of Schoenberg (see [79]). By the famous theorem of Bochner [7] the set of continuous positive definite functions on $\mathbb{R}^n$ coincides with set of Fourier transforms of bounded measures on $\mathbb{R}^n$. Thus, in the case of a continuous negative definite function ψ by condition (iii) there exists a family $(\mu_t)_{t\geq 0}$ of bounded measures on $\mathbb{R}^n$ with Fourier transforms

$$\hat{\mu}_t(\xi) = \int_{\mathbb{R}^n} e^{-i(x,\xi)}\, \mu_t(dx) = e^{-t\psi(\xi)} \tag{2.2}$$

and total mass

$$\mu_t(\mathbb{R}^n) = \hat{\mu}_t(0) = e^{-t\psi(0)} \leq 1.$$

Moreover, the functional equation of the exponential function translates by the convolution theorem immediately into the property

$$\mu_s * \mu_t = \mu_{s+t} \qquad \text{for all } s, t \geq 0. \tag{2.3}$$

Therefore $(\mu_t)_{t\geq 0}$ is a convolution semigroup of (sub-)probability measures on $\mathbb{R}^n$. Moreover, because $e^{-t\psi} \underset{t\to 0}{\longrightarrow} 1$, we see that μ_t converges weakly to the point mass ε_0 at the origin as $t \to 0$ and $(\mu_t)_{t\geq 0}$ is continuous with respect to the weak topology. Conversely, given a weakly continuous convolution semigroup $(\mu_t)_{t\geq 0}$ on $\mathbb{R}^n$, (2.2) consistently defines a function $\psi : \mathbb{R}^n \to \mathbb{C}$, which is negative definite by the theorem of Schoenberg and continuous since Fourier transforms of bounded measures are continuous (cf. [4], 8.3). Thus we have seen

Theorem 2.4. *The set of weakly continuous convolution semigroups $(\mu_t)_{t\geq 0}$ of sub-probability measures on $\mathbb{R}^n$ is in one-to-one correspondence with the set of continuous negative definite functions $\psi : \mathbb{R}^n \to \mathbb{C}$. The correpondence is given by (2.2).*

We summerize some typical examples of continuous negative definite functions and the corresponding convolution semigroups on $\mathbb{R}^n$ in the following table.

$\psi(\xi)$	μ_t	
$\lvert\xi\rvert^2$	$\frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{\lvert x\rvert^2}{4t}\right) \cdot dx$	(Brownian semigroup)
$i(b,\xi)$	$\varepsilon_{-tb}(dx)$	(shift semigroup)
$\lvert\xi\rvert^\alpha,\ 0<\alpha\le 2$	symmetric α-stable semigroup	
$\lvert\xi\rvert$	$\Gamma(\frac{n+1}{2})\frac{t}{[\pi(\lvert x\rvert^2+t^2)]^{\frac{n+1}{2}}} \cdot dx$	(Cauchy semigroup)
$1-e^{i(y,\xi)}$	$\sum_{k=0}^\infty e^{-t}\frac{t^k}{k!}\varepsilon_{ky}(dx)$	(Poisson semigroup with jumps of size $y\in\mathbb{R}^n$)

Note that obviously the convolution semigroup consists of probability measures if and only if $\psi(0)=0$.

In the following we will work with symbols of pseudo differential operators which are defined in terms of continuous negative definite functions. For that purpose estimates which are automatically fulfilled by continuous negative definite functions, for example for the growth behaviour at infinity, are very helpful and will play an important rôle. First we note

Proposition 2.5. *Let $\psi:\mathbb{R}^n\to\mathbb{C}$ be a negative definite function. Then*

$$\sqrt{|\psi(\xi+\eta)|}\le\sqrt{|\psi(\xi)|}+\sqrt{|\psi(\eta)|},\qquad \xi,\eta\in\mathbb{R}^n, \tag{2.4}$$

i.e. $\sqrt{|\psi|}$ *is subadditive.*

Proof: By definition of a negative definite function we have

$$0\le\det\begin{pmatrix}\psi(\xi)+\overline{\psi(\xi)}-\psi(0) & \psi(\xi)+\overline{\psi(-\eta)}-\psi(\xi+\eta)\\ \psi(-\eta)+\overline{\psi(\xi)}-\psi(-\xi-\eta) & \psi(-\eta)+\overline{\psi(-\eta)}-\psi(0)\end{pmatrix}$$

$$=(2\mathrm{Re}\,\psi(\xi)-\psi(0))(2\mathrm{Re}\,\psi(\eta)-\psi(0))-|\psi(\xi+\eta)-\psi(\xi)-\psi(\eta)|^2,$$

where we have used $\overline{\psi(\xi)}=\psi(-\xi)$. Thus, since $\psi(0)\ge 0$,

$$\begin{aligned}2\,|\psi(\xi)|\cdot 2\,|\psi(\eta)| &\ge (2\mathrm{Re}\,\psi(\xi)-\psi(0))(2\mathrm{Re}\,\psi(\eta)-\psi(0))\\ &\ge |\psi(\xi+\eta)-\psi(\xi)-\psi(\eta)|^2\end{aligned}$$

and therefore

$$\begin{aligned}|\psi(\xi+\eta)| &\le |\psi(\xi+\eta)-\psi(\xi)-\psi(\eta)|+|\psi(\xi)|+|\psi(\eta)|\\ &\le 2\sqrt{|\psi(\xi)|}\cdot\sqrt{|\psi(\eta)|}+|\psi(\xi)|+|\psi(\eta)|\\ &= (\sqrt{|\psi(\xi)|}+\sqrt{|\psi(\eta)|})^2.\end{aligned}$$

□

Because $\sqrt{|\psi|}$ is an even function, (2.4) also implies a triangle inequality from below:

$$\left|\sqrt{|\psi(\xi)|} - \sqrt{|\psi(\eta)|}\right| \le \sqrt{|\psi(\xi-\eta)|}, \qquad \xi, \eta \in \mathbb{R}^n. \tag{2.5}$$

Another elementary but important consequence is the validity of a generalized form of Peetre's inequality

$$\left(\frac{1+|\xi|^2}{1+|\eta|^2}\right)^s \le 2^{|s|}(1+|\xi-\eta|^2)^{|s|}, \qquad s \in \mathbb{R},\ \xi, \eta \in \mathbb{R}^n. \tag{2.6}$$

Lemma 2.6 *(Peetre type inequality). Let $\psi : \mathbb{R}^n \to \mathbb{C}$ be a negative definite function and $s \in \mathbb{R}$. Then*

$$\left(\frac{1+|\psi(\xi)|}{1+|\psi(\eta)|}\right)^s \le 2^{|s|}(1+|\psi(\xi-n)|)^{|s|}, \qquad \xi, \eta \in \mathbb{R}^n. \tag{2.7}$$

Proof: By (2.5) and (2.6) we see

$$\begin{aligned}
\left(\frac{1+|\psi(\xi)|}{1+|\psi(\eta)|}\right)^s = \left(\frac{1+\sqrt{|\psi(\xi)|}^2}{1+\sqrt{|\psi(\eta)|}^2}\right)^s &\le 2^{|s|}(1+\left(\sqrt{|\psi(\xi)|} - \sqrt{|\psi(\eta)|}\right)^2)^{|s|} \\
&\le 2^{|s|}(1+\sqrt{|\psi(\xi-\eta)|}^2)^{|s|} \\
&= 2^{|s|}(1+|\psi(\xi-\eta)|)^{|s|}.
\end{aligned}$$

□

Finally by the subadditivity of $\sqrt{|\psi|}$ we see that the growth of $\psi(\xi)$ for large values of ξ is controlled be its values in a neighbourhood of the origin:

Theorem 2.7. *Let $\psi : \mathbb{R}^n \to \mathbb{C}$ be continuous negative definite function. Then*

$$|\psi(\xi)| \le c_\psi(1+|\xi|^2), \qquad \xi \in \mathbb{R}^n, \tag{2.8}$$

where $c_\psi = 2 \sup_{|\xi|\le 1} |\psi(\xi)|$.

Proof: By Proposition 2.5 we have for all $k \in \mathbb{N}$

$$|\psi(\xi)| = \sqrt{|\psi(\xi)|}^2 = \sqrt{\left|\psi\left(k\frac{\xi}{k}\right)\right|}^2 \le \left(k\sqrt{\left|\psi\left(\frac{\xi}{k}\right)\right|}\right)^2 = k^2\left|\psi\left(\frac{\xi}{k}\right)\right|.$$

For given $\xi \in \mathbb{R}^n$ choose $k_0 = \inf\{k \in \mathbb{N} : |\xi| \le k\}$. Then $k_0^2 \le 2(1+|\xi|^2)$ and

$$|\psi(\xi)| \le 2(1+|\xi|^2)\cdot\left|\psi\left(\frac{\xi}{k_0}\right)\right| \le 2 \sup_{|\eta|\le 1}|\psi(\eta)|\cdot(1+|\xi|^2).$$

□

In particular, the map $\xi \mapsto 1 - e^{i(x,\xi)} + \frac{i(x,\xi)}{1+|x|^2}$ defines for fixed $x \in \mathbb{R}^n$ a continuous negative definite function on $\mathbb{R}^n$. Thus by Theorem 2.7

$$\left|1 - e^{i(x,\xi)} + \frac{i(x,\xi)}{1+|x|^2}\right| \leq c_x(1+|\xi|^2)$$

with a constant

$$c_x = 2 \sup_{|\xi|\leq 1} \left|1 - e^{i(x,\xi)} + \frac{i(x,\xi)}{1+|x|^2}\right| = 2 \sup_{\substack{-|x|\leq z\leq |x| \\ z\in\mathbb{R}}} \left|1 - e^{iz} + \frac{iz}{1+|x|^2}\right|$$

depending on x. But for $-|x| \leq z \leq |x|$ we have

$$\left|1 - e^{iz} + \frac{iz}{1+|x|^2}\right| \leq 2 + \frac{|z|}{1+|x|^2} \leq 2 + \frac{|x|}{1+|x|^2} \leq \frac{5}{2}$$

as well as

$$\left|1 - e^{iz} + \frac{iz}{1+|x|^2}\right| \leq \left|1 - e^{iz} + iz\right| + |z| \frac{|x|^2}{1+|x|^2} \leq \frac{1}{2}|x|^2 + \frac{|x|^3}{1+|x|^2} \leq |x|^2,$$

where we have used Taylor formula. Therefore

$$c_x \leq 2\,(\frac{5}{2} \wedge |x|^2) \leq 7\frac{|x|^2}{1+|x|^2}$$

and we have shown

Lemma 2.8. *For all $x \in \mathbb{R}^n$ and $\xi \in \mathbb{R}^n$*

$$\left|1 - e^{i(x,\xi)} + \frac{i(x,\xi)}{1+|x|^2}\right| \leq 7 \cdot \frac{|x|^2}{1+|x|^2} \cdot (1+|\xi|^2). \tag{2.9}$$

2.2 The Lévy–Khinchin formula and the positive maximum principle

We have seen in the introduction that the positivity preserving property of a semigroup is reflected by the positive maximum principle on the level of the generator. The main concern of this section is the result of Ph. Courrége [13] that characterizes those operators as pseudo differential operators having symbols that are defined in terms of continuous negative definite functions. Since this result is fundamental for the sequent we shall give a complete proof. This proof moreover admits an easy approach to the closely related Lévy–Khinchin formula and the representation of the operator as a Lévy-type operator. The proof relies on a result (see Proposition 2.10 below) concerning the representation of linear functional satisfying the positive maximum principle. This result seems to be obtained first by J.P Roth and basing on his ideas a proof was given by F. Hirsch (unpublished manuscript, but see also [27]. p.115). We remark that there are several other proofs of the Lévy–Khinchin formula, which employ different techniques, let us mention among others the proofs given by Courrège [12], Rogalski [75], Harzallah [24] and a recent proof by Jacob and Schilling [48].

Denote by $M(\mathbb{R}^n)$ the space of all real-valued functions on $\mathbb{R}^n$ and let $A : D(A) \to M(\mathbb{R}^n)$ be a linear operator defined on a subspace $D(A)$ of $M(\mathbb{R}^n)$. We are mainly interested in the case $D(A) = C_0^\infty(\mathbb{R}^n)$.
We say that A satisfies the **positive maximum principle** on $D(A)$ if for all $\varphi \in D(A)$ such that $\sup_{x\in\mathbb{R}^n} \varphi(x) = \varphi(x_0) \geq 0$ for some $x_0 \in \mathbb{R}^n$ (depending on φ) we have

$$A\varphi(x_0) \leq 0.$$

Moreover we define for a linear functional $T : C_0^\infty(\mathbb{R}^n) \to \mathbb{R}$:

T is called **almost positive** if $\varphi \in C_0^\infty(\mathbb{R}^n)$, $\varphi \geq 0$, $\varphi(0) = 0$ implies $T\varphi \geq 0$,

T satisfies the **positive maximum principle** if $\varphi \in C_0^\infty(\mathbb{R}^n)$, $\sup_{x\in\mathbb{R}^n} \varphi(x) = \varphi(0) \geq 0$ implies $T\varphi \leq 0$.

Note that the positive maximum principle for T implies that T is almost positive. In fact, for $\varphi \in C_0^\infty(\mathbb{R}^n)$, $\varphi \geq 0$, $\varphi(0) = 0$ it follows $-\varphi \in C_0^\infty(\mathbb{R}^n)$ has a nonnegative maximum in 0, hence by the positive maximum principle $T\varphi = -T(-\varphi) \geq 0$.

Clearly the positive maximum principle for the operator A is a pointwise statement, i.e. $A : C_0^\infty(\mathbb{R}^n) \to M(\mathbb{R}^n)$ satisfies the positive maximum principle if and only if for all $x \in \mathbb{R}^n$ the functionals $A_x : C_0^\infty(\mathbb{R}^n) \to \mathbb{R}$ fulfill the positive maximum principle, where A_x is defined by

$$A_x\varphi := [A(\varphi(\cdot - x))](x).$$

We therefore first concentrate on the positive maximum principle for functionals. To begin with we note that such functional satisfies certain continuity properties by itself.

Lemma 2.9. *Let T be an almost positive functional. Then $T \in \mathcal{D}^{(2)'}(\mathbb{R}^n)$, i.e. is a distribution of order two, this means for all compact sets $K \subset \mathbb{R}^n$ there is a constant c_K such that*

$$|T\varphi| \leq c_K \sum_{\substack{\alpha\in\mathbb{N}_0^n \\ |\alpha|\leq 2}} \sup_{x\in K} |\partial_x^\alpha \varphi(x)| \qquad \textit{for all } \varphi \in C_0^\infty(K).$$

Proof: Let $\varphi \in C_0^\infty(K)$ and $M = \sum_{\substack{\alpha\in\mathbb{N}_0^n \\ |\alpha|\leq 2}} \sup_{x\in K} |\partial_x^\alpha \varphi(x)|$ and define

$$\tilde{\varphi}(x) := \varphi(x) - \varphi(0) \cdot \chi(x) - \sum_{i=1}^n \partial_{x_i}\varphi(0)\chi(x) \cdot x_i, \tag{2.10}$$

where $\chi \in C_0^\infty(\mathbb{R}^n)$ is chosen such that $0 \leq \chi \leq 1$ and $\chi = 1$ in a neighbourhood of 0. Then

$$\begin{aligned} \left|\partial^2_{x_j x_k}\tilde{\varphi}(x)\right| &\leq \left|\partial^2_{x_j x_k}\varphi(x)\right| + |\varphi(0)| \left|\partial^2_{x_j x_k}\chi(x)\right| + \sum_{i=1}^n |\partial_{x_i}\varphi(0)| \left|\partial^2_{x_j x_k}(x_i \cdot \chi(x))\right| \\ &\leq c_\chi \cdot M \qquad \text{for all } x \in \mathbb{R}^n \end{aligned}$$

and $\tilde{\varphi}(0) = \partial_{x_i}\tilde{\varphi}(0) = 0$, $i = 1, \ldots, n$. Therefore $\tilde{\varphi} \in C_0^\infty(\mathbb{R}^n)$ and by Taylor formula

$$|\tilde{\varphi}(x)| \leq \frac{1}{2}\sum_{j,k=1}^{n} \sup\left|\partial^2_{x_j,x_k}\tilde{\varphi}\right| \cdot |x_j x_k| \leq \frac{n}{2} c_\chi \cdot M \cdot |x|^2 \qquad \text{for all } x \in \mathbb{R}^n.$$

Now choose $\Psi_K \in C_0^\infty(\mathbb{R}^n)$ nonnegative such that $\Psi_K(x) = |x|^2$ in $K \cup \operatorname{supp}\chi$. Then

$$\frac{n}{2} c_\chi \cdot M \cdot \Psi_K \pm \tilde{\varphi} \geq 0$$

and vanishes in 0. Thus, since T is almost positive,

$$T(\frac{n}{2} c_\chi \cdot M \cdot \Psi_K \pm \tilde{\varphi}) \geq 0,$$

i.e.

$$|T\tilde{\varphi}| \leq \frac{n}{2} c_\chi \cdot M \cdot T\Psi_K.$$

But by (2.10)

$$|T\varphi| \leq |T\tilde{\varphi}| + |\varphi(0) \cdot T\chi| + \sum_{i=1}^{n} |\partial_{x_i}\varphi(0) T(x_i\chi(x))| \leq c_K \cdot M.$$

□

Functionals satisfying the positive maximum principle are now characterized in the following way.

Proposition 2.10. *Let $T : C_0^\infty(\mathbb{R}^n) \to \mathbb{R}$ be a linear functional that satisfies the positive maximum principle. Then there are unique constants $a_{ij}, b_i, c \in \mathbb{R}$, $i, j = 1, \ldots, n$ and a unique Borel measure μ on $\mathbb{R}^n \setminus \{0\}$ such that*

(i) $(a_{ij})_{i,j=1,\ldots,n}$ *is a symmetric nonnegative definite matrix,*

(ii) $c \geq 0$,

(iii) $\int_{\mathbb{R}^n\setminus\{0\}} \frac{|x|^2}{1+|x|^2}\, \mu(dx) < \infty$

and we have for all $\varphi \in C_0^\infty(\mathbb{R}^n)$

$$\begin{aligned} T\varphi &= \sum_{i,j=1}^{n} a_{ij}\partial_{x_i}\partial_{x_j}\varphi(0) - \sum_{i=1}^{n} b_i\partial_{x_i}\varphi(0) - c \cdot \varphi(0) \\ &\quad + \int_{\mathbb{R}^n\setminus\{0\}} (\varphi(x) - \varphi(0) - \frac{(x, \nabla\varphi(0))}{1+|x|^2})\, \mu(dx). \end{aligned} \tag{2.11}$$

Proof: By Lemma 2.9 T is a distribution in $\mathcal{D}^{(2)\prime}(\mathbb{R}^n)$. It follows that $|\cdot|^2 \cdot T \in \mathcal{D}'(\mathbb{R}^n)$ is a positive distribution, because for all $\varphi \in C_0^\infty(\mathbb{R}^n)$, $\varphi \geq 0$ clearly $|x|^2 \cdot \varphi(x) \geq 0$ and vanishes in 0. Thus, since T is almost positive,

$$\langle |\cdot|^2 \cdot T, \varphi\rangle = \langle T, |\cdot|^2 \cdot \varphi\rangle \geq 0.$$

Therefore there is a Borel measure ν on $\mathbb{R}^n$ such that $|\cdot|^2 \cdot T = \nu$.
Let $\mu = \frac{1}{|x|^2} \cdot \nu|_{\mathbb{R}^n \setminus \{0\}}$. Then μ is a Borel measure on $\mathbb{R}^n \setminus \{0\}$ and we have

$$\mu = T|_{\mathbb{R}^n \setminus \{0\}} \qquad \text{in } \mathcal{D}'(\mathbb{R}^n \setminus \{0\}). \tag{2.12}$$

We show that (iii) holds. Clearly

$$\int_{0<|x|\leq 1} |x|^2 \; \mu(dx) = \nu(\{0 < |x| \leq 1\}) < \infty. \tag{2.13}$$

Moreover for $\varphi, g \in C_0^\infty(\mathbb{R}^n)$, $0 \leq \varphi, g \leq 1$, $\operatorname{supp}\varphi \subset \{|x| > 1\}$ and $\operatorname{supp} g \subset \{|x| \leq 1\}$ such that $g(0) = 1$ we have

$$\sup_{x \in \mathbb{R}^n} (g + \varphi)(x) = g(0) + \varphi(0) = 1.$$

Thus by the positive maximum principle $T(g + \varphi) \leq 0$, which yields

$$\int_{\mathbb{R}^n \setminus \{0\}} \varphi \, d\mu = \langle T, \varphi \rangle \leq -\langle T, g \rangle.$$

Taking the supremum over all possible choices of φ we see that $\mu\{|x| > 1\} < \infty$. This together with (2.13) gives (iii).

Next define for $\varphi \in C_0^\infty(\mathbb{R}^n)$ the linear functional

$$S(\varphi) = \int_{\mathbb{R}^n \setminus \{0\}} (\varphi(x) - \varphi(0) - \frac{(x, \nabla\varphi(0))}{1 + |x|^2}) \, \mu(dx). \tag{2.14}$$

The integrand in (2.14) is bounded and vanishes of second order in 0, so the integral is well-defined. Furthermore, S clearly is almost positive and hence $S \in \mathcal{D}^{(2)'}(\mathbb{R}^n)$ by Lemma 2.9. Therefore

$$P := T - S$$

is also a distribution of order two and moreover for $\varphi \in C_0^\infty(\mathbb{R}^n \setminus \{0\})$

$$S(\varphi) = \int_{\mathbb{R}^n \setminus \{0\}} \varphi(x) \, \mu(dx) = T\varphi, \tag{2.15}$$

that is $P(\varphi) = 0$ and consequently $\operatorname{supp} P \subset \{0\}$. But any distribution supported in the origin is a linear combination of derivatives of the point mass in 0, hence

$$P\varphi = \sum_{i,j=1}^{n} a_{ij} \partial_{x_i} \partial_{x_j} \varphi(0) - \sum_{i=1}^{n} b_i \partial_{x_i} \varphi(0) - c \cdot \varphi(0)$$

with $a_{ij} = a_{ji}$. Thus $T = P + S$ has the form as claimed in (2.11).
We check the properties (i) and (ii). Choose a sequence $(\varphi_k) \subset C_0^\infty(\mathbb{R}^n)$ such that $\varphi_k \geq 0$, $\varphi_k = 1$ in a neighbourhood of 0 and $\varphi_k \uparrow 1$ as $k \to \infty$. Then

$$S(\varphi_k) = \int_{\mathbb{R}^n \setminus \{0\}} (\varphi_k(x) - 1) \, \mu(dx) \underset{k \to \infty}{\longrightarrow} 0,$$

hence

$$T\varphi_k = P\varphi_k + S\varphi_k = -c\varphi_k(0) + S\varphi_k \underset{k\to\infty}{\longrightarrow} -c. \tag{2.16}$$

But $T\varphi_k \le 0$ by the positive maximum principle, which gives $c \ge 0$.

Furthermore, choose a sequence $(\chi_k) \subset C_0^\infty(\mathbb{R}^n)$, $0 \le \chi_k \le 1$, $\chi_k = 1$ in a neighbourhood of 0 such that supp$\chi_k \subset B_{\frac{1}{k}}(0)$. Since T is almost positive, we have for all $\xi = (\xi_1, \dots \xi_n) \in \mathbb{R}^n$

$$T(\chi_k \cdot \frac{1}{2}(\xi, \cdot)^2) \ge 0$$

and further

$$\begin{aligned} T(\chi_k \cdot \frac{1}{2}(\xi, \cdot)^2) &= P(\chi_k \cdot \frac{1}{2}(\xi, \cdot)^2) + S(\chi_k \cdot \frac{1}{2}(\xi, \cdot)^2) && (2.17)\\ &= \sum_{i,j=1}^n a_{ij}\xi_i\xi_j + \int_{\mathbb{R}^n\setminus\{0\}} \chi_k(x) \cdot \frac{1}{2}(\xi, x)^2)\, \mu(dx) \\ &\underset{k\to\infty}{\longrightarrow} \sum_{i,j=1}^n a_{ij}\xi_i\xi_j, && (2.18) \end{aligned}$$

that is $\sum_{i,j=1}^n a_{ij}\xi_i\xi_j \ge 0$.

It remains to prove the uniqueness of the representation. But by (2.16) and (2.17) the constant c and the symmetric matrix (a_{ij}) are uniquely determined. Moreover, for all $\varphi \in C_0^\infty(\mathbb{R}^n \setminus \{0\})$ by (2.15)

$$T\varphi = \int_{\mathbb{R}^n\setminus\{0\}} \varphi(x)\, \mu(dx),$$

which determines the measure μ. Consequently we can calculate $\sum_{i=1}^n b_i \partial_{x_i}\varphi(0)$ for all $\varphi \in C_0^\infty(\mathbb{R}^n)$, which also fixes b_i. □

Remark 2.11. *Note that conversely every functional of type (2.11) obviously satisfies the positive maximum principle. So (2.11) gives a complete and unique representation of linear functionals that fulfill the positive maximum principle.*

By the remark preceding Lemma 2.9, Proposition 2.10 immediately yields

Theorem 2.12. *Let $A : C_0^\infty(\mathbb{R}^n) \to M(\mathbb{R}^n)$ be a linear operator. Then A satisfies the positive maximum principle if and only if there exist, uniquely determined by A for every $x \in \mathbb{R}^n$,*

- *a nonnegative definite, real, symmetric matrix $(a_{ij}(x))_{i,j=1,\dots,n}$,*
- *a vector $b(x) = (b_1(x), \dots, b_n(x)) \in \mathbb{R}^n$,*
- *a constant $c(x) \ge 0$ and*
- *a Borel measure $\mu(x, dy)$ on $\mathbb{R}^n \setminus \{0\}$ satisfying $\int_{\mathbb{R}^n\setminus\{0\}} \frac{|y|^2}{1+|y|^2}\, \mu(x, dy) < \infty$*

such that

$$A\varphi(x) = \sum_{i,j=1}^{n} a_{ij}(x)\partial_{x_i}\partial_{x_j}\varphi(x) - \sum_{i=1}^{n} b_i(x)\partial_{x_i}\varphi(x) - c\cdot\varphi(x) + \int_{\mathbb{R}^n\setminus\{0\}} (\varphi(x+y) - \varphi(x) - \frac{(y,\nabla\varphi(x))}{1+|y|^2})\,\mu(x,dy). \tag{2.19}$$

Operators of type (2.19) are called **Lévy-type operators**, the family of measures $\mu(x,dy)$ the corresponding **Lévy-kernel**.

Proposition 2.10 also admits an easy proof of the **Lévy-Khinchin formula**:

Theorem 2.13. *Let $\psi : \mathbb{R}^n \to \mathbb{C}$ be a continuous negative definite function. Then there are a real, nonnegative definite, symmetric matrix $a = (a_{jk})_{j,k=1,\dots,n}$, a vector $b = (b_j)_{j=1,\dots,n} \in \mathbb{R}^n$, a constant $c \geq 0$ and a Borel measure μ on $\mathbb{R}^n \setminus \{0\}$, the so-called* **Lévy-measure***, satisfying $\int_{\mathbb{R}^n\setminus\{0\}} \frac{|x|^2}{1+|x|^2}\,\mu(dx) < \infty$ such that*

$$\psi(\xi) = \sum_{j,k=1}^{n} a_{jk}\xi_j\xi_k + i\sum_{j=1}^{n} b_j\xi_j + c + \int_{\mathbb{R}^n\setminus\{0\}} (1 - e^{i(x,\xi)} + \frac{i(x,\xi)}{1+|x|^2})\,\mu(dx). \tag{2.20}$$

Here a, b, c and μ are uniquely determined. Conversely, each such choice of a, b, c and μ defines by (2.20) a continuous negative definite function.

Proof: First note that for given a, b, c and μ as above (2.20) defines a continuous negative definite function as a superposition of continuous negative definite functions. Therefore let conversely $\psi : \mathbb{R}^n \to \mathbb{C}$ be a continuous negative definite function and let $(\mu_t)_{t\geq 0}$ be the convolution semigroup of sub-probability measures associated to the continuous negative definite function $\overline{\psi}$ by Theorem 2.4, i.e. μ_t is defined by its Fourier transform

$$\hat{\mu}_t(\xi) = e^{-t\overline{\psi}(\xi)}.$$

Let $T : C_0^\infty(\mathbb{R}^n) \to \mathbb{R}$ be the linear functional defined by

$$T\varphi = -\int_{\mathbb{R}^n} \psi(\xi)\cdot\hat{\varphi}(\xi)đ\xi. \tag{2.21}$$

By Theorem 2.7 $\psi(\xi)\cdot\hat{\varphi}(\xi)$ is integrable for all $\varphi \in C_0^\infty(\mathbb{R}^n)$ and T is well-defined (Recall that $đ\xi = (2\pi)^{-n}\,d\xi$). By Fubini's theorem we have for any $\varphi \in C_0^\infty(\mathbb{R}^n)$

$$\begin{aligned}\int_{\mathbb{R}^n} (\varphi(0) - \varphi(x))\,\mu_t(dx) &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (1 - e^{i(x,\xi)})\hat{\varphi}(\xi)\,đ\xi\mu_t(dx) \\ &= \int_{\mathbb{R}^n} (\hat{\mu}_t(0) - \hat{\mu}_t(-\xi))\hat{\varphi}(\xi)\,đ\xi = \int_{\mathbb{R}^n} (e^{-t\overline{\psi}(0)} - e^{-t\overline{\psi}(-\xi)})\hat{\varphi}(\xi)\,đ\xi \\ &= e^{-t\psi(0)}\cdot\int_{\mathbb{R}^n} (1 - e^{-t(\psi(\xi)-\psi(0))})\hat{\varphi}(\xi)\,đ\xi,\end{aligned}$$

where we have used $\overline{\psi(-\xi)} = \psi(\xi)$. Recall also that $\operatorname{Re}(\psi(\xi) - \psi(0)) \geq 0$. Because for all $z \in \mathbb{C}$, $\operatorname{Re} z \geq 0$,

$$\lim_{t \downarrow 0} \frac{1 - e^{-tz}}{t} = z$$

and

$$\left| \frac{1 - e^{-tz}}{t} \right| \leq |z| \qquad \text{for all } t > 0,$$

we obtain by dominated convergence

$$\lim_{t \downarrow 0} \frac{1}{t} \int_{\mathbb{R}^n} (\varphi(0) - \varphi(x))\, \mu_t(dx) = \int_{\mathbb{R}^n} (\psi(\xi) - \psi(0)) \hat{\varphi}(\xi)\, đ\xi = \int_{\mathbb{R}^n} \psi(\xi) \hat{\varphi}(\xi)\, đ\xi = -T\varphi.$$

Moreover, if φ attains its nonnegative maximum in 0, we have $\int_{\mathbb{R}^n} (\varphi(0) - \varphi(x))\, \mu_t(dx) \geq 0$, hence $T\varphi \leq 0$, i.e. T satisfies the positive maximum principle and by Proposition 2.10 has the representation

$$\begin{aligned} T\varphi \; &= \; \sum_{j,k=1}^n a_{jk} \partial_{x_j} \partial_{x_k} \varphi(0) - \sum_{j=1}^n b_j \partial_{x_j} \varphi(0) - c \cdot \varphi(0) \\ &\quad + \int_{\mathbb{R}^n \setminus \{0\}} (\varphi(x) - \varphi(0) - \frac{(x, \nabla \varphi(0))}{1 + |x|^2})\, \mu(dx) \end{aligned} \tag{2.22}$$

with coefficients as claimed in the theorem. Using $\varphi(x) = \int_{\mathbb{R}^n} e^{i(x,\xi)} \hat{\varphi}(\xi)\, đ\xi$ we see that

$$\begin{aligned} T\varphi \; &= \; \int_{\mathbb{R}^n} \left(- \sum_{j,k=1}^n a_{jk} \xi_j \xi_k - i \sum_{j=1}^n b_j \xi_j - c \right) \hat{\varphi}(\xi) đ\xi \\ &\quad - \int_{\mathbb{R}^n \setminus \{0\}} \int_{\mathbb{R}^n} (1 - e^{i(x,\xi)} + \frac{i(x,\xi)}{1 + |x|^2}) \hat{\varphi}(\xi)\, đ\xi \mu(dx). \end{aligned}$$

By Lemma 2.8 we may change the order of integration in the second term and obtain

$$T\varphi = - \int_{\mathbb{R}^n} \left[\sum_{j,k=1}^n a_{jk} \xi_j \xi_k + i \sum_{j=1}^n b_j \xi_j + c + \int_{\mathbb{R}^n \setminus \{0\}} (1 - e^{i(x,\xi)} + \frac{i(x,\xi)}{1 + |x|^2})\, \mu(dx) \right] \hat{\varphi}(\xi)\, đ\xi.$$

Since this holds for all $\varphi \in C_0^\infty(\mathbb{R}^n)$ a comparision with (2.21) yields the representation (2.20). Finally, because for all $\varphi \in C_0^\infty(\mathbb{R}^n)$ with $\operatorname{supp} \varphi \subset \mathbb{R}^n \setminus \{0\}$ by (2.22)

$$T\varphi = \int_{\mathbb{R}^n \setminus \{0\}} \varphi(x)\, \mu(dx),$$

the measure μ is uniquely determined by ψ. In turn also the coefficients a_{jk}, b_j and c of the polynomial part of (2.20) are unique. □

Remark: Later on we shall be interested mainly in the case of continuous negative definite function which are real-valued. Replacing in (2.20) all terms by their complex conjugates we see by the uniqueness of the representation that for real-valued ψ the linear term $i \sum_{j=1}^n b_j \xi_j$ vanishes and the Lévy-measure is symmetric. Therefore

Corollary 2.14. *Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a real-valued continuous negative definite function. Then*

$$\psi(\xi) = \sum_{j,k=1}^{n} a_{jk}\xi_j\xi_k + c + \int_{\mathbb{R}^n\setminus\{0\}} (1 - \cos(x,\xi))\,\mu(dx), \tag{2.23}$$

where (a_{jk}) is a real-valued, nonnegative definite, symmetric matrix, $c \geq 0$ and μ is a symmetric Borel measure on $\mathbb{R}^n \setminus \{0\}$ such that $\int_{\mathbb{R}^n\setminus\{0\}} \frac{|x|^2}{1+|x|^2}\,\mu(dx) < \infty$.
Moreover, (a_{jk}), c and μ are uniquely determined by ψ.

A particular case of Corollary 2.14 will turn out to be useful to estimate the Lévy-measure of a continuous negative definite function. Note that the function $y \mapsto \frac{|y|^2}{1+|y|^2}$ is a bounded continuous negative definite function on $\mathbb{R}^n$. It has a bounded Lévy-measure. More precisely:

Lemma 2.15. *There is a bounded measure ν on $\mathbb{R}^n$ such that $\int_{\mathbb{R}^n} (1 + |\xi|^2)\,\nu(d\xi) < \infty$ and*

$$\int_{\mathbb{R}^n} (1 - \cos(y,\xi))\,\nu(d\xi) = \frac{|y|^2}{1+|y|^2}, \qquad y \in \mathbb{R}^n. \tag{2.24}$$

Proof: Define

$$\nu(d\xi) = \left(\int_0^\infty e^{-t}(4\pi t)^{-n/2} \exp(-\frac{|\xi|^2}{4t})\,dt \right) \cdot d\xi.$$

Then

$$\int_{\mathbb{R}^n} \nu(d\xi) = \int_0^\infty e^{-t} \int_{\mathbb{R}^n} (4\pi t)^{-n/2} \exp(-\frac{|\xi|^2}{4t})\,d\xi dt = \int_0^\infty e^{-t}\,dt = 1$$

and

$$\int_{\mathbb{R}^n} |\xi|^2\ \nu(d\xi) = \int_0^\infty e^{-t} \int_{\mathbb{R}^n} |\xi|^2\ (4\pi t)^{-n/2} \exp(-\frac{|\xi|^2}{4t})\,d\xi dt = \int_0^\infty 2nt \cdot e^{-t}\,dt = 2n$$

and finally

$$\begin{aligned} \int_{\mathbb{R}^n} (1 - \cos(y,\xi))\,\nu(d\xi) &= 1 - \int_0^\infty e^{-t} \int_{\mathbb{R}^n} \cos(y,\xi)(4\pi t)^{-n/2} \exp(-\frac{|\xi|^2}{4t})\,d\xi dt \\ &= 1 - \int_0^\infty e^{-t}\,e^{-t|y|^2}\,dt = 1 - \frac{1}{1+|y|^2} = \frac{|y|^2}{1+|y|^2}. \end{aligned}$$

□

We now combine the Lévy-type representation and the Lévy–Khinchin formula. The result is essential for the following and provides another equivalent representation of an operator satisfying the positive maximum principle as a pseudo differential operator. This representation was first observed by Ph. Courrège [13].
Recall that by a **pseudo differential operator** we understand an operator $p(x, D)$ of the type

$$p(x,D)\varphi(x) = \int_{\mathbb{R}^n} e^{i(x,\xi)} p(x,\xi)\hat{\varphi}(\xi)\,đ\xi, \qquad \varphi \in C_0^\infty(\mathbb{R}^n). \tag{2.25}$$

The function $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ which defines the operator is called the **symbol** of the operator. If not stated otherwise we will always assume that the pseudo differential operator $p(x, D)$ is defined on $C_0^\infty(\mathbb{R}^n)$. For symbols $p(\xi)$ which do not depend on x we simply write $p(D)$. Note that unlike the situation of usual symbol classes we do not a priori assume strong smoothness assumption for the symbol.

Theorem 2.16. *Let $A : C_0^\infty(\mathbb{R}^n) \to M(\mathbb{R}^n)$ be a linear operator. Then A satisfies the positive maximum principle if and only if A is a pseudo differential operator*

$$A = -p(x, D),$$

where the symbol $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ has the property that

$$\xi \mapsto p(x, \xi)$$

is a continuous negative definite function for all $x \in \mathbb{R}^n$.
The symbol $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ is uniquely determined by $p(x, D)$.

Thus there is a one-to-one correspondence between this type of operators and Lévy-type operators (2.19) in Theorem 2.12, which is given in the following way: For each $x \in \mathbb{R}^n$ the coefficients and the Lévy-measure of the continuous negative definite function $p(x, \cdot)$ in the Lévy–Khinchin representation (2.20) coincide with the corresponding terms of the Lévy-type operator in (2.19).

Proof: By (2.25) it is clear that the operator determines the symbol p in a unique way. As the same holds true for the coefficients of the Lévy-type operator (2.19) it remains to prove the equivalence of both representations. Replacing φ by $\int_{\mathbb{R}^n} e^{i(x,\xi)} \hat{\varphi}(\xi)\, d\!\!{}^-\xi$ in (2.19) we obtain

$$\begin{aligned} & A\varphi(x) = \\ = \; & -\int_{\mathbb{R}^n} \Big[\sum_{j,k=1}^n a_{jk}(x) \xi_j \xi_k + i \sum_{j=1}^n b_j(x) \xi_j + c(x) \\ & \quad + \int_{\mathbb{R}^n \setminus \{0\}} (1 - e^{i(y,\xi)} + \frac{i(y,\xi)}{1 + |y|^2})\, \mu(x, dy) \Big] e^{i(x,\xi)} \hat{\varphi}(\xi)\, d\!\!{}^-\xi \\ = \; & -\int_{\mathbb{R}^n} e^{i(x,\xi)} p(x, \xi) \hat{\varphi}(\xi)\, d\!\!{}^-\xi, \end{aligned}$$

where we have used again Lemma 2.8 to interchange the order of integration. □

We therefore introduce the following notion:

Definition 2.17. *A symbol $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ such that $\xi \mapsto p(x, \xi)$ is a continuous negative definite function for all $x \in \mathbb{R}^n$ is called a* **negative definite symbol**.

We remark that a negative definite symbol satisfies $p(x, -\xi) = \overline{p(x, \xi)}$ by Proposition 2.2 (ii) and therefore the corresponding pseudo differential operator maps real-valued functions into real-valued functions. This of course can also be seen from the Lévy-type representation of the operator.

If we assume additional smoothness for the symbol with respect to x it is easy to see that $p(x, D)$ then maps testfunctions into classes of smooth functions. Our minimal standard assumption will be that the symbol $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ is a continuous function of both variables. In this case for x in some open, relatively compact subset U of $\mathbb{R}^n$ we know by Theorem 2.7

$$|p(x,\xi)| \le c_{p,U}(1+|\xi|^2),$$

where $c_{p,U} = 2 \sup\limits_{\substack{|\xi|\le 1\\ x\in U}} |p(x,\xi)| < \infty$. Therefore by dominated convergence

$$x \mapsto \int_{\mathbb{R}^n} e^{i(x,\xi)} p(x,\xi)\hat{\varphi}(\xi)\, đ\xi$$

depends continuously on x for all $\varphi \in C_0^\infty(\mathbb{R}^n)$ and we have shown

Theorem 2.18. *Let $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ be a continuous negative definite symbol. Then*

$$-p(x,D) : C_0^\infty(\mathbb{R}^n) \to C(\mathbb{R}^n)$$

is a linear operator that satisfies the positive maximum principle.

Note that the converse is not true, i.e. for a negative definite symbol p the property that $-p(x, D)$ maps $C_0^\infty(\mathbb{R}^n)$ into $C(\mathbb{R}^n)$ does not imply that $p(x,\xi)$ is a continuous function, see [13] for a counter example.

Let us also mention that for continuous negative definite symbols the Lévy-kernel in the Lévy–Khinchin representation (2.19) actually is a kernel in the sense of measure theory, that is $\mu(x, dy)$ depends on x in measurable way. This follows from the following lemma.

Lemma 2.19. *Let $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ be a continuous negative definite symbol with Lévy–Khinchin representation (2.19) and Lévy-kernel $\mu(x, dy)$. Then for all $\varphi \in C_0^\infty(\mathbb{R}^n \setminus \{0\})$ the map*

$$x \mapsto \int_{\mathbb{R}^n\setminus\{0\}} \varphi(y)\frac{|y|^2}{1+|y|^2}\,\mu(x,dy)$$

is continuous on $\mathbb{R}^n$. In particular the map $x \mapsto \mu(x, A)$ is measurable for all Borel sets $A \subset \mathbb{R}^n \setminus \{0\}$.

Proof: Let $\varphi \in C_0^\infty(\mathbb{R}^n \setminus \{0\})$ and define $\tilde{\varphi} \in C_0^\infty(\mathbb{R}^n)$ by $\tilde{\varphi}(y) = \frac{|y|^2}{1+|y|^2}\varphi(y)$ for $y \neq 0$ and $\tilde{\varphi}(0) = 0$. Then $\partial_{x_j}\tilde{\varphi}(0) = \partial^2_{x_j x_k}\tilde{\varphi}(0) = 0$ and therefore by Theorem 2.16

$$\begin{aligned}
\int_{\mathbb{R}^n\setminus\{0\}} \varphi(y)\frac{|y|^2}{1+|y|^2}\,\mu(x,dy) &= \int_{\mathbb{R}^n\setminus\{0\}} \tilde{\varphi}(y)\,\mu(x,dy)\\
&= \sum_{j,k=1}^n a_{jk}(x)\partial^2_{x_j x_k}\tilde{\varphi}(0) - \sum_{i=j}^n b_j(x)\partial_{x_j}\tilde{\varphi}(0) - c(x)\cdot\tilde{\varphi}(0)\\
&\quad + \int_{\mathbb{R}^n\setminus\{0\}} (\tilde{\varphi}(x) - \tilde{\varphi}(0) - \frac{(y,\nabla\tilde{\varphi}(0))}{1+|y|^2})\,\mu(x,dy)\\
&= -\int_{\mathbb{R}^n} e^{i(0,\xi)} p(x,\xi)\hat{\tilde{\varphi}}(\xi)\,đ\xi = -\int_{\mathbb{R}^n} p(x,\xi)\hat{\tilde{\varphi}}(\xi)\,đ\xi,
\end{aligned}$$

which depends continuously on x. The measurability statement now follows from a standard approximation argument. □

Often we will consider some extension of the operator $-p(x,D)$ and we need the positive maximum principle also on the larger domain of definition. For that purpose the following proposition is useful, it is based on the proof of Theorem 9.3 in [42]. We recall that $C_\infty(\mathbb{R}^n)$ denotes the space of all continuous functions on $\mathbb{R}^n$ that vanish at infinity.

Proposition 2.20. *Let $(A, D(A))$ be a linear operator in $C_\infty(\mathbb{R}^n)$. If $C_0^\infty(\mathbb{R}^n) \subset D(A)$ is a core of A and A satisfies the positive maximum principle on $C_0^\infty(\mathbb{R}^n)$, then A satisfies the positive maximum principle also on $D(A)$.*

Proof: Let $u \in D(A)$ such that $\sup_{x\in\mathbb{R}^n} u(x) = u(x_0) \geq 0$ for some $x_0 \in \mathbb{R}^n$. Choose a function $\chi \in C_0^\infty(\mathbb{R}^n)$ such that $\chi(x_0) = 1$, but $\chi(x) < 1$ for all $x \neq x_0$. Then for all $\eta > 0$

$$\sup_{x\in\ \mathbb{R}^n} (u + \eta\chi)(x) = u(x_0) + \eta > 0.$$

Since $C_0^\infty(\mathbb{R}^n)$ is a core of A we find a sequence $(\varphi_k^\eta)_{k\in\mathbb{N}}$ of testfunctions such that

$$\varphi_k^\eta \to u + \eta\chi,\ A\varphi_k^\eta \to A(u + \eta\chi) \qquad \text{uniformly as } k \to \infty. \tag{2.26}$$

For each $k \in \mathbb{N}$ choose a point $x_k \in \mathbb{R}^n$, where the testfunction φ_k^η attains its maximum. By the uniform convergence of φ_k^η to $u + \eta\chi$ we have

$$\varphi_k^\eta(x_k) \to u(x_0) + \eta \tag{2.27}$$

as $k \to \infty$. Suppose that there is a neighbourhood $U(x_0)$ of x_0 such that $x_k \notin U(x_0)$ for all k. Then by the properties of χ there is an $\varepsilon > 0$ such that

$$(u + \eta\chi)(x) < u(x_0) + \eta - \varepsilon \qquad \text{for all } x \in \mathbb{R}^n \setminus U(x_0).$$

But this contradicts (2.27). Thus there is a subsequence of $(x_k)_{k\in\mathbb{N}}$ that converges to x_0. Without loss of generality we may assume that $x_k \to x_0$ and that $\varphi_k^\eta(x_k) \geq 0$ by (2.27). But A satisfies the positive maximum principle on $C_0^\infty(\mathbb{R}^n)$, hence by (2.26)

$$Au(x_0) = A(u + \eta\chi)(x_0) - \eta A\chi(x_0) = \lim_{k\to\infty} A\varphi_k^\eta(x_k) - \eta A\chi(x_0) \leq -\eta A\chi(x_0).$$

Since $\eta > 0$ was arbitray, we have $Au(x_0) \leq 0$. □

Chapter 3

The martingale problem: Existence of solutions

3.1 The martingale problem for jump processes

The martingale problem presents a possibility to describe the relation between a stochastic process and an associated generator in a purely probabilistic way. Introduced by Stroock and Varadhan [82] in order to characterize diffusion processes corresponding to diffusion operators with merely continuous coefficients, the martingale problem was also extended to jump processes, see the monograph [83] as a standard reference. A good introduction to the topic is also contained in [17].

Let (E, d) be a separable metric space. By D_E we denote the space of all càdlàg–paths with values in E, i.e.

$$D_E = \{\omega : [0, \infty) \to E : \ \omega \text{ is right continuous, } \lim_{s \uparrow t} \omega(s) \text{ exists for all } t > 0\}.$$

Clearly, $\omega \in D_E$ has at most countably many points of discontinuity and all discontinuities are of first kind. Moreover the oscillation of ω is bounded in the following way, see [5], p.110:

Lemma 3.1. *Let $\omega \in D_E$ and $T > 0$. Then for all $\varepsilon > 0$ there is an $m \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_{m-1} \le T < t_m$ such that*

$$\sup_{s,t \in [t_{i-1}, t_i)} d(\omega(s), \omega(t)) < \varepsilon \qquad (i = 1, \ldots, m).$$

In particular $\omega \in D_E$ has at most finitely many jumps of size greater than any given $\varepsilon > 0$ in any compact interval of time.

D_E is equipped with the Skorohod topology. This topology generalizes the topology of locally uniform convergence for continuous paths to the càdlàg–case. It is defined in the following way: Let Λ be the set of all bijective, monotonously increasing Lipschitz functions $\lambda : [0, \infty) \to [0, \infty)$ such that

$$\|\lambda\|_\Lambda := \sup_{s > t \ge 0} \left| \log \frac{\lambda(s) - \lambda(t)}{s - t} \right| < \infty.$$

For any ω_1, $\omega_2 \in D_E$, $\lambda \in \Lambda$ and $u \geq 0$ we define

$$\tilde{D}(\omega_1, \omega_2, \lambda, u) = \sup_{t \geq 0} d(\omega_1(t \wedge u), \omega_2(\lambda(t) \wedge u)) \wedge 1$$

and

$$D(\omega_1, \omega_2) = \inf_{\lambda \in \Lambda} \left[\|\lambda\|_\Lambda \vee \int_0^\infty e^{-u} \tilde{D}(\omega_1, \omega_2, \lambda, u)\, du \right].$$

Then D is a metric on D_E which turns D_E into a complete separable metric space and induces the Skorohod topology, see [17], 3.5. Note that convergence in the Skorohod topology does not imply pointwise convergence everywhere in the usual sense, but locally uniform convergence if we allow an arbitrary small distortion of the time scale, i.e. $\omega_k \to \omega$ in D_E if and only if there are functions $\lambda_k \in \Lambda$ such that $\|\lambda_k\|_\Lambda \underset{k\to\infty}{\longrightarrow} 0$ and

$$\lim_{k \to 0} \sup_{0 \leq t \leq T} |d(\omega_k(t), \omega(\lambda_k(t)))| = 0 \qquad \text{for all } T > 0,$$

see [17], 3.5.3. In particular, $\omega_k(t) \to \omega(t)$ at all points t of continuity of ω, whereas at points $t > 0$ of discontinuity of ω there is always a sequence (ω_k) which converges in the Skorohod topology to ω, but not pointwise at t. In other words the map $\omega \mapsto \omega(t)$ is discontinuous unless $t = 0$ and its points of continuity are given by those $\omega \in D_E$ which are continuous at t.

We will consider the canonical D_E-process in the following sense: let

$$X_t : D_E \to E, \qquad t \in [0, \infty),$$

be the position map

$$X_t(\omega) = \omega(t).$$

We define the σ-algebras of the corresponding canonical filtration

$$\mathcal{F}_t = \sigma(X_s : s \leq t), \qquad t \in [0, \infty),$$

and the σ-algebra

$$\mathcal{F} = \sigma(X_t : t \in [0, \infty))$$

on D_E. Moreover let as usually

$$X_{t-} = \lim_{s \uparrow t} X_s \qquad \text{for } t > 0,$$

and

$$\mathcal{F}_{t+} = \bigcap_{s > t} \mathcal{F}_s.$$

Note that the σ-algebra $\mathcal{F}$ coincides with the Borel σ-algebra generated by the Skorohod topology.

By $\mathcal{M}_1(D_E)$ we denote the set of probability measures on $(D_E, \mathcal{F})$. Then $\mathcal{M}_1(D_E)$ equipped with the weak topology is a separable and completely metrizable space (see [17], 3.5.6). Hence by the Prohorov theorem, see [17], 3.2.2, relative compactness and tightness of subsets of $\mathcal{M}_1(D_E)$ are the same.

For every probability measure $P \in \mathcal{M}_1(D_E)$ the family $(X_t)_{t\geq 0}$ defines a càdlàg–process which is obviously adapted and by right-continuity even progressively measurable with respect to the filtration $(\mathcal{F}_t)$. Conversely, for every process with paths in D_E there is a unique probability measure $P \in \mathcal{M}_1(D_E)$ such $(X_t)_{t\geq 0}$ has the same finite–dimensional distributions under P.

We already have mentioned that the set of paths ω where X_t is continuous is given by the set of all paths which do not jump at time t. Therefore the weak convergence of a sequence of probability measures P_k to P in $\mathcal{M}_1(D_E)$ in general does not imply the weak convergence of the corresponding finite-dimensional distributions. Nevertheless, such a result is true, if we take a small exceptional set into account (see [17], chap.3.7):

Proposition 3.2. *Let $(P_k)_{k\in\mathbb{N}}$ be a sequence of probability measures in $\mathcal{M}_1(D_E)$ which converges weakly to a probability measure P. Then the set*

$$T_P = \{t \geq 0 : P(X_{t-} = X_t) = 1\}$$

has an at most countable complement in $[0, \infty)$ and for all $t_1, \dots, t_m \in T_P$ the finite-dimensional distributions $(X_{t_1}, \dots, X_{t_m})(P_k)$ converge weakly to $(X_{t_1}, \dots, X_{t_m})(P)$.

Definition 3.3. *Let $A : D(A) \to B(E)$ be a linear operator with domain $D(A) \subset B(E)$. A probability measure $P \in \mathcal{M}_1(D_E)$ is called a solution of the (D_E –)* **martingale problem** *for the operator A if for every $\varphi \in D(A)$*

$$\varphi(X_t) - \int_0^t A\varphi(X_s)\, ds, \qquad t \geq 0, \tag{3.1}$$

is a martingale under P with respect to the filtration $(\mathcal{F}_t)$.

If for every probability measure $\mu \in \mathcal{M}_1(E)$ there is a unique solution P_μ of the martingale problem for A with initial distribution

$$P_\mu \circ X_0^{-1} = \mu,$$

then the martingale problem for A is called well-posed.

Remark 3.4.

(i) Recall that if A is the generator of a Feller semigroup then the fact that (3.1) is a martingale is just a probabilistic reformulation of the semigroup property in terms of the corresponding Feller process. Thus the martingale problem in general gives a weak formulation of the correspondence of a Markov process and its generator. However, if the martingale problem is well-posed then the process is even strong Markov and typically defines a Feller semigroup.

(ii) Note that the fact that an E-valued process $(Y_t)_{t\geq 0}$, which is progressively measurable with respect to the canonical filtration, turns

$$\varphi(Y_t) - \int_0^t A\varphi(Y_s)\, ds, \qquad t \geq 0, \tag{3.2}$$

into a martingale only depends on the finite dimensional distributions of the process and so any progessively measurable version of the process will do.

Therefore it is no serious restriction in Definition 3.3 to assume a priori that the solution of the martingale problem has càdlàg paths, since typically a càdlàg version exists. This result mainly relies on the result of Doob ([15], p.363) that there always is a càdlàg version of the martingale in (3.2) and hence of $(\varphi(Y_t))$. This property extends to (Y_t) itself if the domain of A satisfies a certain point seperation condition, which is fulfilled in all cases we are interested in. See [17], 4.3 for details concerning this question.

In order to prove existence of solutions to the martingale problem it is a standard procedure to first approximate the operator A by a sequence of operators, for which the martingale problem is solvable in an elementary way. In a second step one proves tightness of the solutions for the approximating operators and identifies an accumulation point as a solution of the martingale problem for A.

The class of approximating operator we have in mind are operators $K : B(E) \to B(E)$ of the type

$$Kf(x) = \lambda \int_E (f(y) - f(x))\, \mu(x, dy), \tag{3.3}$$

where $\lambda > 0$ and $\mu(x, dy)$ is a kernel of sub-probability measures on E. By replacing the kernel $\mu(x, dy)$ by $\mu(x, dy) + (1 - \mu(x, E)) \cdot \varepsilon_x(dy)$ we may and do assume that $\mu(x, dy)$ is even a kernel of probability measures.
Note that in the case $E = \mathbb{R}^n$ the operator K is a Lévy-type operator having no diffusion part and a Lévy-kernel which consists of uniformly bounded measures. In particular K is a bounded operator in $B(E)$ and generates a strongly continuous (Markovian) semigroup

$$T_t = e^{tK} = \sum_{k=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} \Gamma^k \tag{3.4}$$

in $B(E)$, where $\Gamma f(x) = \int_E f(y)\, \mu(x, dy)$ is the operator in $B(E)$ generated by the kernel $\mu(x, dy)$, i.e. $K = \lambda(\Gamma - \mathrm{Id})$, see [16], II 2.2.2, II 2.2.18. It is well-known that for any initial distribution $\nu \in \mathcal{M}_1(E)$ there is a Markov process $(Z_t)_{t\geq 0}$ in E generated by the transition semigroup $(T_t)_{t\geq 0}$, which can be defined by

$$Z_t = Y(V_t). \tag{3.5}$$

Here $(V_t)_{t\geq 0}$ is a standard Poisson process with parameter λ and $(Y(l))_{l\in\mathbb{N}_0}$ is an E-valued Markov process, independent of (V_t), with discrete parameter set $\mathbb{N}_0$, initial distribution ν and transition function $\mu(x, dy)$ (see [17], 4.2 for a detailed derivation). We then have

$$\mathbb{E}\left[f(Z_{t+s}) | \mathcal{F}_t^Z\right] = T_s f(Z_t) \qquad \text{a.s.} \tag{3.6}$$

for all $f \in B(E)$, $s, t \geq 0$, where $(\mathcal{F}_t^Z)$ is the canonical filtration of (Z_t). It is now easy to show

Proposition 3.5. *Let the operator K with domain $B(E)$ be defined as above. Then for any initial distribution $\nu \in \mathcal{M}_1(E)$ there is a solution of the martingale problem for K.*

Proof: The process (Z_t) has paths in D_E by construction and the distribution of $Z_0 = Y(V_0) = Y(0)$ equals the given measure ν. It remains to verify that

$$f(Z_t) - \int_0^t Kf(Z_s)\,ds, \qquad t \geq 0,$$

is an $(\mathcal{F}_t^Z)$-martingale, then the distribution of (Z_t) on the path space is a solution of the martingale problem. To this end note that the semigroup (T_t) in $B(E)$ is strongly continuous. Hence by (3.6) for $0 \leq t_1 \leq t_2$

$$\begin{aligned}
&\mathbb{E}\left[f(Z_{t_2}) - \int_0^{t_2} (Kf)(Z_u)\,du|\mathcal{F}_{t_1}^Z\right] \\
&= T_{t_2-t_1}f(Z_{t_1}) - \int_{t_1}^{t_2} T_{u-t_1}(Kf)(Z_{t_1})\,du - \int_0^{t_1} (Kf)(Z_u)\,du \\
&= T_{t_2-t_1}f(Z_{t_1}) - \int_0^{t_2-t_1} T_u(Kf)(Z_{t_1})\,du - \int_0^{t_1} (Kf)(Z_u)\,du \\
&= f(Z_{t_1}) - \int_0^{t_1} (Kf)(Z_u)\,du \qquad \text{a.s.}
\end{aligned}$$

□

Let us also mention that the solvability of the martingale problem is stable under perturbations of type (3.3). For the proof of the following proposition we refer [17], Prop.4.10.2.

Proposition 3.6. *Let A be a linear operator in $B(E)$ such that the martingale problem for A is solvable for all initial distributions and let K be given by (3.3). Then for all initial distributions there is a solution of the martingale problem for $A + K$.*

The idea of the proof is based on the fact that the process (Z_t) as defined in (3.5) has only finitely many jumps in finite intervals of time and remains constant in between. The solution of the martingale problem for $A+K$ then can be obtained by interlacing the solution of A with jumps corresponding to K.

Once we have found solutions $P_n \in \mathcal{M}_1(D_E)$ of the martingale problem for a sequence of operators of type (3.3) that approximates the general operator A, we have construct in the second step a limit element of the sequence (P_n). For that purpose a tightness criterion for subsets of $\mathcal{M}_1(D_E)$ is needed. Therefore we define the following expression, which generalizes the modulus of continuity for a continuous path to the case of càdlàg paths: for $\omega \in D_E$, $\delta > 0$ and $T > 0$ let

(3.7) $$w'(\omega, \delta, T) = \inf_{\mathcal{Z}(\delta,T)} \sup_{\substack{s,t\in[t_i,t_{i+1})\\ i=0,\ldots,m \\ (t_0,\ldots,t_{m+1})\in\mathcal{Z}(\delta,T)}} d(\omega(s), \omega(t)),$$

where $\mathcal{Z}(\delta, T)$ denotes all finite partitions $0 = t_0 < t_1 < \ldots < t_m \leq T < t_{m+1}$ with $m \in \mathbb{N}$ and $\inf_{i=0,\ldots,m} (t_{i+1} - t_i) > \delta$. Note that the map $\omega \mapsto w'(\omega, \delta, T)$ is $\mathcal{F}$-measurable for fixed δ and T (cf. [17], 3.6.2). Moreover, we have obviously by the triangle inequality for $\omega, \omega' \in D_E$

(3.8) $$w'(\omega, \delta, T) \leq w'(\omega', \delta, T) + 2 \cdot \sup_{0\leq t<T+\delta} d(\omega(t), \omega'(t)).$$

Now analoguous to the theorem of Arzela–Ascoli the relatively compact subsets of D_E are characterized as sets of paths which are locally uniformly bounded and whose oscillation measured by $w'(\cdot, \delta, T)$ tends uniformly to 0 as $\delta \to 0$.
By the Prohorov theorem this leads to the following standard characterization of tightness of probability measures in $\mathcal{M}_1(D_E)$ (see [5], p.125 or [17], 3.7.2, 3.7.3).

Theorem 3.7. *Let $(P_\alpha)_{\alpha\in I}$ be a family of probability measures in $\mathcal{M}_1(D_E)$ indexed by a set I. Then the family $(P_\alpha)_{\alpha\in I}$ is tight if and only if*

(i) the compact containment condition holds:
For all $T > 0$ and all $\varepsilon > 0$ there is a compact set $K_{T,\varepsilon} \subset E$ such that

$$\sup_{\alpha\in I} P_\alpha(\omega \in D_E : \omega(t) \notin K_{T,\varepsilon} \text{ for some } 0 \le t \le T) \le \varepsilon \tag{3.9}$$

and

(ii) for all $T > 0$ and all $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\sup_{\alpha\in I} P_\alpha(\omega \in D_E : w'(\omega, \delta, T) \ge \varepsilon) \le \varepsilon. \tag{3.10}$$

The condition (3.10) on the oscillation of the paths can be replaced by an upper bound for the variation of the paths. The following condition is due to T.G. Kurtz [56], (4.20), see also [17], Th. 3.8.6. We use the notation $q = d \wedge 1$, where d is the metric in E, the expectation with respect to P_α is denoted by $\mathbb{E}^\alpha$.

Theorem 3.8. *Let $(P_\alpha)_{\alpha\in I}$ be a family of probability measures in $\mathcal{M}_1(D_E)$. If the compact containment condition (3.9) is satisfied and if*

(i)

$$\sup_{\alpha\in I} \mathbb{E}^\alpha \left[q(X_t, X_0)^2\right] \underset{t\to 0}{\longrightarrow} 0 \tag{3.11}$$

and

(ii) for all $T > 0$ there is a family of measurable functions

$$\gamma_\delta : D_E \to [0, \infty), \qquad 0 < \delta \le 1,$$

such that

$$\sup_{\alpha\in I} \mathbb{E}^\alpha \left[\gamma_\delta\right] \underset{\delta\to 0}{\longrightarrow} 0 \tag{3.12}$$

and

$$\mathbb{E}^\alpha \left[q(X_{t+u}, X_t)^2 \cdot q(X_t, X_{t-v})^2 | \mathcal{F}_t\right] \le \mathbb{E}^\alpha \left[\gamma_\delta | \mathcal{F}_t\right] \qquad (P_\alpha\text{-a.s.}) \tag{3.13}$$

for all $0 \le t \le T$, $0 \le u \le \delta \le 1$ and $0 \le v \le \delta \wedge t$,

then $(P_\alpha)_{\alpha\in I}$ is tight.

Provided we already know that the measures P_α are solutions of the martingale problem for certain operators A_α the situation fortunately is more easy. If in this case we have a uniform bound for the operators A_α, it is then enough to check the compact containment condition (3.9), whereas the uniform control on the oscillation of the paths can be derived directly from properties of martingales of type (3.1). To see this we first note

Lemma 3.9. *Let $(P_\alpha)_{\alpha\in I}$ be a family of probability measures in $\mathcal{M}_1(D_E)$ that fulfills the compact containment condition (3.9). Moreover let D be a subspace of $C_b(E)$ which is dense with respect to uniform convergence on compact sets. For every $f \in C_b(E)$ and every P_α let $P_{f,\alpha} \in \mathcal{M}_1(D_\mathbb{R})$ be the distribution of the paths of the $\mathbb{R}$-valued càdlàg process $(f(X_t))_{t\geq 0}$ under P_α.*
If for every $f \in D$ the set $(P_{f,\alpha})_{\alpha\in I}$ is tight in $\mathcal{M}_1(D_\mathbb{R})$ then $(P_\alpha)_{\alpha\in I}$ is tight in $\mathcal{M}_1(D_E)$.

Proof: (see [17], 3.9.1) We first prove that $(P_{f,a})_{\alpha\in I}$ is tight for all $f \in C_b(E)$. Let $f \in C_b(E)$ and $(f_n)_{n\in\mathbb{N}}$ a sequence in D that converges uniformly to f on compact sets. Moreover fix $T > 0$ and $\varepsilon > 0$. Then by (3.9) there is compact set $K_{T,\varepsilon} \subset E$ such that

$$M_{T,\varepsilon} := \{\omega \in D_E : \omega(t) \in K_{T,\varepsilon} \text{ for all } 0 \leq t \leq T+1\}$$

satisfies

$$P_\alpha(M_{T,\varepsilon}) \geq 1 - \frac{\varepsilon}{2} \qquad \text{for all } \alpha \in I. \tag{3.14}$$

Since for n suffciently large we have $\sup_{x\in K_{T,\varepsilon}} |f(x) - f_n(x)| \leq \frac{\varepsilon}{4}$, (3.8) implies for such n and $0 < \delta \leq 1$ sufficiently small

$$\begin{aligned}
P_{f,\alpha}(w'(\cdot,\delta,T) \geq \varepsilon) &= P_\alpha(w'(f(X.),\delta,T) \geq \varepsilon) \\
&\leq P_\alpha((w'(f(X.),\delta,T) \geq \varepsilon) \cap M_{T,\varepsilon}) + P_\alpha(M^{\complement}_{T,\varepsilon}) \\
&\leq P_\alpha((w'(f_n(X.),\delta,T) + 2\sup_{0\leq t\leq T+1}|f(X_t) - f_n(X_t)| \geq \varepsilon) \cap M_{T,\varepsilon}) + \frac{\varepsilon}{2} \\
&\leq P_\alpha(w'(f_n(X.),\delta,T) \geq \frac{\varepsilon}{2}) + \frac{\varepsilon}{2} \\
&= P_{f_n,\alpha}(w'(\cdot,\delta,T) \geq \frac{\varepsilon}{2}) + \frac{\varepsilon}{2} \leq \varepsilon
\end{aligned}$$

by tightness of $(P_{f_n,\alpha})_{\alpha\in I}$ and Theorem 3.7, condition (ii). Thus $(P_{f,\alpha})_{\alpha\in\ I}$ also satisfies condition (ii) of Theorem 3.7 and moreover clearly satisfies the compact containment condition (3.9), since f is bounded. Consequently by the converse direction of Theorem 3.7 $(P_{f,\alpha})_{\alpha\in I}$ is tight for all $f \in C_b(E)$.

In particular, this holds true for the bounded and continuous functions $x \mapsto q(x,z)$, $z \in E$, where $q = d \wedge 1$ and d is the metric of E. We prove tightness of $(P_\alpha)_{\alpha\in I}$ using Theorem 3.8. Let $\varepsilon > 0$, $T > 0$ and $K_{T,\varepsilon}$, $M_{T,\varepsilon}$ as above. Then we find finitely many points $z_1,\ldots,z_{N_\varepsilon} \in K_{T,\varepsilon}$ such that the open balls $B_\varepsilon(z_i)$ of radius ε cover the compact set $K_{T,\varepsilon}$. Therefore, for $y \in K_{T,\varepsilon}$ there is an $i \in \{1,\ldots,N_\varepsilon\}$ such that

$$q(x,y) \leq q(x,z_i) + q(y,z_i) \leq |q(x,z_i) - q(y,z_i)| + 2\varepsilon \qquad \text{for all } x \in E.$$

Consequently for $0 \le t \le T$, $0 < \delta \le 1$ and $0 \le u < \delta$, $0 \le v \le \delta \wedge t$

$$
\begin{aligned}
(3.15) \quad & q^2(X_{t+u}, X_t) \cdot q^2(X_t, X_{t-v}) \le q(X_{t+u}, X_t) \cdot q(X_t, X_{t-v}) \\
& \le \max_{i=1,\dots,N_\varepsilon} \{|q(X_{t+u}, z_i) - q(X_t, z_i)| \cdot |q(X_t, z_i) - q(X_{t-v}, z_i)|\} + 4(\varepsilon + \varepsilon^2) + 1_{M_{T,\varepsilon}^{\complement}} \\
& \le \max_{i=1,\dots,N_\varepsilon} w'(q(X_\cdot, z_i), 2\delta, T+1) \wedge 1 \; + 4(\varepsilon + \varepsilon^2) + 1_{M_{T,\varepsilon}^{\complement}} =: \gamma_\delta,
\end{aligned}
$$

where the last inequality follows from the fact that $w'(\cdot, 2\delta, T+1)$ is calculated from the oscillation of the path on intervals of length not smaller than 2δ. Thus at least one of the intervals $[t-v, t)$ or $[t, t+u)$ is completely contained in such intervals. In particular, (3.13) in Theorem 3.8 holds with this choice of γ_δ.

For $\eta > 0$ we fix ε and hence N_ε such that

$$4(\varepsilon + \varepsilon^2) + \frac{\varepsilon}{2} \le \frac{\eta}{2}.$$

By Theorem 3.7

$$\sup_{\alpha \in I} \mathbb{E}^\alpha \left[w'(q(X_\cdot, z_i), 2\delta, T+1) \wedge 1\right] \le \frac{\eta}{2N_\varepsilon} \qquad \text{for all } i \in \{1, \dots, N_\varepsilon\}$$

if δ is sufficiently small. Thus by (3.14) and (3.15)

$$\sup_{\alpha \in I} \mathbb{E}^\alpha [\gamma_\delta] \le \eta,$$

i.e. (3.12) holds, since $\eta > 0$ was arbitrary.

Finally we have by definition of w' for $0 \le t \le T$ and $\varepsilon > 0$ with the notation as above

$$
\begin{aligned}
q^2(X_t, X_0) & \le q(X_t, X_0) \le \max_{i=1,\dots,N_\varepsilon} |q(X_t, z_i) - q(X_0, z_i)| + 2\varepsilon + 1_{M_{T,\varepsilon}^{\complement}} \\
& \le \max_{i=1,\dots,N_\varepsilon} w'(q(X_\cdot, z_i), t, T) + 2\varepsilon + 1_{M_{T,\varepsilon}^{\complement}}
\end{aligned}
$$

and thus again with the argument as above by Theorem 3.7 and (3.14)

$$\sup_{\alpha \in I} \mathbb{E}^\alpha \left[q^2(X_t, X_0)\right] \xrightarrow[t \to 0]{} 0,$$

i.e. (3.11) is satisfied and $(P_\alpha)_{\alpha \in I}$ is tight. □

The tightness criterion for solutions of the martingale problem now reads as follows (cf. [17], 3.9.4).

Theorem 3.10. *Let D be a subalgebra of $C_b(E)$ which is dense with respect to uniform convergence on compact sets and let $((A_\alpha, D))_{\alpha \in I}$ be a family of linear operators in $C_b(E)$ with common domain D. Assume*

$$(3.16) \qquad \sup_{\alpha \in I} \|A_\alpha f\|_\infty < \infty \qquad \textit{for all } f \in D.$$

If for each $\alpha \in I$ the measure $P_\alpha \in \mathcal{M}_1(D_E)$ is a solution of the martingale problem for A_α and $(P_\alpha)_{\alpha \in I}$ satisfies the compact containment condition (3.9), then $(P_\alpha)_{\alpha \in I}$ is tight.

Proof: By Lemma 3.9 it is enough to check that for all $f \in D$ the process $(f(X_t))_{t\geq 0}$ defines a tight family of distributions $(P_{f,\alpha})_{\alpha\in I}$ in $\mathcal{M}_1(D_\mathbb{R})$. Let $X_t^\mathbb{R} : D_\mathbb{R} \to \mathbb{R}$, $t \geq 0$, the canonical position map on $D_\mathbb{R}$ and $\mathcal{F}^\mathbb{R} = \sigma(X_t^\mathbb{R} : t \geq 0)$, $\mathcal{F}_t^\mathbb{R} = \sigma(X_s^\mathbb{R} : s \leq t)$ the corresponding σ-algebra and filtration. A function $f \in C_b(E)$ induces a map

$$\begin{aligned} \tilde{f} : D_E &\to D_\mathbb{R} \\ \omega &\mapsto f \circ \omega. \end{aligned}$$

$\tilde{f}$ is $\mathcal{F} - \mathcal{F}^\mathbb{R}$ - and $\mathcal{F}_t - \mathcal{F}_t^\mathbb{R}$ -measurable and $P_{f,\alpha}$ is the image of P_α under $\tilde{f}$. We use Theorem 3.8 to prove tightness of $(P_{f,\alpha})_{\alpha\in I}$. Clearly, $(P_{f,\alpha})_{\alpha\in I}$ satisfies the compact containment condition (3.9), since f is bounded.
Let $T > 0$ and $0 < \delta \leq 1$. For every set $A \in \mathcal{F}_t^\mathbb{R}$, $0 \leq t \leq T$ and $0 \leq u \leq \delta$ we have

$$\begin{aligned} \int_A (X_{t+u}^\mathbb{R} - X_t^\mathbb{R})^2 \, dP_{f,\alpha} &= \int_{\tilde{f}^{-1}(A)} (X_{t+u}^\mathbb{R} \circ \tilde{f} - X_t^\mathbb{R} \circ \tilde{f})^2 \, dP_\alpha \\ &= \int_{\tilde{f}^{-1}(A)} (f(X_{t+u}) - f(X_t))^2 \, dP_\alpha \\ &= \int_{\tilde{f}^{-1}(A)} \mathbb{E}^{P_\alpha} \left[(f(X_{t+u}) - f(X_t))^2 | \mathcal{F}_t \right] \, dP_\alpha. \end{aligned} \tag{3.17}$$

But P_α is a solution for the martingale problem for A_α and $f, f^2 \in D$. Therefore

$$\begin{aligned} &\mathbb{E}^{P_\alpha} \left[(f(X_{t+u}) - f(X_t))^2 | \mathcal{F}_t \right] \\ &= \mathbb{E}^{P_\alpha} \left[f^2(X_{t+u}) - f^2(X_t) | \mathcal{F}_t \right] - 2 f(X_t) \, \mathbb{E}^{P_\alpha} \left[f(X_{t+u}) - f(X_t) | \mathcal{F}_t \right] \\ &= \mathbb{E}^{P_\alpha} \left[\int_t^{t+u} A_\alpha(f^2)(X_s) \, ds | \mathcal{F}_t \right] - 2 f(X_t) \, \mathbb{E}^{P_\alpha} \left[\int_t^{t+u} A_\alpha f(X_s) \, ds | \mathcal{F}_t \right] \\ &\leq \delta \left\| A_\alpha(f^2) \right\|_\infty + 2 \left\| f \right\|_\infty \cdot \delta \cdot \left\| A_\alpha f \right\|_\infty = c_f \cdot \delta \qquad P_\alpha\text{-a.s.} \end{aligned}$$

with a constant c_f independent of α. Thus by (3.17)

$$\mathbb{E}^{P_{f,\alpha}} \left[(X_{t+u}^\mathbb{R} - X_t^\mathbb{R})^2 | \mathcal{F}_t^\mathbb{R} \right] \leq c_f \cdot \delta.$$

In particular, conditions (3.11) and (3.13) of Theorem 3.8 hold true for $(X_t^\mathbb{R})_{t\geq 0}$ and $\gamma_\delta = c_f \cdot \delta$ and we conclude that $(P_{f,\alpha})_{\alpha\in I}$ is tight for all $f \in D$. □

3.2 The solution of the martingale problem for pseudo differential operators

We now turn to the concrete situation that the operator A is a pseudo differential operator. For that purpose let

$$p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$$

be a continuous negative definite symbol. Then by Theorem 2.18

$$-p(x, D) : C_0^\infty(\mathbb{R}^n) \to C(\mathbb{R}^n)$$

is a linear operator that satisfies the positive maximum principle. To prove existence of solution for the corresponding martingale problem, we want to solve the martingale problem for an approximating sequence of operators and then apply Theorem 3.10 to the so obtained sequence of probability measures. In order to do so, in principle we have to check the compact containment condition, but our strategy will be a little bit different. We first extend the operator $-p(x, D)$ to an operator defined on functions on the one-point compactification of $\mathbb{R}^n$. Then we get the compact containment condition for free. The difficulty of course is only postponed to a later stage and we will have to check in the end that the solution of the martingale problem does not approach the point at infinity in finite time.

Therefore denote by $\overline{\mathbb{R}^n} = \mathbb{R}^n \cup \{\Delta\}$ the one-point compactification of $\mathbb{R}^n$, Δ is the point at infinity. We equip $\overline{\mathbb{R}^n}$ with a metric d that induces the topology of $\overline{\mathbb{R}^n}$ such that $(\overline{\mathbb{R}^n}, d)$ is a separable and complete metric space.

Unfortunately, for a function $\varphi \in C_0^\infty(\mathbb{R}^n)$ it is in general not possible to extend $p(x, D)\varphi$ to $\overline{\mathbb{R}^n}$ in a continuous way, since the limit of $p(x, D)\varphi$ at infinity may not exist, even under boundedness condition on the symbol which are uniform with respect to x. The problem is due to the nonlocality of the pseudo differential operator $p(x, D)$. Unlike in the situation of a differential operator, which maps functions with compact support to functions with compact support, this is not the case for $-p(x, D)$. Looking to the Lévy-type representation (2.19) of $-p(x, D)$ we see that the nonlocality is governed by the Lévy-kernel $\mu(x, dy)$. Thus, in order to control the nonlocal behaviour we will cut off the Lévy-kernel by multiplication with a density θ which vanishes at infinity. On the level of the symbol this corresponds to a convolution-like operation. The remainder part of the symbol then corresponds to the Lévy-kernel $(1-\theta(y))\,\mu(x, dy)$, which consists of finite measures and thus can be handled as a bounded perturbation. The next proposition clarifies the situation.

Proposition 3.11. *Let $\theta \in \mathcal{S}(\mathbb{R}^n)$, $0 \le \theta \le 1$, $\theta(0) = 1$ and let $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ be a continuous negative definite symbol with Lévy–Khinchin representation*

$$p(x,\xi) = q(x,\xi) + i(b(x),\xi) + \int_{\mathbb{R}^n\setminus\{0\}} (1 - e^{i(y,\xi)} + \frac{i(y,\xi)}{1+|y|^2})\, \mu(x,dy), \tag{3.18}$$

where $q(x,\xi)$ for fixed x denotes the quadradic form. Define

$$p_1^\theta(x,\xi) = \int_{\mathbb{R}^n} (p(x,\xi+\eta) - p(x,\eta))\,\hat{\theta}(\eta)\, đ\eta$$

and

$$p_2^\theta(x,\xi) = \int_{\mathbb{R}^n} (p(x,\xi) - p(x,\xi+\eta) + p(x,\eta))\,\hat{\theta}(\eta)\, đ\eta.$$

Then $p_1^\theta, p_2^\theta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ are continuous negative definite symbols such that

$$p(x,\xi) = p_1^\theta(x,\xi) + p_2^\theta(x,\xi) \tag{3.19}$$

and the Lévy–Khinchin representations

$$p_1^\theta(x,\xi) = q(x,\xi) + i(b(x)+\tilde{b}(x),\xi) + \int_{\mathbb{R}^n\setminus\{0\}} (1 - e^{i(y,\xi)} + \frac{i(y,\xi)}{1+|y|^2})\,\theta(y)\, \mu(x,dy), \tag{3.20}$$

and

$$(3.21)\qquad p_2^\theta(x,\xi) = \int_{\mathbb{R}^n\setminus\{0\}} (1-e^{i(y,\xi)})\,(1-\theta(y))\,\mu(x,dy),$$

hold, where $\tilde b(x) = (\tilde b_1(x),\dots,\tilde b_n(x)) \in \mathbb{R}^n$ *is given by* $\tilde b_j(x) = \int_{\mathbb{R}^n\setminus\{0\}} (1-\theta(y))\frac{y_j}{1+|y|^2}\,\mu(x,dy)$.

Proof: By Theorem 2.7 we know that

$$p(x,\xi) \le c(1+|\xi|^2),$$

where the constant c can be chosen locally uniform with respect to x. Thus clearly p_1^θ and p_2^θ are continuous functions, which satisfy (3.19). Moreover, once (3.20) and (3.21) are established, we see immediately from the Lévy–Khinchin formula that p_1^θ and p_2^θ are negative definite symbols. Note that for $\xi = (\xi_1,\dots,\xi_n)$, $\eta = (\eta_1,\dots,\eta_n) \in \mathbb{R}^n$

$$\begin{aligned}&\int_{\mathbb{R}^n} ((\xi_j+\eta_j)(\xi_k+\eta_k) - \eta_j\eta_k)\,\hat\theta(\eta)\,đ\eta \\ &= \xi_j\xi_k\,\theta(0) - i\xi_j\,(\partial_{x_k}\theta)(0) - i\xi_k\,(\partial_{x_j}\theta)(0) = \xi_j\xi_k,\end{aligned}$$

since $\nabla\theta(0) = 0$, and

$$\int_{\mathbb{R}^n} ((\xi_j+\eta_j) - \eta_j)\,\hat\theta(\eta)\,đ\eta = \xi_j\,\theta(0) = \xi_j,$$

thus also

$$(3.22)\int_{\mathbb{R}^n} \{q(x,\xi+\eta) + i(b(x),\xi+\eta) - q(x,\eta) - i(b(x),\eta)\}\,\hat\theta(\eta)\,đ\eta = q(x,\xi) + i(b(x),\xi).$$

On the other hand by Lemma 2.8

$$\begin{aligned}&\left|(1-e^{i(y,\xi+\eta)} + \frac{i(y,\xi+\eta)}{1+|y|^2}) - (1-e^{i(y,\eta)} + \frac{i(y,\eta)}{1+|y|^2})\right| \\ &\le c\frac{|y|^2}{1+|y|^2}(1+|\xi+\eta|^2+|\eta|^2) \le c_\xi\frac{|y|^2}{1+|y|^2}(1+|\eta|^2).\end{aligned}$$

Therefore Fubini's theorem yields

$$\begin{aligned}&\int_{\mathbb{R}^n}\int_{\mathbb{R}^n\setminus\{0\}} \left((1-e^{i(y,\xi+\eta)} + \frac{i(y,\xi+\eta)}{1+|y|^2}) - (1-e^{i(y,\eta)} + \frac{i(y,\eta)}{1+|y|^2})\right)\mu(x,dy)\,\hat\theta(\eta)đ\eta \\ &= \int_{\mathbb{R}^n\setminus\{0\}}\int_{\mathbb{R}^n} \left(e^{i(y,\eta)} - e^{i(y,\xi)}e^{i(y,\eta)} + \frac{i(y,\xi)}{1+|y|^2}\right)\hat\theta(\eta)đ\eta\,\mu(x,dy) \\ &= \int_{\mathbb{R}^n\setminus\{0\}} \left(\theta(y) - e^{i(y,\xi)}\theta(y) + \frac{i(y,\xi)}{1+|y|^2}\right)\mu(x,dy) \\ &= \int_{\mathbb{R}^n\setminus\{0\}} \left(1-e^{i(y,\xi)} + \frac{i(y,\xi)}{1+|y|^2}\right)\theta(y)\,\mu(x,dy) + \int_{\mathbb{R}^n\setminus\{0\}} (1-\theta(y))\,\frac{i(y,\xi)}{1+|y|^2}\,\mu(x,dy).\end{aligned}$$

Note that the last integral exists, since $1-\theta(y)$ vanishes of order 2 in the origin, and this integral equals $i(\tilde{b}(x),\xi)$. Together with (3.22) this implies (3.20).
Finally by (3.19)

$$\begin{aligned} p_2^\theta(x,\xi) &= p(x,\xi)-p_1^\theta(x,\xi) \\ &= \int_{\mathbb{R}^n\setminus\{0\}} \left(1-e^{i(y,\xi)}+\frac{i(y,\xi)}{1+|y|^2}\right)(1-\theta(y))\,\mu(x,dy) - i(\tilde{b}(x),\xi) \\ &= \int_{\mathbb{R}^n\setminus\{0\}} \left(1-e^{i(y,\xi)}\right)(1-\theta(y))\,\mu(x,dy). \end{aligned}$$

that is (3.21). □

Our aim is to decompose the operator $p(x,D)$ by Proposition 3.11 into an operator $p_1^\theta(x,D)$, which extends to functions on $\overline{\mathbb{R}^n}$, and an operator $p_2^\theta(x,D)$, which has a uniformly bounded Lévy-kernel and which in the end will be handled as a perturbation as in Proposition 3.6. To do so, we choose θ in Proposition 3.11 to have compact support.

Theorem 3.12. *Let $p:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{C}$ be a continuous negative definite symbol such that $p(x,0)=0$ and let $\theta\in C_0^\infty(\mathbb{R}^n)$ be a cut-off function such that $0\le\theta\le 1$ and $\theta(0)=1$. Let*

$$p(x,\xi)=p_1^\theta(x,\xi)+p_2^\theta(x,\xi)$$

be the decomposition as in Proposition 3.11. Then $-p_1^\theta(x,D)$ maps $C_0^\infty(\mathbb{R}^n)$ into $C_0(\mathbb{R}^n)$ and $-p_2^\theta(x,D)$ is a Lévy-type operator of pure jump type

$$-p_2^\theta(x,D)\,u(x)=\int_{\mathbb{R}^n\setminus\{0\}}(u(x+y)-u(x))\,\tilde{\mu}(x,dy),\qquad u\in C_0^\infty(\mathbb{R}^n), \tag{3.23}$$

where the Lévy-kernel $\tilde{\mu}(x,dy)$ consists of finite measures, which satisfy

$$\tilde{\mu}(x,\mathbb{R}^n\setminus\{0\})\le c_\theta\int_{\mathbb{R}^n} Re\,p(x,\xi)\,\nu(d\xi)<\infty.$$

Here ν is the finite measure defined in Lemma 2.15.

Proof: We know that $p_1^\theta(x,\xi)$ is continuous negative definite symbol und thus $p_1^\theta(x,D)\varphi$ is continuous for all $\varphi\in C_0^\infty(\mathbb{R}^n)$ by Theorem 2.18. To prove the first statement it remains to show that $p_1^\theta(x,D)\varphi$ has compact support. The Lévy-type representation of $-p_1^\theta(x,D)$ is given by a diffusion operator, which is local and preserves the compact support, plus an integro-differential part. Thus for $x\notin\operatorname{supp}\varphi$ by Proposition 3.11

$$\begin{aligned} -p_1^\theta(x,D)\varphi(x) &= \int_{\mathbb{R}^n\setminus\{0\}}\left(\varphi(x+y)-\varphi(x)-\frac{(y,\nabla\varphi(x))}{1+|y|^2}\right)\theta(y)\,\mu(x,dy) \\ &= \int_{\mathbb{R}^n\setminus\{0\}}\varphi(x+y)\,\theta(y)\,\mu(x,dy), \end{aligned}$$

where $\mu(x,dy)$ is the Lévy-kernel of $p(x,\xi)$. But this expression vanishes for $|x|$ sufficiently large, since then $\operatorname{supp}\varphi(x+\cdot)\cap\operatorname{supp}\theta=\emptyset$.

By (3.21) p_2^θ is given by

$$p_2^\theta(x,\xi) = \int_{\mathbb{R}^n\setminus\{0\}} (1-e^{i(y,\xi)})(1-\theta(y))\,\mu(x,dy)$$

and thus as in the proof of Theorem 2.16 we find (3.23) with a Lévy-kernel $\tilde{\mu}(x,dy) = (1-\theta(y))\,\mu(x,dy)$. Because $1-\theta$ vanishes of order 2 in the origin and is bounded, we have with a suitable constant $c_\theta \geq 0$

$$1-\theta(y) \leq c_\theta \frac{|y|^2}{1+|y|^2}$$

and therefore by Lemma 2.15

$$\begin{aligned}\tilde{\mu}(x,\mathbb{R}^n\setminus\{0\}) &\leq c_\theta \int_{\mathbb{R}^n\setminus\{0\}} \frac{|y|^2}{1+|y|^2}\,\mu(x,dy) = \int_{\mathbb{R}^n\setminus\{0\}}\int_{\mathbb{R}^n}(1-\cos(y,\xi))\,\nu(d\xi)\,\mu(x,dy)\\ &\leq c_\theta \int_{\mathbb{R}^n} \operatorname{Re} p(x,\xi)\,\nu(d\xi).\end{aligned}$$

□

Let $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ be a continuous negative definite symbol and θ as in the above theorem. We consider $C_\infty(\mathbb{R}^n)$ as a subspace of $C(\overline{\mathbb{R}^n})$ with the convention $\varphi(\Delta) = 0$ for all $\varphi \in C_\infty(\mathbb{R}^n)$.

Define an operator A_θ in the Banach space $C(\overline{\mathbb{R}^n})$ with domain

(3.24) $$D(A_\theta) = \{\varphi \in C(\overline{\mathbb{R}^n}) : (\varphi - \varphi(\Delta)) \in C_0^\infty(\mathbb{R}^n)\}$$

by

(3.25) $$A_\theta\varphi(x) = \begin{cases} -p_1^\theta(x,D)\,[(\varphi-\varphi(\Delta))|_{\mathbb{R}^n}]\,(x) & \text{if } x \in \mathbb{R}^n,\\ 0 & \text{if } x = \Delta.\end{cases}$$

Then by Theorem 3.12 A_θ maps $D(A_\theta)$ into $C(\overline{\mathbb{R}^n})$, hence A_θ actually is an operator in $C(\overline{\mathbb{R}^n})$, which extends $-p_1^\theta(x,D)$. Moreover A_θ is densely defined, $1 \in D(A_\theta)$ and $A_\theta 1 = 0$. Finally A_θ satisfies the positive maximum principle on $D(A_\theta)$, since $-p_1^\theta(x,D)$ does.

For operators of this type J.P. Roth has shown that they can be approximated by Lévy-type operators with bounded Lévy-kernel, see [76], Prop. I.2.3 (see also [17], 4.5.3):

Theorem 3.13. *Let A_θ be as above. Then there are kernels of probability measures $\mu_k(x,dy)$ on $\overline{\mathbb{R}^n}$, $k \in \mathbb{N}$, such that*

(3.26) $$k\int_{\overline{\mathbb{R}^n}} (\varphi(y)-\varphi(x))\,\mu_k(x,dy) \underset{k\to\infty}{\longrightarrow} A_\theta\varphi(x)$$

uniformly on $\overline{\mathbb{R}^n}$ for all $\varphi \in D(A_\theta)$.

The proof mimics the Yosida approximation of a generator L by its resolvent (R_λ):

$$Lu = \lim_{k\to\infty} k(kR_k - \mathrm{Id})u, \qquad u \in D(L).$$

Proof: Since A_θ satisfies the positive maximum principle, it is dissipative in $C(\overline{\mathbb{R}^n})$, see [17], 4.2.1. Therefore, $k - A_\theta$ is continuously invertible in $C(\overline{\mathbb{R}^n})$, but the inverse is not necessarily densely defined. For $k \in \mathbb{N}$ and $x \in \overline{\mathbb{R}^n}$ define a linear functional T_k^x on the range $\mathrm{R}(k - A_\theta)$ by

$$T_k^x \varphi = k(k - A_\theta)^{-1}\varphi(x).$$

Then $|T_k^x \varphi| \le \|\varphi\|_\infty$ by the dissipativity of A_θ. Moreover $1 = (k - A_\theta)\frac{1}{k} \in \mathrm{R}(k - A_\theta)$ and $T_k^x 1 = 1$. Thus for $\varphi \ge 0$

$$T_k^x \varphi = T_k^x \|\varphi\|_\infty + T_k^x(\varphi - \|\varphi\|_\infty) \ge \|\varphi\|_\infty - \|(\varphi - \|\varphi\|_\infty)\|_\infty \ge 0.$$

Hence T_k^x is a positive linear functional of norm 1 on $\mathrm{R}(k - A_\theta)$, which extends by the Hahn-Banach theorem to a (not uniquely determined) positive linear functional of norm 1 on $C(\overline{\mathbb{R}^n})$. Thus by the Riesz representation theorem there is a probability measure $\mu_k^x \in \mathcal{M}_1(\overline{\mathbb{R}^n})$ such that

$$T_k^x \varphi = \int_{\overline{\mathbb{R}^n}} \varphi(y)\, \mu_k^x(dy) \qquad \text{for all } \varphi \in \mathrm{R}(k - A_\theta).$$

Let $M_k^x = \{\mu \in \mathcal{M}_1(\overline{\mathbb{R}^n}) : T_k^x \varphi = \int_{\overline{\mathbb{R}^n}} \varphi(y)\, \mu(dy) \quad \text{for all } \varphi \in \mathrm{R}(k - A_\theta)\} \neq \emptyset$.
As usual we equip $\mathcal{M}_1(\overline{\mathbb{R}^n})$ with the weak topology. For fixed $k \in \mathbb{N}$ we like to construct a kernel $\mu_k(x, dy)$ on $\overline{\mathbb{R}^n}$ such that for each $x \in \overline{\mathbb{R}^n}$ we have $\mu_k(x, \cdot) \in M_k^x$. This measurable selection is possible by the selection theorem of K. Kuratowski and C. Ryll-Nardzewski [55], see also C.J. Himmelberg [26], Theorem 3.5, provided $M_k^x \subset \mathcal{M}_1(\overline{\mathbb{R}^n})$ is closed for all $x \in \overline{\mathbb{R}^n}$ and the set $\{x \in \overline{\mathbb{R}^n} : M_k^x \cap C \neq \emptyset\} \subset \overline{\mathbb{R}^n}$ is measurable for all closed subsets $C \subset \mathcal{M}_1(\overline{\mathbb{R}^n})$. The first property is clear from the definition of M_k^x and the properties of weak convergence. To see the second, let $(x_l)_{l \in \mathbb{N}}$ be a sequence in $\overline{\mathbb{R}^n}$ such that there are $\mu_l \in M_k^{x_l} \cap C$ and $x_l \underset{l \to \infty}{\longrightarrow} x \in \overline{\mathbb{R}^n}$. Since $\mathcal{M}_1(\overline{\mathbb{R}^n})$ is compact with respect to the weak topology, there is a subsequence of $(\mu_l)_{l \in \mathbb{N}}$ that converges to a measure $\mu_\infty \in C$. Moreover $\mu_\infty \in M_k^x$, because for all $\varphi \in \mathrm{R}(k - A_\theta)$

$$T_k^x \varphi = k(k - A_\theta)^{-1}\varphi(x) = \lim_{l \to \infty} k(k - A_\theta)^{-1}\varphi(x_l) = \lim_{l \to \infty} T_k^{x_l}\varphi = \lim_{l \to \infty} \int_{\overline{\mathbb{R}^n}} \varphi\, d\mu_l = \int_{\overline{\mathbb{R}^n}} \varphi\, d\mu_\infty.$$

Thus $\{x \in \overline{\mathbb{R}^n} : M_k^x \cap C \neq \emptyset\}$ is even closed and hence measurable.

Having chosen a kernel $\mu_k(x, dy)$ as above it remains to show (3.26). For $\varphi \in D(A_\theta)$

$$\begin{aligned}
\int_{\overline{\mathbb{R}^n}} \varphi(y)\, \mu_k(x, dy) &= \frac{1}{k} \int_{\overline{\mathbb{R}^n}} (k - A_\theta)\varphi(y)\, \mu_k(x, dy) + \frac{1}{k} \int_{\overline{\mathbb{R}^n}} A_\theta \varphi(y)\, \mu_k(x, dy) \\
&= \frac{1}{k} T_k^x[(k - A_\theta)\varphi] + \frac{1}{k} \int_{\overline{\mathbb{R}^n}} A_\theta \varphi(y)\, \mu_k(x, dy) \\
&= \varphi(x) + \frac{1}{k} \int_{\overline{\mathbb{R}^n}} A_\theta \varphi(y)\, \mu_k(x, dy).
\end{aligned}$$

But $A_\theta \varphi$ is bounded, so we have

$$\text{(3.27)} \qquad \int_{\overline{\mathbb{R}^n}} \varphi(y)\, \mu_k(x, dy) \underset{k \to \infty}{\longrightarrow} \varphi(x) \qquad \text{uniformly on } \overline{\mathbb{R}^n}.$$

Moreover, since $D(A_\theta)$ is dense in $C(\overline{\mathbb{R}^n})$ with respect to uniform convergence, (3.27) holds true for all $\varphi \in C(\overline{\mathbb{R}^n})$. Finally we have for $\varphi \in D(A_\theta)$ as above

$$\begin{aligned} k\int_{\overline{\mathbb{R}^n}} (\varphi(y)-\varphi(x))\,\mu_k(x,dy) &= k\int_{\overline{\mathbb{R}^n}} \varphi(y)\,\mu_k(x,dy) - k\varphi(x) \\ &= \int_{\overline{\mathbb{R}^n}} A_\theta\varphi(y)\,\mu_k(x,dy), \end{aligned}$$

which tends uniformly to $A_\theta\varphi$ by (3.27) as $k\to\infty$, that is (3.26) holds. □

We now solve the martingale problem for the operator A_θ

Proposition 3.14. *Let A_θ be as above. Then for any initial distribution $\nu \in \mathcal{M}_1(\overline{\mathbb{R}^n})$ there is a solution $P \in \mathcal{M}_1(D_{\overline{\mathbb{R}^n}})$ of the $D_{\overline{\mathbb{R}^n}}$-martingale problem for A_θ.*

Proof: Let A_k be the operators defined by

$$A_k\varphi = k\int_{\overline{\mathbb{R}^n}} (\varphi(y)-\varphi(x))\,\mu_k(x,dy), \qquad \varphi \in D(A_\theta),$$

where $\mu_k(x,dy)$ are the kernels of probability measures defined in Theorem 3.13. By Proposition 3.5 for all $k\in\mathbb{N}$ there is a solution $P_k \in \mathcal{M}_1(D_{\overline{\mathbb{R}^n}})$ of the martingale problem for A_k with initial distribution ν. Moreover, since $A_k\varphi \underset{k\to\infty}{\longrightarrow} A_\theta\varphi$ uniformly for all $\varphi\in D(A_\theta)$, we have

$$\sup_{k\in\mathbb{N}} \|A_k\varphi\|_\infty < \infty$$

for all $\varphi \in D(A_\theta)$. Thus, since $\overline{\mathbb{R}^n}$ is compact, the compact containment condition is fulfilled by $(P_k)_{k\in\mathbb{N}}$ and Theorem 3.10 implies that a subsequence of (P_k) converges to a probability measure $P \in \mathcal{M}_1(\overline{\mathbb{R}^n})$. We claim that P solves the martingale problem for A_θ with initial distribution ν.

To simplify the notation we denote the subsequence again by $(P_k)_{k\in\mathbb{N}}$. Because the position map $X_0 : D_{\overline{\mathbb{R}^n}} \to \overline{\mathbb{R}^n}$ is continuous, this implies

$$P_k \circ X_0^{-1} \underset{k\to\infty}{\longrightarrow} P\circ X_0^{-1} \qquad \text{in } \mathcal{M}_1(\overline{\mathbb{R}^n}),$$

hence $P\circ X_0^{-1} = \nu$. Finally we have to check that for $\varphi\in D(A_\theta)$

$$\varphi(X_t) - \int_0^t A_\theta\varphi(X_u)\,du, \qquad t\ge 0$$

is a martingale under P. That amounts to show that

$$\mathbb{E}^P\left[\left(\varphi(X_{t_2}) - \varphi(X_{t_1}) - \int_{t_1}^{t_2} A_\theta\varphi(X_u)\,du\right)\cdot \prod_{m=1}^M h_m(X_{s_m})\right] = 0$$

for all $0\le s_1 < \ldots < s_M \le t_1 \le t_2$ and all $h_m \in C(\overline{\mathbb{R}^n})$. Let T_P be the dense subset of $[0,\infty)$ defined in Proposition 3.2. By the right-continuity of the paths it suffices to consider

$s_1, \ldots, s_M, t_1, t_2 \in T_P$. Then the distributions of $(X_{s_1}, \ldots, X_{s_M}, X_{t_1}, X_{t_2})$ under P_k converge weakly to that under P and, since P_k is a solution of the martingale problem for A_k,

$$\mathbb{E}^{P_k}\left[\left(\varphi(X_{t_2}) - \varphi(X_{t_1}) - \int_{t_1}^{t_2} A_k\varphi(X_u)\,du\right) \cdot \prod_{m=1}^{M} h_m(X_{s_m})\right] = 0.$$

Hence, it is left to show

$$\begin{aligned} \lim_{k\to\infty} \mathbb{E}^{P_k}\left[\int_{t_1}^{t_2} A_k\varphi(X_u)\,du \cdot \prod_{m=1}^{M} h_m(X_{s_m})\right] \\ = \mathbb{E}^{P}\left[\int_{t_1}^{t_2} A_\theta\varphi(X_u)\,du \cdot \prod_{m=1}^{M} h_m(X_{s_m})\right]. \end{aligned} \tag{3.28}$$

$(A_k\varphi)_{k\in\mathbb{N}}$ is uniformly bounded and converges uniformly to $A_\theta\varphi$. Thus, if we interchange the order of integration in (3.28), it is enough to show that

$$\lim_{k\to\infty} \mathbb{E}^{P_k}\left[A_k\varphi(X_u) \cdot \prod_{m=1}^{M} h_m(X_{s_m})\right] = \mathbb{E}^{P}\left[A_\theta\varphi(X_u) \cdot \prod_{m=1}^{M} h_m(X_{s_m})\right]$$

for all $u \in T_P$. But this is evident, because

$$\lim_{k\to\infty} \mathbb{E}^{P_k}\left[A_\theta\varphi(X_u) \cdot \prod_{m=1}^{M} h_m(X_{s_m})\right] = \mathbb{E}^{P}\left[A_\theta\varphi(X_u) \cdot \prod_{m=1}^{M} h_m(X_{s_m})\right]$$

by the weak convergence of (P_k) and

$$\lim_{k\to\infty} \mathbb{E}^{P_k}\left[(A_\theta\varphi(X_u) - A_k\varphi(X_u)) \cdot \prod_{m=1}^{M} h_m(X_{s_m})\right] = 0$$

by the uniform convergence of $(A_k\varphi)$. This completes the proof. □

The main result of this section now reads as follows.

Theorem 3.15. *Let $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ be a continuous negative definite symbol such that $p(x, 0) = 0$ for all $x \in \mathbb{R}^n$. Assume that*

$$|p(x,\xi)| \le c(1 + |\xi|^2), \qquad x \in \mathbb{R}^n,\ \xi \in \mathbb{R}^n. \tag{3.29}$$

Then for all initial distributions $\mu \in \mathcal{M}_1(\mathbb{R}^n)$ there is a solution of the $D_{\mathbb{R}^n}$-martingale problem for $-p(x, D)$.

Remark: We know that (3.29) is satisfied with a constant which is locally uniform with respect to x. The global bound thus should be considered as a generalization of a boundedness condition for the coefficient of a diffusion operator. Note also that (3.29) implies

$$p(x, D) : C_0^\infty(\mathbb{R}^n) \to C_b(\mathbb{R}^n).$$

To prove Theorem 3.15 we need the following lemma.

Lemma 3.16. *Suppose $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ is a continuous negative definite symbol that satisfies $p(x, 0) = 0$ and (3.29). Then there is a sequence $(\varphi_k)_{k\in\mathbb{N}}$ of $C_0^\infty(\mathbb{R}^n)$-functions, such that $(\varphi_k)_{k\in\mathbb{N}}$ as well as $(p(x, D)\varphi_k)_{k\in\mathbb{N}}$ are uniformly bounded and converge pointwise in $\mathbb{R}^n$ to 1 and 0, respectively, as $k \to \infty$.*

Proof: Choose $\varphi \in C_0^\infty(\mathbb{R}^n)$, $0 \le \varphi \le 1$, $\operatorname{supp}\varphi \subset B_1(0)$, $\varphi|_{B_{\frac{1}{2}}(0)} = 1$ and define $\varphi_k(x) = \varphi(\frac{x}{k})$. Then $(\varphi_k)_{k\in\mathbb{N}}$ is uniformly bounded, converges pointwise to 1 and, since $\hat{\varphi} \in \mathcal{S}(\mathbb{R}^n)$ we have

$$\begin{aligned}
\sup_{x\in\mathbb{R}^n} |p(x, D)\varphi_k(x)| &\le \sup_{x\in\mathbb{R}^n} \int_{\mathbb{R}^n} |p(x, \xi)|\, |\hat{\varphi}_k(\xi)|\, đ\xi \\
&\le c \int_{\mathbb{R}^n} (1 + |\xi|^2) k^n\, |\hat{\varphi}(k\xi)|\, đ\xi \\
&\le c \cdot \sup_{|\xi|\le 1} (1 + |\xi|^2) \cdot \int_{|\xi|\le 1} k^n\, |\hat{\varphi}(k\xi)|\, đ\xi + c \int_{|\xi|>1} (1 + |\xi|^2) k^n\, |\hat{\varphi}(k\xi)|\, đ\xi \\
&\le c \int_{\mathbb{R}^n} k^n\, |\hat{\varphi}(k\xi)|\, đ\xi + c \int_{|\xi|>1} |\xi|^2\, k^n\, |k\xi|^{-(n+3)}\, đ\xi \\
&\le c \int_{\mathbb{R}^n} |\hat{\varphi}(\xi)|\, đ\xi + \frac{c}{k^3} \int_{|\xi|>1} |\xi|^{-(n+1)}\, đ\xi
\end{aligned}$$

and this is bounded as $k \to \infty$. Similarly we find for each fixed $x_0 \in \mathbb{R}^n$

$$|p(x, D)\varphi_k(x_0)| \le \int_{\mathbb{R}^n} |p(x_0, \xi)|\, k^n\, |\hat{\varphi}(k\xi)|\, đ\xi \tag{3.30}$$

$$\begin{aligned}
&\le \sup_{|\xi|\le \frac{1}{\sqrt{k}}} |p(x_0, \xi)| \cdot \int_{|\xi|\le \frac{1}{\sqrt{k}}} k^n\, |\hat{\varphi}(k\xi)|\, đ\xi + c \int_{|\xi|> \frac{1}{\sqrt{k}}} (1 + |\xi|^2) k^n\, |\hat{\varphi}(k\xi)|\, đ\xi \\
&\le \sup_{|\xi|\le \frac{1}{\sqrt{k}}} |p(x_0, \xi)| \cdot \int_{\mathbb{R}^n} |\hat{\varphi}(\xi)|\, đ\xi + c \int_{\frac{1}{\sqrt{k}}<|\xi|\le 1} k^n\, |k\xi|^{-(n+3)}\, đ\xi + c \int_{|\xi|>1} |\xi|^2\, k^n\, |k\xi|^{-(n+3)}\, đ\xi \\
&\le c \left(\sup_{|\xi|\le \frac{1}{\sqrt{k}}} |p(x_0, \xi)| + \frac{1}{k^3} \left(k^{3/2} - 1\right) + \frac{1}{k^3} \right)
\end{aligned}$$

and this tends to 0 as $k \to \infty$, since $p(x_0, 0) = 0$. □

Proof of Theorem 3.15: We decompose $p(x, \xi)$ as in Theorem 3.12:

$$p(x, \xi) = p_1^\theta(x, \xi) + p_2^\theta(x, \xi).$$

Denote again by A_θ the extension of $-p_1^\theta(x, D) : C_0^\infty(\mathbb{R}^n) \to C_0(\mathbb{R}^n)$ defined by (3.24), (3.25). Then by Proposition 3.14 there is a solution $P \in \mathcal{M}_1(D_{\overline{\mathbb{R}^n}})$ of the $D_{\overline{\mathbb{R}^n}}$-martingale problem for A_θ and the initial distribution $\mu \in \mathcal{M}_1(\mathbb{R}^n) \subset \mathcal{M}_1(\overline{\mathbb{R}^n})$. We show that P-a.s. the paths of the solution are in $D_{\mathbb{R}^n}$.
Define the $\{\mathcal{F}_{t+}\}$-stopping times

$$\tau_m = \inf\{t \ge 0 : d(\Delta, X_t) < \frac{1}{m}\}, \qquad m \in \mathbb{N},$$

where d is the metric $\overline{\mathbb{R}^n}$. It is sufficient to show that for all $T > 0$ and $Z_T = \lim_{m\to\infty} \overline{X}_{\tau_m \wedge T}$ we have $P(Z_T \in \mathbb{R}^n) = 1$.

For this purpose consider the sequence of $C_0^\infty(\mathbb{R}^n)$-functions $(\varphi_k)_{k\in\mathbb{N}}$ of Lemma 3.16 and let $\varphi_k(\Delta) = 0$ as usual. Then the sequences (φ_k) and $(A_\theta \varphi_k)$ of functions on $\overline{\mathbb{R}^n}$ are uniformly bounded and tend pointwise to $1_{\mathbb{R}^n}$ and 0, respectively, as $k \to \infty$. Moreover

$$\varphi_k(X_t) - \int_0^t A_\theta \varphi_k(X_u)\, du, \qquad t \geq 0,$$

is an $\{\mathcal{F}_{t+}\}$-martingale by right-continuity of the paths and optional sampling yields

$$\mathbb{E}^P \left[\varphi_k(X_{\tau_m \wedge T}) - \int_0^{\tau_m \wedge T} A_\theta \varphi_k(X_u)\, du \right] = \mathbb{E}^P \left[\varphi_k(X_0) \right], \qquad k, m \in \mathbb{N}.$$

Thus for $m \to \infty$

$$\mathbb{E}^P \left[\varphi_k(Z_T) \right] - \mathbb{E}^P \left[\int_0^{\sup_{m\in\mathbb{N}}(\tau_m \wedge T)} A_\theta \varphi_k(X_u)\, du \right] = \mathbb{E}^P \left[\varphi_k(X_0) \right]$$

and for $k \to \infty$

$$\mathbb{E}^P \left[1_{\mathbb{R}^n}(Z_T) \right] = \mathbb{E}^P \left[1_{\mathbb{R}^n}(X_0) \right] = \mu(\mathbb{R}^n) = 1$$

and we may assume that there is a solution to the martingale problem with sample paths in $D_{\mathbb{R}^n}$. This means the $D_{\mathbb{R}^n}$-martingale problem for $-p_1^\theta(x, D)$ is solvable.

To complete the proof it is enough to note that by Theorem 3.12 the operator $-p_2^\theta(x, D)$ is a pure jump Lévy-type operator

$$-p_2^\theta(x, D)\varphi(x) = \int_{\mathbb{R}^n} (\varphi(x+y) - \varphi(x))\, \tilde{\mu}(x, dy) = \int_{\mathbb{R}^n} (\varphi(y) - \varphi(x))\, \tilde{\mu}(x, dy - x), \qquad \varphi \in C_0^\infty(\mathbb{R}^n).$$

The Lévy-kernel $\tilde{\mu}$ consists of uniformly bounded measures, since by (3.29), Theorem 3.12 and Lemma 2.15

$$\sup_{x\in\mathbb{R}^n} \|\tilde{\mu}(x, \cdot)\|_\infty \leq c_\theta \sup_{x\in\mathbb{R}^n} \int_{\mathbb{R}^n} \operatorname{Re} p(x, \xi)\, \nu(d\xi) \leq c \int_{\mathbb{R}^n} (1 + |\xi|^2)\, \nu(d\xi) < \infty.$$

Here ν denotes the measure defined in Lemma 2.15. Hence $-p_2^\theta(x, D)$ is a perturbation in the sense of Proposition 3.6 and the $D_{\mathbb{R}^n}$-martingale problem for $-p(x, D) = -p_1^\theta(x, D) - p_2^\theta(x, D)$ is solvable for all initial distributions. □

Chapter 4

Generators of Feller semigroups

4.1 Technical preliminaries

A diffusion process can be regarded as a certain perturbation of Brownian motion, i.e of the Lévy-process corresponding to the continuous negative definite function $\xi \mapsto |\xi|^2$. Now the philosophy for all the subsequent will be the following:

We fix a continuous negative definite function ψ as a reference function and we consider continuous negative definite symbols $p(x, \xi)$ which for fixed x are compatible with ψ in a suitable sense. This means that we take the Lévy-process corresponding to ψ as a reference process and we investigate jump processes which are in a certain sense comparable with this given Lévy-process.
Here and in the following we will limit our examinations to the case of real-valued continuous negative definite symbols. Therefore let

(4.1) $$\psi : \mathbb{R}^n \to \mathbb{R}$$

be a fixed continuous negative definite reference function. It turns out that we will need a mild no-degeneracy condition on ψ, namely a minimal growth condition for $|\xi| \to \infty$. More precisely, we assume that there is a constant $r > 0$, possibly very small, such that

(4.2) $$\psi(\xi) \geq c\,|\xi|^r \qquad \text{for } |\xi| \geq 1$$

with a suitable constant $c > 0$. In the following we will always state estimates for the symbols in terms of the function

(4.3) $$\lambda(\xi) = (1 + \psi(\xi))^{1/2}$$

instead of the function ψ itself, because this often turns out to be more convenient. Clearly (4.2) implies

(4.4) $$\lambda(\xi) \geq c\,|\xi|^{r/2}$$

for some $c > 0$. Note that by (2.8) we always have $r \leq 2$ and the case $r = 2$ is attained for the continuous negative definite reference function $|\xi|^2$ which corresponds to Brownian motion and diffusion processes.

A very useful tool to investigate pseudo differential operators with symbols compatible with the reference function ψ (or λ) will be an appropriate scale of Sobolev spaces, which are defined in termes of the function λ: For $s \in \mathbb{R}$ we define an **anisotropic Sobolev space** by

$$H^{s,\lambda}(\mathbb{R}^n) = \{u \in \mathcal{S}'(\mathbb{R}^n) : \int_{\mathbb{R}^n} \lambda^{2s}(\xi) \, |\hat{u}|^2 \, đ\xi < \infty\} \tag{4.5}$$

equipped with the norm

$$\|u\|_{s,\lambda} = \left(\int_{\mathbb{R}^n} \lambda^{2s}(\xi) \, |\hat{u}|^2 \, đ\xi \right)^{1/2}. \tag{4.6}$$

These anisotropic Sobolev spaces in the context of continuous negative definite functions and pseudo differential operators where first introduced by Jacob, see [42] and the references there in, but see also [4]. It is easy to see that $H^{s,\lambda}(\mathbb{R}^n)$ is a Hilbert space with inner product

$$(u, v)_{s,\lambda} = \int_{\mathbb{R}^n} \lambda^{2s}(\xi) \, \hat{u}(\xi) \, \overline{\hat{v}(\xi)} \, đ\xi. \tag{4.7}$$

The continuous embeddings

$$H^{s_2,\lambda}(\mathbb{R}^n) \hookrightarrow H^{s_1,\lambda}(\mathbb{R}^n), \qquad s_2 \geq s_1 \tag{4.8}$$

hold and $C_0^\infty(\mathbb{R}^n)$ is dense in $H^{s,\lambda}(\mathbb{R}^n)$ for all $s \in \mathbb{R}$. In particular $H^{0,\lambda}(\mathbb{R}^n) = L^2(\mathbb{R}^n)$ and therefore we mostly denote the norm and the inner product in $L^2(\mathbb{R}^n)$ by $\|\cdot\|_0$ and $(\cdot,\cdot)_0$.

In the special case $\psi(\xi) = |\xi|^2$ we recover the ordinary fractional L^2-Sobolev spaces: $H^{s,(1+|\xi|^2)^{1/2}}(\mathbb{R}^n) = H^s(\mathbb{R}^n)$. The spaces $H^{s,\lambda}(\mathbb{R}^n)$ thus can be regarded as a generalization of the usual Sobolev spaces to an in general anisotropic situation. If we identify $L^2(\mathbb{R}^n)$ with its dual space, $H^{s,\lambda}(\mathbb{R}^n)'$ becomes canonically isomorphic to $H^{-s,\lambda}(\mathbb{R}^n)$ and

$$\|u\|_{-s,\lambda} = \sup_{\substack{v \in C_0^\infty(\mathbb{R}^n) \\ v \neq 0}} \frac{|(u,v)_0|}{\|v\|_{s,\lambda}}, \tag{4.9}$$

where we extend the $L^2(\mathbb{R}^n)$-inner product $(\cdot,\cdot)_0$ in the obvious way.

Note that by (4.4) and (2.8) we have

$$c_1(1 + |\xi|^2)^{r/4} \leq \lambda(\xi) \leq c_2(1 + |\xi|^2)^{1/2},$$

which implies

$$H^s(\mathbb{R}^n) \hookrightarrow H^{s,\lambda}(\mathbb{R}^n) \hookrightarrow H^{\frac{r}{2}\cdot s}(\mathbb{R}^n) \tag{4.10}$$

for all $s \geq 0$. In particular, since $H^s(\mathbb{R}^n) \hookrightarrow C_\infty(\mathbb{R}^n)$ for $s > \frac{n}{2}$, see [54], Lemma 3.2.5, we have the following Sobolev embedding.

Proposition 4.1 *Assume (4.4). Then*

$$H^{s,\lambda}(\mathbb{R}^n) \hookrightarrow C_\infty(\mathbb{R}^n)$$

if $s > \frac{n}{r}$.

Moreover note that for all $\varepsilon > 0$ and $s_3 > s_2 > s_1$ the elementary estimate

$$\lambda^{2s_2}(\xi) \leq \varepsilon^2 \, \lambda^{2s_3}(\xi) + c(\varepsilon)^2 \, \lambda^{2s_1}(\xi)$$

implies

$$\|u\|_{s_2,\lambda} \leq \varepsilon \, \|u\|_{s_3,\lambda} + c(\varepsilon) \, \|u\|_{s_1,\lambda} \, , \qquad u \in H^{s_3,\lambda}(\mathbb{R}^n). \tag{4.11}$$

Finally, let us remark that by definition

$$\|u\|_{t,\lambda} = \left\| \lambda^t(D) \, u \right\|_0 \qquad \text{for all } u \in H^{t,\lambda}(\mathbb{R}^n) \tag{4.12}$$

or more general

$$\|u\|_{s+t,\lambda} = \left\| \lambda^t(D) \, u \right\|_{s,\lambda} \qquad \text{for all } u \in H^{s+t,\lambda}(\mathbb{R}^n) \tag{4.13}$$

for $s, t \in \mathbb{R}$. In this context it is natural to replace the usual notion of the order of an operator by the one given by the anisotropic Sobolev spaces $H^{s,\lambda}(\mathbb{R}^n)$. Then (4.13) just means that $\lambda^t(D)$ can be regarded as an operator of order t between these Sobolev spaces $H^{s,\lambda}(\mathbb{R}^n)$, which should be regarded as a generalization of an elliptic operator.

For a general pseudo differential operator $p(x, D)$ with a negative definite symbol defined in terms of the given reference function $\psi(\xi)$ it now reasonable to expect a similar behaviour. In fact, estimates for $p(x, D)$ in the Sobolev spaces $H^{s,\lambda}(\mathbb{R}^n)$ will play an central rôle in order to handle the operator by an analytic approach. We start with the following easy, but essential lemma.

Lemma 4.2. *Let $M \in \mathbb{N}$ and $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be M-times continuously differentable with repect to the first variable. Moreover assume that for every $\beta \in \mathbb{N}_0^n$, $|\beta| \leq M$ there is a function $\Phi_\beta \in L^1(\mathbb{R}^n)$ such that*

$$\left| \partial_x^\beta p(x, \xi) \right| \leq \Phi_\beta(x) \cdot \lambda^2(\xi), \qquad x \in \mathbb{R}^n, \ \xi \in \mathbb{R}^n \tag{4.14}$$

holds for all $|\beta| \leq M$. Let

$$\hat{p}(\eta, \xi) = \int_{\mathbb{R}^n} e^{-i(x,\xi)} p(x, \xi) \, dx, \qquad \eta \in \mathbb{R}^n,$$

be the Fourier transform of $p(x, \xi)$ with respect to x. Then there is a constant $C_M > 0$ depending only on M and the space dimension such that

$$|\hat{p}(\eta, \xi)| \leq C_M \cdot \sum_{|\beta| \leq M} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \, \langle \eta \rangle^{-M} \cdot \lambda^2(\xi), \tag{4.15}$$

where we have used the notation $\langle \eta \rangle = (1 + |\eta|^2)^{1/2}$.

Proof: Choose C_M such that

$$\langle \eta \rangle^M \leq C_M \cdot \sum_{|\beta| \leq M} \left| \eta^\beta \right| . \tag{4.16}$$

For $\beta \in \mathbb{N}_0^n$, $|\beta| \leq M$ we have by partial integration

$$\begin{aligned} \left|\eta^\beta \cdot \hat{p}(\eta,\xi)\right| &= \left|\eta^\beta \int_{\mathbb{R}^n} e^{-i(x,\eta)} p(x,\xi)\,dx\right| \\ &= \left|\int_{\mathbb{R}^n} e^{-i(x,\eta)} \partial_x^\beta p(x,\xi)\,dx\right| \\ &\leq \int_{\mathbb{R}^n} |\Phi_\beta(x)| \cdot \lambda^2(\xi)\,dx \\ &= \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \cdot \lambda^2(\xi). \end{aligned}$$

Summing up this estimate for all $|\beta| \leq M$ and using (4.16) proves the assertion □

As usual we denote by $[A, B] = AB - BA$ the commutator of two operators. The main auxiliary tool to treat the variable coefficient case, i.e. the case of symbols that depend on x, will be certain commutator estimates. In the case of two differential operators of order m_1 and m_2, respectively, it is well-known that their commutator is of order $m_1 + m_2 - 1$ and hence the order of the commutator is strictly less than the sum of the order of both operators. This result carries over to pseudo differential operators in classical symbol classes. It is this behaviour of a commutator which allows to treat the effect of variable coefficients as a lower order perturbation in an appropriate sense. But in the case of negative definite symbols considered here the standard symbolic calculus for pseudo differential operators is not available so the arguments used for classical pseudo differential operators fail.
Nevertheless we can show order reducing properties of the commutator using a different proof, which relies on estimates for negative definite functions.

Theorem 4.3. *Let $M \in \mathbb{N}$ and $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a symbol that satisfies (4.14) for that M. Then for all $s, t \in \mathbb{R}$ such that $|s-1| + 1 + |t| + n < M$ we have*

$$\|[\lambda^s(D), p(x,D)]\, u\|_{t,\lambda} \leq c_{M,s,t,\psi} \sum_{\substack{\beta\in\mathbb{N}_0^n \\ |\beta|\leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \cdot \|u\|_{t+s+1,\lambda} \qquad \text{for all } u \in C_0^\infty(\mathbb{R}^n).$$

Here $c_{M,s,t,\psi}$ is independent of the particular choice of p.

Note that under assumption (4.14) we expect $p(x, D)$ to be an opertor of order 2 and $\lambda^s(D)$ is of order s. Thus Theorem 4.3 states that the order of the commutator is $s + 1$, which in fact is one order less than the sum of the operator orders.

Proof: First note that by Theorem 2.7 the reference function satisfies $\psi(\xi) \leq c_\psi(1 + |\xi|^2) = c_\psi\langle\xi\rangle$, hence

$$\lambda(\xi) \leq (1 + c_\psi)^{1/2} \cdot \langle\xi\rangle \qquad \text{for all } \xi \in \mathbb{R}^n. \tag{4.17}$$

Moreover note that $\lambda(\xi) = (1 + \psi(\xi))^{1/2}$ is the square root of a continuous negative definite function. Thus by (2.5)

$$|\lambda(\xi) - \lambda(\eta)| \leq \lambda(\xi - \eta) \tag{4.18}$$

and by Lemma 2.6 the generalized Peetre inequality

$$\frac{\lambda^s(\xi)}{\lambda^s(\eta)} \leq 2^{|s|/2}\, \lambda^{|s|}(\xi - \eta) \tag{4.19}$$

holds for all $\xi, \eta \in \mathbb{R}^n$. An elementary calculation yields

$$([\lambda^s(D), p(x,D)]\, u)\hat{}\,(\xi) = \int_{\mathbb{R}^n} \hat{p}(\xi - \eta, \eta)\, (\lambda^s(\xi) - \lambda^s(\eta))\, \hat{u}(\eta)\, đ\eta$$

and thus for all $v \in C_0^\infty(\mathbb{R}^n)$ by Plancherel's theorem

$$|([\lambda^s(D), p(x,D)]\, u, v)_0| \leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |\hat{p}(\xi - \eta, \eta)|\, |\lambda^s(\xi) - \lambda^s(\eta)|\, |\hat{u}(\eta)|\, |\hat{v}(\xi)|\, đ\eta đ\xi. \tag{4.20}$$

Next note that for all $x, y > 0$ by the mean value theorem

$$|x^s - y^s| \leq |s| \cdot |x - y| \cdot (x^{s-1} + y^{s-1})$$

holds, which gives by (4.18) and (4.19)

$$\begin{aligned}
|\lambda^s(\xi) - \lambda^s(\eta)| &\leq |s| \cdot |\lambda(\xi) - \lambda(\eta)| \cdot (\lambda^{s-1}(\xi) + \lambda^{s-1}(\eta)) \\
&\leq |s| \cdot \lambda(\xi - \eta) \cdot \left(2^{|s-1|/2} \lambda^{s-1}(\eta) \cdot \lambda^{|s-1|}(\xi - \eta) + \lambda^{s-1}(\eta)\right)\right) \cdot \lambda^t(\xi) \cdot \lambda^{-t}(\xi) \\
&\leq 2 \cdot 2^{|s-1|/2}\, |s| \cdot \lambda(\xi - \eta) \cdot \lambda^{|s-1|}(\xi - \eta) \cdot \lambda^{s-1}(\eta) \cdot 2^{|t|/2} \lambda^{|t|}(\xi - \eta) \cdot \lambda^t(\eta) \cdot \lambda^{-t}(\xi) \\
&= 2^{(|s-1|+|t|+2)/2}\, |s| \cdot \lambda^{|s-1|+|t|+1}(\xi - \eta) \cdot \lambda^{t+s-1}(\eta) \cdot \lambda^{-t}(\xi).
\end{aligned}$$

Hence by (4.15), Lemma 4.2 and (4.17)

$$\begin{aligned}
&|([\lambda^s(D), p(x,D)]\, u, v)_0| \\
&\quad \leq C_M \cdot 2^{(|s-1|+|t|+2)/2}\, |s| \cdot \\
&\qquad \cdot \sum_{\substack{\beta \in \mathbb{N}_0^n \\ |\beta| \leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \langle \xi - \eta \rangle^{-M} \lambda^{|s-1|+|t|+1}(\xi - \eta) \cdot \lambda^{t+s+1}(\eta)\, |\hat{u}(\eta)| \cdot \lambda^{-t}(\xi)\, |\hat{v}(\xi)|\, đ\eta đ\xi \\
&\quad \leq C_M \cdot 2^{1/2} \cdot (2(1 + c_\psi))^{(|s-1|+|t|+1)/2}\, |s| \cdot \\
&\qquad \cdot \sum_{\substack{\beta \in \mathbb{N}_0^n \\ |\beta| \leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \langle \xi - \eta \rangle^{-M+|s-1|+|t|+1} \cdot \lambda^{t+s+1}(\eta)\, |\hat{u}(\eta)| \cdot \lambda^{-t}(\xi)\, |\hat{v}(\xi)|\, đ\eta đ\xi.
\end{aligned}$$

By the assumption on s and t the function $\langle \cdot \rangle^{-M+|s-1|+|t|+1}$ is integrable on $\mathbb{R}^n$. Therefore, using first Cauchy-Schwarz- and then Young's inequality finally yields

$$|([\lambda^s(D), p(x,D)]\, u, v)_0| \leq c_{M,s,t,\psi} \sum_{\substack{\beta \in \mathbb{N}_0^n \\ |\beta| \leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \cdot \|u\|_{t+s+1,\lambda} \cdot \|v\|_{-t,\lambda}$$

with a constant

$$c_{M,s,t,\psi} = C_M \cdot 2^{1/2} \cdot (2(1 + c_\psi))^{(|s-1|+|t|+1)/2}\, |s| \cdot \left\| \langle \cdot \rangle^{-M+|s-1|+|t|+1} \right\|_{L^1(\mathbb{R}^n)}. \tag{4.21}$$

By (4.9) the theorem now follows immediately. □

For later purposes we introduce the Friedrichs mollifier $J_\varepsilon : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$, $\varepsilon > 0$, defined by $J_\varepsilon u = j_\varepsilon * u$, where

$$j_\varepsilon(x) = \varepsilon^{-n} j(\frac{x}{\varepsilon}),\ x \in \mathbb{R}^n,\ \text{and}\ j(x) := \begin{cases} c_0 \cdot e^{\frac{1}{|x|^2-1}} & \text{for } |x| < 1 \\ 0 & \text{for } |x| \geq 1 \end{cases}, \tag{4.22}$$

and c_0 is chosen such that $\int_{\mathbb{R}^n} j(x)\,dx = 1$. Because $(J_\varepsilon u)\hat{}(\xi) = \hat{j}(\varepsilon\xi) \cdot \hat{u}(\xi)$ and $\hat{j} \in \mathcal{S}(\mathbb{R}^n)$, we have $J_\varepsilon u \in H^{s,\lambda}(\mathbb{R}^n)$ for all $s \geq 0$. Moreover, since $|\hat{j}(\varepsilon\xi)| \leq \hat{j}(0) = 1$, we find for $u \in H^{t,\lambda}(\mathbb{R}^n)$

$$\|J_\varepsilon u\|_{t,\lambda} \leq \|u\|_{t,\lambda}$$

and

$$J_\varepsilon u \underset{\varepsilon \to 0}{\longrightarrow} u \quad \text{in } H^{t,\lambda}(\mathbb{R}^n).$$

For the Friedrichs mollifier the following uniform commutator estimates hold.

Theorem 4.4. *Let $M \in \mathbb{N}$ and $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a symbol that satisfies (4.14) for that M. Then for all $s \geq 0$ such that $|s-1| + 1 + n < M$ and all $0 < \varepsilon \leq 1$*

$$\|[J_\varepsilon, p(x,D)]\,u\|_{s,\lambda} \leq c\,\|u\|_{s+1,\lambda} \qquad \text{for all } u \in C_0^\infty(\mathbb{R}^n)$$

with a constant c independent of ε.

Proof: Note that

$$([J_\varepsilon, p(x,D)]\,u)\hat{}(\xi) = \int_{\mathbb{R}^n} \hat{p}(\xi-\eta,\eta)(\hat{j}(\varepsilon\xi) - \hat{j}(\varepsilon\eta))\hat{u}(\eta)\,đ\eta \tag{4.23}$$

and

$$|\hat{j}(\varepsilon\xi) - \hat{j}(\varepsilon\eta)| \cdot (1+|\xi|^2)^{1/2} \leq c(1+|\xi-\eta|^2)^{1/2} \tag{4.24}$$

with a constant c independent of $0 < \varepsilon \leq 1$, since (4.24) is trivial, if $|\xi - \eta| \geq \frac{|\xi|}{2}$ and otherwise the meanvalue theorem and the estimate $|\nabla\hat{j}(\xi)| \leq c(1+|\xi|^2)^{-1/2}$ yield

$$\begin{aligned} |\hat{j}(\varepsilon\xi) - \hat{j}(\varepsilon\eta)| \cdot (1+|\xi|^2)^{1/2} &\leq |\varepsilon\xi - \varepsilon\eta| \cdot c\left(1 + \left|\frac{\varepsilon\xi}{2}\right|^2\right)^{-1/2} \cdot (1+|\xi|^2)^{1/2} \\ &\leq c(1+|\xi-\eta|^2)^{1/2}\left(\frac{1}{\varepsilon^2} + \frac{|\xi|^2}{2}\right)^{-1/2} \cdot (1+|\xi|^2)^{1/2}. \end{aligned}$$

Combining (4.23), (4.24), Lemma 4.2 and (4.17), (4.19) we obtain

$$\begin{aligned} &|(\lambda^s(D)\,[J_\varepsilon, p(x,D)]\,u)\hat{}(\xi)| \\ &= \left|\int_{\mathbb{R}^n} \hat{p}(\xi-\eta,\eta)(\hat{j}(\varepsilon\xi) - \hat{j}(\varepsilon\eta))\lambda^s(\xi)\hat{u}(\eta)\,đ\eta\right| \\ &\leq c\int_{\mathbb{R}^n} \langle\xi-\eta\rangle^{-M}\lambda^2(\eta)\langle\xi-\eta\rangle \cdot \frac{\lambda(\xi)}{\langle\xi\rangle} \cdot \lambda^{s-1}(\xi)\,|\hat{u}(\eta)|\,đ\eta \\ &\leq c\int_{\mathbb{R}^n} \langle\xi-\eta\rangle^{-M+1}(1+c_\psi)^{1/2} \cdot \lambda^{s+1}(\eta) \cdot \lambda^{|s-1|}(\xi-\eta)\,|\hat{u}(\eta)|\,đ\eta \\ &\leq c\int_{\mathbb{R}^n} \langle\xi-\eta\rangle^{-M+1+|s-1|} \cdot \lambda^{s+1}(\eta)\,|\hat{u}(\eta)|\,đ\eta \end{aligned}$$

with a constant c not depending on ε. Observe that by assumption $\langle\cdot\rangle^{-M+1+|s-1|}$ is integrable. Thus by (4.12) Young's inequality again yields the desired estimate. □

4.2 The construction of Feller semigroups

In this section we develop the direct approach to Feller semigroups via the Hille-Yosida theorem as it was carried out by Jacob. In this context first results for pseudo differential operators with continuous negative definite symbols were obtained in [42] for symbols having a particular sum structure. These resulte were improved in [29] by using better commutator estimates based on properties of the underlying negative definite functions as described in the previous section. Finally the specific sum structure of the symbol could be dropped and in this general case Feller semigroups were constructed by Jacob in [43]. The results presented here are taken essentially from [43].

We recall the definition of a Feller semigroup.

Definition 4.5. *A* **Feller semigroup** *on* $\mathbb{R}^n$ *is a family of bounded linear operators* $T_t : C_\infty(\mathbb{R}^n) \to C_\infty(\mathbb{R}^n)$, $t \geq 0$ *such that* $T_0 = Id$ *and*

(i) $T_s \circ T_t = T_{s+t}$ *for all* $s, t \geq 0$ *(semigroup property),*

(ii) $T_t u \underset{t\to 0}{\longrightarrow} u$ *in* $C_\infty(\mathbb{R}^n)$ *for all* $u \in C_\infty(\mathbb{R}^n)$ *(strong continuity),*

(iii) For all $u \in C_\infty(\mathbb{R}^n)$ *such that* $0 \leq u \leq 1$ *and all* $t \geq 0$ *we have*

$$0 \leq T_t u \leq 1 \qquad \textit{(submarkovian property).}$$

In other word a Feller semigroup is a strongly continuous contraction semigroup on $C_\infty(\mathbb{R}^n)$ which is also positivity preserving. As usual we can define the generator $(L, D(L))$ of this strongly continuous semigroup with domain

$$D(L) = \{u \in C_\infty(\mathbb{R}^n) : \lim_{t\to 0} \frac{1}{t}(T_t u - u) \text{ exists in } C_\infty(\mathbb{R}^n)\}$$

by

$$Lu = \lim_{t\to 0} \frac{1}{t}(T_t u - u).$$

As it is well-known, $(L, D(L))$ is a densely defined and closed operator in $C_\infty(\mathbb{R}^n)$ and determines the semigroup $(T_t)_{t\geq 0}$ in a unique way, see [90]. By the theorem of Hille–Yosida we have a complete characterization of all generators of strongly continuous semigroups. Let us recall this theorem for generators of strongly continuous contraction semigroups in a general Banach space $(B, \|\cdot\|)$, see [17], 1.2.6:

Theorem 4.6. *A linear operator* $(A, D(A))$ *in a Banach space* $(B, \|\cdot\|)$ *is closable and the closure is the generator of a strongly continuous semigroup of contractions* $(T_t)_{t\geq 0}$ *if and only if*

(i) $D(A)$ *is dense in* B,

(ii) A is dissipative:

$$\|\tau u - Au\| \geq \tau \, \|u\| \qquad \textit{for all } \tau > 0, \; u \in D(A),$$

(iii) the range $R(\tau - A)$ is dense in B for some $\tau > 0$.

Note that in the case that B is a Hilbert space H, (ii) follows from

$$(-Au, u)_H \geq 0 \qquad \text{for all } u \in D(A).$$

There is a refined version of this theorem, which characterizes generators of Feller semigroups, see [17], 4.2.2.

Theorem 4.7. *A linear operator $(A, D(A))$ in $C_\infty(\mathbb{R}^n)$ is closable and the closure is the generator of a Feller semigroup if and only if*

(i) $D(A) \subset C_\infty(\mathbb{R}^n)$ is dense,

(ii) A satisfies the positive maximum principle on $D(A)$,

(iii) the range $R(\tau - A) \subset C_\infty(\mathbb{R}^n)$ is dense for some $\tau > 0$.

Note that the positive maximum principle implies that the operator A is dissipative, i.e.

$$\|\tau u - Au\|_\infty \geq \tau \, \|u\|_\infty \qquad \text{for all } f \in D(A), \; \tau > 0,$$

see [17], 4.2.1. Therefore Theorem 4.7 is based on the Hille–Yosida theorem 4.6. The stronger assumption of the positive maximum principle implies that the semigroup is in addition positivity preserving, i.e. a Feller semigroup. In other words a strongly continuous contraction semigroup on $C_\infty(\mathbb{R}^n)$ is a Feller semigroup if and only if the generator satisfies the positive maximum principle. Hence, provided $C_0^\infty(\mathbb{R}^n)$ is contained in the domain of the generator, by Theorem 2.16 the generator must be a pseudo differential operator $-p(x, D)$ with a negative definite symbol $p(x, \xi)$. Conversely an operator of this type satisfies the conditions (i) and (ii) of Theorem 4.7. Our aim therfore will be to find conditions for the symbol such that the range condition (iii) of the Hille–Yosida theorem is also fulfilled. This amounts to show that the equation

$$(p(x, D) + \tau)u = f \tag{4.25}$$

has solutions for sufficiently many right hand sides f.

To solve this problem we will apply a Hilbert space approach based on the anisotropic Sobolev spaces which were introduced in the previous section. Let again $\psi : \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite reference function which satisfies the non-degeneracy condition (4.2). Moreover let

$$\lambda(\xi) = (1 + \psi(\xi))^{1/2}$$

and

$$p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

be a continuous negative definite symbol. For an arbitrary point $x_0 \in \mathbb{R}^n$ we split the symbol into two parts

$$p_1(\xi) = p(x_0, \xi)$$

and

$$p_2(x, \xi) = p(x, \xi) - p(x_0, \xi),$$

i.e.

$$p(x, \xi) = p_1(\xi) + p_2(x, \xi). \tag{4.26}$$

The symbol $p_1(\xi)$ is independent of x, so the corresponding operator $-p_1(D)$ is just a Fourier multiplier and well-understood. Actually $p_1(\xi)$ is a continuous negative definite function and $-p_1(D)$ is the generator of a Lévy process.

Now the idea is to regard $p_2(x, \xi)$ as small perturbation of the symbol $p_1(\xi)$. We therefore consider the following assumptions:

Let $M \in \mathbb{N}$. For $M > n + 1$ the function $\langle \cdot \rangle^{-M+1}$ is integrable over $\mathbb{R}^n$. Then by γ_M we denote the constant

$$\gamma_M = \left(8 C_M (2(1 + c_\psi))^{1/2} \left\| \langle \cdot \rangle^{-M+1} \right\|_{L^1(\mathbb{R}^n)} \right)^{-1}, \tag{4.27}$$

where the constants C_M and c_ψ are defined in (4.16) and (4.17). We assume:

(A.1) There are constants $c_0, c_1 > 0$ such that

$$c_0 \lambda^2(\xi) \leq p_1(\xi) \leq c_1 \lambda^2(\xi) \qquad \text{for all } |\xi| \geq 1 \tag{4.28}$$

and for $M \in \mathbb{N}$:

(A.2.M) the symbol $p_2(x, \xi)$ is M-times continuously differentiable with respect to x and for $\beta \in \mathbb{N}_0^n$, $|\beta| \leq M$ there are functions $\Phi_\beta \in L^1(\mathbb{R}^n)$ such that

$$\left| \partial_x^\beta p_2(x, \xi) \right| \leq \Phi_\beta(x) \cdot \lambda^2(\xi), \tag{4.29}$$

and for $M > n + 1$:

(A.3.M) with the constant $\gamma_M > 0$ as in (4.27) we have

$$\sum_{\substack{\beta \in \mathbb{N}_0^n \\ |\beta| \leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \leq \gamma_M \cdot c_0, \tag{4.30}$$

where c_0 is the constant given in (A.1).

The first assumption says that the symbol $p_1(\xi)$ has the same behaviour as the reference function $\psi(\xi)$. Of course we could have taken $p_1(\xi)$ itself as the reference function. But the slightly more general assumption (A.1) enables us to obtain uniform estimates for a whole class of symbols. (A.2.M) relates also the x-dependent part to the reference function.
The condition (A.3.M) states that the perturbation of $p_1(\xi)$ by $p_2(x, \xi)$ is small, where the size of the perturbation relative to p_1 is measured by the constant γ_M.

We first note that $p(x, D)$ is a well-defined object in the anisotropic Sobolev spaces $H^{s,\lambda}(\mathbb{R}^n)$.

Theorem 4.8. *Assume* (A.1) *and* (A.2.M). *Then for* $s \in \mathbb{R}$ *such that* $|s-1|+1+n < M$ *we have*

$$\|p(x,D)u\|_{s,\lambda} \leq c\,\|u\|_{s+2,\lambda} \qquad \text{for all } u \in C_0^\infty(\mathbb{R}^n). \tag{4.31}$$

For $s = 0$ *this estimate holds even if* $M > n$.

Proof: The condition (A.1) implies by (4.12) and Plancherel's theorem

$$\begin{aligned} \|p_1(D)u\|_{s,\lambda} &= \|\lambda^s(D)\,p_1(D)u\|_0 \\ &= \left(\int_{\mathbb{R}^n} |\lambda^s(\xi)p_1(\xi)\hat{u}(\xi)|^2\, đ\xi\right)^{1/2} \leq c_1' \left(\int_{\mathbb{R}^n} \left|\lambda^{s+2}(\xi)\hat{u}(\xi)\right|^2\, đ\xi\right)^{1/2} \\ &= c_1' \left\|\lambda^{s+2}(D)u\right\|_0 = c_1'\,\|u\|_{s+2,\lambda}\,. \end{aligned} \tag{4.32}$$

Concerning an estimate for $p_2(x,D)$ we first consider the case $s = 0$. Then by Plancherel's theorem and Lemma 4.2

$$\begin{aligned} \|p_2(x,D)u\|_0 &= \sup_{\substack{v\in C_0^\infty(\mathbb{R}^n)\\ \|v\|_0=1}} |(p_2(x,D)u,v)_0| \\ &= \sup_{\substack{v\in C_0^\infty(\mathbb{R}^n)\\ \|v\|_0=1}} \left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \hat{p}_2(\xi-\eta,\eta)\hat{u}(\eta)\overline{\hat{v}(\xi)}\, đ\eta đ\xi\right| \\ &\leq \sup_{\substack{v\in C_0^\infty(\mathbb{R}^n)\\ \|v\|_0=1}} C_M \sum_{\substack{\beta\in\mathbb{N}_0^n\\ |\beta|\leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \langle\xi-\eta\rangle^{-M}\,\lambda^2(\eta)\;|\hat{u}(\eta)|\;|\hat{v}(\xi)|\; đ\eta đ\xi \\ &\leq \sup_{\substack{v\in C_0^\infty(\mathbb{R}^n)\\ \|v\|_0=1}} C_M \sum_{\substack{\beta\in\mathbb{N}_0^n\\ |\beta|\leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \cdot \left\|\langle\cdot\rangle^{-M}\right\|_{L^1(\mathbb{R}^n)} \cdot \|u\|_{2,\lambda}\cdot\|v\|_{0,\lambda} \\ &= C_M \sum_{\substack{\beta\in\mathbb{N}_0^n\\ |\beta|\leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \cdot \left\|\langle\cdot\rangle^{-M}\right\|_{L^1(\mathbb{R}^n)} \cdot \|u\|_{2,\lambda} \end{aligned} \tag{4.33}$$

where for the last inequality we used Cauchy-Schwarz- and Youngs' inequality. Note that by assumption $M > n$ and hence $\langle\cdot\rangle^{-M}$ is integrable. In particular $p_2(x,D)$ extends continuously to $H^{2,\lambda}(\mathbb{R}^n)$

Now let $s \in \mathbb{R}$ be arbitrary. Then

$$\begin{aligned} \|p_2(x,D)u\|_{s,\lambda} &= \|\lambda^s(D)\,p_2(x,D)u\|_0 \\ &\leq \|p_2(x,D)\,\lambda^s(D)u\|_0 + \|[\lambda^s(D),p_2(x,D)]\,u\|_0 \end{aligned}$$

But by (4.33)

$$\begin{aligned} \|p_2(x,D)\,\lambda^s(D)u\|_0 &\leq C_M \sum_{\substack{\beta\in\mathbb{N}_0^n\\ |\beta|\leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \left\|\langle\cdot\rangle^{-M}\right\|_{L^1(\mathbb{R}^n)} \cdot \|\lambda^s(D)u\|_{2,\lambda} \\ &= C_M \sum_{\substack{\beta\in\mathbb{N}_0^n\\ |\beta|\leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \left\|\langle\cdot\rangle^{-M}\right\|_{L^1(\mathbb{R}^n)} \cdot \|u\|_{s+2,\lambda} \end{aligned}$$

and by Theorem 4.3

$$\begin{aligned}\|[\lambda^s(D), p_2(x,D)]\,u\|_0 &\leq c_{M,s,\psi} \sum_{\substack{\beta\in\mathbb{N}_0^n\\|\beta|\leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \|u\|_{s+1,\lambda}\\ &\leq c_{M,s,\psi} \sum_{\substack{\beta\in\mathbb{N}_0^n\\|\beta|\leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \|u\|_{s+2,\lambda}\,,\end{aligned}$$

which together with (4.32) implies (4.31). □

Therefore under the assumptions of Therorem 4.8 we may extend $p(x,D)$ to a continuous operator

$$p(x,D) : H^{s+2,\lambda}(\mathbb{R}^n) \to H^{s,\lambda}(\mathbb{R}^n)$$

and $p(x,D)$ is consistently defined for different values of s.
Theorem 4.8 shows that $p(x,D)$ indeed behaves like an operator of order 2 in the scale of Sobolev spaces $H^{s,\lambda}(\mathbb{R}^n)$.

In order to find solutions for the equation (4.25)

$$(p(x,D)+\tau)u = f,$$

we are first seeking for a weak solution. We therefore define the bilinear form

$$B_\tau(u,v) = ((p(x,D)+\tau)u, v)_0, \qquad u,v \in C_0^\infty(\mathbb{R}^n). \tag{4.34}$$

It turns out that B_τ extends to a continuous and coercive bilinear form on the space $H^{1,\lambda}(\mathbb{R}^n)$.

Theorem 4.9. *Assume* (A.1) *and* (A.2.M) *with $M > n+1$. Then*

$$|B_\tau(u,v)| \leq c\,\|u\|_{1,\lambda}\cdot\|v\|_{1,\lambda} \qquad \text{for all } u,v\in C_0^\infty(\mathbb{R}^n), \tag{4.35}$$

i.e. B_τ extends continuously to $H^{1,\lambda}(\mathbb{R}^n)\times H^{1,\lambda}(\mathbb{R}^n)$.
If moreover (A.3.M) *is satisfied, then there is a constant $\tau_0\in\mathbb{R}$ such that for all $\tau \geq \tau_0$*

$$B_\tau(u,u) \geq \frac{c_0}{2}\|u\|_{1,\lambda}^2 \qquad \text{for all } u\in H^{1,\lambda}(\mathbb{R}^n). \tag{4.36}$$

Proof: First we obtain by Plancherel's theorem and (A.1)

$$\begin{aligned}|(p_1(D)u,v)_0| &= \left|\int_{\mathbb{R}^n} p_1(\xi)\hat{u}(\xi)\overline{\hat{v}(\xi)}\,đ\xi\right|\\ &\leq c_1' \int_{\mathbb{R}^n} \lambda(\xi)\,|\hat{u}(\xi)|\cdot\lambda(\xi)\,|\hat{v}(\xi)|\,đ\xi\\ &\leq c_1'\,\|u\|_{1,\lambda}\cdot\|v\|_{1,\lambda}\,.\end{aligned}$$

Moreover we find as in (4.33) using (4.17) and (4.19)

$$
\begin{aligned}
|(p_2(x,D)u,v)_0| &\leq C_M \sum_{\substack{\beta\in\mathbb{N}_0^n \\ |\beta|\leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \langle \xi-\eta\rangle^{-M} \lambda(\eta)\, |\hat{u}(\eta)| \cdot \lambda(\eta)\, |\hat{v}(\xi)|\, đ\eta đ\xi \\
&\leq C_M \cdot (2(1+c_\psi))^{1/2} \cdot \\
&\qquad \cdot \sum_{\substack{\beta\in\mathbb{N}_0^n \\ |\beta|\leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \langle \xi-\eta\rangle^{-M+1} \lambda(\eta)\, |\hat{u}(\eta)| \cdot \lambda(\xi)\, |\hat{v}(\xi)|\, đ\eta đ\xi \qquad (4.37) \\
&\leq C_M \cdot (2(1+c_\psi))^{1/2} \cdot \sum_{\substack{\beta\in\mathbb{N}_0^n \\ |\beta|\leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \left\| \langle\cdot\rangle^{-M+1} \right\|_{L^1(\mathbb{R}^n)} \cdot \|u\|_{1,\lambda} \cdot \|v\|_{1,\lambda} \,,
\end{aligned}
$$

which implies (4.35), since of course

$$
|\tau(u,v)_0| \leq |\tau|\, \|u\|_0 \cdot \|v\|_0 \,.
$$

To prove the lower estimate first note that by (A.1) with a suitable constant $\tau_0 \geq 0$

$$
\begin{aligned}
(p_1(D)u,u)_0 &= \int_{\mathbb{R}^n} p_1(\xi)\, |\hat{u}(\xi)|^2\, đ\xi \geq \int_{\mathbb{R}^n} (c_0\lambda^2(\xi) - \tau_0)\, |\hat{u}(\xi)|^2\, đ\xi \\
&= c_0 \|u\|_{1,\lambda}^2 - \tau_0 \|u\|_0^2 \,. \qquad (4.38)
\end{aligned}
$$

But by (4.37), the choice of γ_M and (A.3.M)

$$
|(p_2(x,D)u,u)_0| \leq \frac{1}{8\gamma_M} \sum_{\substack{\beta\in\mathbb{N}_0^n \\ |\beta|\leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \cdot \|u\|_{1,\lambda}^2 \leq \frac{c_0}{8} \|u\|_{1,\lambda}^2 \,,
$$

Therefore combining the above estimates yields

$$
\begin{aligned}
(p(x,D)u,u)_0 &\geq (p_1(D)u,u)_0 - |(p_2(x,D)u,u)_0| \\
&\geq (c_0 - \frac{c_0}{8}) \|u\|_{1,\lambda}^2 - \tau_0 \|u\|_0^2 \\
&\geq \frac{c_0}{2} \|u\|_{1,\lambda}^2 - \tau_0 \|u\|_0^2 \,,
\end{aligned}
$$

that is (4.36). □

For each $f \in L^2(\mathbb{R}^n)$ the map $v \mapsto (f,v)_0$ is a bounded linear functional on $H^{1,\lambda}(\mathbb{R}^n)$. Thus under the conditions of Theorem 4.9 the Lax-Milgram theorem (cf. [96], p.92) immediately yields a weak solution of the equation (4.25) in the following sense:

Theorem 4.10. *Assume* (A.1), (A.2.M) *and* (A.3.M) *with* $M > n+1$. *Then there is a* $\tau_0 \in \mathbb{R}$ *such that for all* $\tau \geq \tau_0$ *and for all* $f \in L^2(\mathbb{R}^n)$ *there is a unique* $u \in H^{1,\lambda}(\mathbb{R}^n)$ *such that*

$$
B_\tau(u,v) = (f,v)_0 \qquad \text{for all } v \in C_0^\infty(\mathbb{R}^n).
$$

Of course a weak solution in $H^{1,\lambda}(\mathbb{R}^n)$ is not sufficient for our purposes, since we are aiming for a solution in $C_\infty(\mathbb{R}^n)$ to satisfy the conditions of the Hille-Yosida theorem. By the Sobolev embedding given in Proposition 4.1 it is therefore necessary to look for a regularity result for the solution in higher order Sobolev spaces. Because of the lower bound for the symbol given in (A.1) we expect an elliptic-like situation. In fact we can prove

Theorem 4.11. *Let $s \geq 0$ and assume that* (A.1), (A.2.M) *and* (A.3.M) *with $M > |s-1|+1+n$ hold true. Then for all $u \in H^{s+2,\lambda}(\mathbb{R}^n)$ we have*

$$\|u\|_{s+2,\lambda} \leq c(\|p(x,D)u\|_{s,\lambda} + \|u\|_0). \tag{4.39}$$

Proof: By continuity it is enough to prove (4.39) for $u \in C_0^\infty(\mathbb{R}^n)$. For the operator $p_1(D)$ the assumption (A.1) and (4.11) yield

$$\begin{aligned} \|p_1(D)u\|_{s,\lambda}^2 &= \int_{\mathbb{R}^n} \lambda^{2s}(\xi) p_1(\xi)^2 |\hat{u}(\xi)|^2 \, đ\xi \\ &\geq \int_{\mathbb{R}^n} \lambda^{2s}(\xi)(c_0^2\lambda^4(\xi) - c) |\hat{u}(\xi)|^2 \, đ\xi \\ &= c_0^2 \|u\|_{s+2,\lambda}^2 - c\|u\|_{s,\lambda}^2 \geq \frac{1}{4}c_0^2 \|u\|_{s+2,\lambda}^2 - c\|u\|_0^2 \end{aligned} \tag{4.40}$$

or

$$\|p_1(D)u\|_{s,\lambda} \geq \frac{1}{2}c_0 \|u\|_{s+2,\lambda} - c\|u\|_0 .$$

Moreover we see from the proof of Theorem 4.8

$$\|p_2(x,D)u\|_{s,\lambda} \leq C_M \sum_{\substack{\beta\in\mathbb{N}_0^n \\ |\beta|\leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \left\| \langle\cdot\rangle^{-M} \right\|_{L^1(\mathbb{R}^n)} \cdot \|u\|_{s+2,\lambda} + c_{M,s,\psi} \sum_{\substack{\beta\in\mathbb{N}_0^n \\ |\beta|\leq M}} \|\Phi_\beta\|_{L^1(\mathbb{R}^n)} \|u\|_{s+1,\lambda} .$$

Thus by assumption (A.3.M) and again by (4.11)

$$\begin{aligned} \|p_2(x,D)u\|_{s,\lambda} &\leq \frac{1}{8}c_0\|u\|_{s+2,\lambda} + \frac{1}{8}c_0\|u\|_{s+2,\lambda} + c\|u\|_0 \\ &= \frac{1}{4}c_0\|u\|_{s+2,\lambda} + c\|u\|_0 . \end{aligned}$$

We combine this result with the esimate for $p_1(D)$ and obtain

$$\begin{aligned} \|p(x,D)u\|_{s,\lambda} &\geq \|p_1(D)u\|_{s,\lambda} - \|p_2(x,D)u\|_{s,\lambda} \\ &\geq \frac{c_0}{4}\|u\|_{s+2,\lambda} - c\|u\|_0 \end{aligned}$$

and (4.39) follows. □

Theorem 4.11 implies the following regularity result for solutions of the equation (4.25)

Theorem 4.12. *Let $s \geq 1$ and assume that* (A.1), (A.2.M) *and* (A.3.M) *hold with an $M > s+n$. Then for any $f \in H^{s,\lambda}(\mathbb{R}^n)$ there is a unique solution $u \in H^{s+2,\lambda}(\mathbb{R}^n)$ of the equation*

$$(p(x,D)+\tau)u = f \tag{4.41}$$

for all $\tau \geq \tau_0$ and τ_0 is chosen as in Theorem 4.9.

Proof: With this choice of s and M the assumptions of the Theorems 4.4, 4.8, 4.9, 4.10 and 4.11 are satisfied. In particular $p(x,D) : H^{t+2,\lambda}(\mathbb{R}^n) \to H^{t,\lambda}(\mathbb{R}^n)$ is continuous for all $1 \le t \le s$ and for $s = 0$. By Theorem 4.10 there exists a unique weak solution $u \in H^{1,\lambda}(\mathbb{R}^n)$ of (4.41):

$$B_\tau(u,v) = (f,v)_0 \qquad \text{for all } v \in C_0^\infty(\mathbb{R}^n).$$

We first prove that $u \in H^{2,\lambda}(\mathbb{R}^n)$. Choose a sequence $(u_k)_{k\in\mathbb{N}}$ in $C_0^\infty(\mathbb{R}^n)$ which converges to u in $H^{1,\lambda}(\mathbb{R}^n)$. Let J_ε be the Friedrichs mollifier defined in (4.22). Note that J_ε is a symmetric operator in $L^2(\mathbb{R}^n)$ and that J_ε commutes with $p_1(D)$. Then for all $v \in C_0^\infty(\mathbb{R}^n)$

$$\begin{aligned}(4.42)\quad ((p(x,D)+\tau)J_\varepsilon\, u_k, v)_0 &= (J_\varepsilon(p(x,D)+\tau)\, u_k, v)_0 - ([J_\varepsilon, p(x,D)+\tau]\ u_k, v)_0\\ &= ((p(x,D)+\tau)\, u_k, J_\varepsilon v)_0 - ([J_\varepsilon, p(x,D)]\ u_k, v)_0\\ &= B_\tau(u_k, J_\varepsilon v) - ([J_\varepsilon, p_2(x,D)]\ u_k, v)_0.\end{aligned}$$

Note that J_ε maps $H^{t,\lambda}(\mathbb{R}^n)$ continuously into $H^{t',\lambda}(\mathbb{R}^n)$ for all $t,t' \ge 0$. Moreover, by the commutator estimate, Theorem 4.4, we know that

$$[J_\varepsilon, p_2(x,D)]\ u_k \underset{k\to\infty}{\longrightarrow} w_\varepsilon \in L^2(\mathbb{R}^n)$$

in $L^2(\mathbb{R}^n)$ and $\|w_\varepsilon\|_0 \le K$ is uniformly bounded for all $0 < \varepsilon \le 1$. Therefore we obtain for $k \to \infty$

$$\begin{aligned}(4.43)\qquad ((p(x,D)+\tau)J_\varepsilon\, u, v)_0 &= B_\tau(u, J_\varepsilon v) - (w_\varepsilon, v)_0\\ &= (f, J_\varepsilon v)_0 - (w_\varepsilon, v)_0\\ &= (J_\varepsilon f, v)_0 - (w_\varepsilon, v)_0.\end{aligned}$$

Taking the supremum over all v with $\|v\|_{L^2(\mathbb{R}^n)} = 1$ yields

$$\|(p(x,D)+\tau)J_\varepsilon\, u\|_0 \le \|J_\varepsilon f\|_0 + \|w_\varepsilon\|_0 \le \|f\|_0 + K.$$

Thus both $\|(p(x,D)+\tau)J_\varepsilon\, u\|_0$ and $\|J_\varepsilon u\|_0$ are uniformly bounded for $0 < \varepsilon \le 1$ and hence by Theorem 4.11

$$\sup_{0<\varepsilon\le 1} \|J_\varepsilon u\|_{2,\lambda} < \infty.$$

Since $H^{2,\lambda}(\mathbb{R}^n)$ is a Hilbert space, there is a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ tending to 0, such that $(J_{\varepsilon_k} u)$ converges weakly to an element $\tilde{u} \in H^{2,\lambda}(\mathbb{R}^n)$. But $J_{\varepsilon_k} u \underset{k\to\infty}{\longrightarrow} u$ in $H^{1,\lambda}(\mathbb{R}^n)$, thus $\tilde{u} = u$ and we have $u \in H^{2,\lambda}(\mathbb{R}^n)$.

We now iterate the argument. Suppose that $u \in H^{t,\lambda}(\mathbb{R}^n)$ for some $2 \le t \le s+1$ and use in (4.42) a sequence $(u_k)_{k\in\mathbb{N}}$ of testfunctions that converges to u in $H^{t,\lambda}(\mathbb{R}^n)$. We then obtain (4.43) with $\sup_{0<\varepsilon\le 1} \|w_\varepsilon\|_{t-1,\lambda} < \infty$. Taking in (4.43) the supremum over all v such that $\|v\|_{1-t,\lambda} = 1$ yields that $(p(x,D)+\tau)J_\varepsilon u$ and $J_\varepsilon u$ are uniformly bounded in $H^{t-1,\lambda}(\mathbb{R}^n)$. Consequently by Theorem 4.11 the same holds true for $J_\varepsilon u$ in $H^{t+1,\lambda}(\mathbb{R}^n)$ and we conclude as above that $u \in H^{t+1,\lambda}(\mathbb{R}^n)$.

Therefore, after a finite number of iterations the result $u \in H^{s+2,\lambda}(\mathbb{R}^n)$ follows.

Finally we have to remark that u is a solution of (4.41) in the strong sense. But clearly for any sequence $(u_k)_{k\in\mathbb{N}}$ of testfunctions that converges to u in $H^{s+2,\lambda}(\mathbb{R}^n)$ and all $v \in C_0^\infty(\mathbb{R}^n)$ we have

$$((p(x,D)+\tau)u_k, v)_0 \underset{k\to\infty}{\longrightarrow} ((p(x,D)+\tau)u, v)_0$$

as well as

$$((p(x,D)+\tau)u_k, v)_0 = B_\tau(u_k, v) \underset{k\to\infty}{\longrightarrow} B_\tau(u,v) = (f,v)_0,$$

that is (4.41) holds. □

Using this regularity result it is now easy to find a solution of the equation (4.25) in $C_\infty(\mathbb{R}^n)$. In combination with the Hille-Yosida theorem we finally can prove that pseudo differential operators with continuous negative definite symbols give examples of generators of Feller semigroups. We have the following theorem, cf. [43], Theo. 5.2.

Theorem 4.13. *Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite reference function that satisfies (4.2) for some $r > 0$ and define $\lambda(\xi)$ as in (4.3). Assume that the conditions* (A.1), (A.2.M) *and* (A.3.M) *hold with an integer $M > \left(\frac{n}{r} \vee 1\right) + n$. Then $-p(x,D)$ with domain $C_0^\infty(\mathbb{R}^n)$ is closable in $C_\infty(\mathbb{R}^n)$ and the closure is the generator of a Feller semigroup.*

Proof: Fix $s \in \mathbb{R}$ as a number strictly between $\frac{n}{r} \vee 1$ and $M - n$. Then $s > \frac{n}{r}$ and $M > s+n$. Therefore the assumptions of Theorem 4.13 are satisfied and the Sobolev embedding $H^{t,\lambda}(\mathbb{R}^n) \hookrightarrow C_\infty(\mathbb{R}^n)$ is valid for all $t \geq s$ by Proposition 4.1. Define a linear operator $(A, D(A))$ in $C_\infty(\mathbb{R}^n)$ by

$$D(A) = H^{s+2,\lambda}(\mathbb{R}^n) \subset C_\infty(\mathbb{R}^n)$$

and

$$Au = -p(x,D)u \qquad \text{for } u \in D(A).$$

Then A is an well-defined operator in $C_\infty(\mathbb{R}^n)$, since $p(x,D) : H^{s+2,\lambda}(\mathbb{R}^n) \to H^{s,\lambda}(\mathbb{R}^n) \subset C_\infty(\mathbb{R}^n)$. Moreover by Sobolev embedding, Theorem 4.8 and Theorem 4.11 we see that the graph norm of A in $C_\infty(\mathbb{R}^n)$

$$\|u\|_A = \|u\|_\infty + \|Au\|_\infty, \qquad u \in D(A) = H^{s+2,\lambda}(\mathbb{R}^n),$$

is weaker than the norm $\|u\|_{s+2,\lambda}$ and therefore $C_0^\infty(\mathbb{R}^n)$ is a core of A.

Since $p(x,\xi)$ is a continuous negative definite symbol, $-p(x,D)$ satisfies the positive maximum principle on $C_0^\infty(\mathbb{R}^n)$ by Theorem 2.18. But then by Proposition 2.20 A also satisfies the positive maximum principle on $D(A)$. Hence A fulfills the conditions (i) and (ii) of the Hille-Yosida theorem 4.7. Moreover for $\tau > 0$ chosen as in Theorem 4.9, for any $f \in H^{s,\lambda}(\mathbb{R}^n)$ there is a $u \in H^{s+2,\lambda}(\mathbb{R}^n) = D(A)$ such that

$$(A-\tau)u = f.$$

As $H^{s,\lambda}(\mathbb{R}^n)$ is dense in $C_\infty(\mathbb{R}^n)$ all conditions of the Hille-Yosida theorem are satisfied. Thus A is closable in $C_\infty(\mathbb{R}^n)$ and the closure generates a Feller semigroup. But, since $C_0^\infty(\mathbb{R}^n)$ is a core of A, this closure coincides with the closure of $A = -p(x,D)$ with domain $C_0^\infty(\mathbb{R}^n)$. □

Of course the derivation of this result has yielded more informations on the generator expressed in terms of estimates in appropriate Sobolev spaces. These estimates are of interest for their own, because they imply results for the semigroup and accordingly, they have probabilistic implications. We will focus on this point in Chapter 8 below.

Chapter 5

The martingale problem: Uniqueness of solutions

5.1 Localization

In the previous chapter we have seen that modified Hilbert space methods represent a useful tool for the investigation of pseudo differential operators with continuous negative definite symbols. In particular we constructed a non-trivial class of examples of generators of Feller semigroups, which are controlled in terms of a general continuous negative definite reference function. A certain drawback of the method is that the assumptions on the symbol are restrictive in the following sense. Condition (A.2.M) demands that the x-dependent part $p_2(x, \xi)$ in the decomposition (4.26) of the symbol is bounded by an integrable function with respect to x. Since the same is assumed also for the derivatives up to a certain order higher than the space dimension, (A.2.M) in the end means that $p_2(x, \xi)$ vanishes for $|x| \to \infty$. Therefore, for $|x| \to \infty$ the symbol $p(x, \xi)$ is asymptotically independent of x. Moreover, condition (A.3.M) imposes a strict bound on the size of the x-dependent perturbation $p_2(x, \xi)$ of the x-independent part $p_1(\xi)$.

One way to overcome these difficulties is the use of the martingale problem. This is mainly due to a localization procedure for solution of the martingale problem. We will briefly explain this technique.

We are looking for conditions for the pseudo differential operator such that the martingale problem is well-posed. This is appropriate, because in this case the martingale problem gives a unique characterization of a Markov process generated by the operator. Consider the following situation. Let again E be a separable, complete metric space and let $A : D(A) \to B(E)$ be linear operator with domain $D(A) \subset C_b(E)$. Define the D_E-martingale problem as in Section 3.1. We need the notion of the stopped martingale problem. Let $P \in \mathcal{M}_1(D_E)$ be a solution of the martingale problem for A, i.e.

$$\varphi(X_t) - \int_0^t A\varphi(X_s)\, ds, \qquad t \geq 0,$$

is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ under P. Now let $U \subset E$ be an open set and let τ_U be the $(\mathcal{F}_t)$-stopping time

$$\tau_U = \inf\{t \geq 0 : X_t \notin U \text{ or } (t > 0 \text{ and } X_{t-} \notin U)\}.$$

Note that $\tau_U(\omega)$ is the first contact time of the closed set $U^{\complement}$ by the càdlàg path $\omega \in D_E$, i.e. the first time t such that the closure of the path ω up to time t hits $U^{\complement}$. Therefore (see [17], 2.1.5), τ_U actually is a stopping time with repect to the canonical filtration $(\mathcal{F}_t)$.
Then by optional stopping obviously

$$\varphi(X_{t\wedge\tau_U}) - \int_0^{t\wedge\tau_U} A\varphi(X_s)\,ds, \qquad t \geq 0, \tag{5.1}$$

is an $(\mathcal{F}_t)$-martingale under P for all $\varphi \in D(A)$. Therefore we say that $P \in \mathcal{M}_1(D_E)$ is a solution of the stopped martingale problem for A and U, if (5.1) is an $(\mathcal{F}_t)$-martingale under P and

$$P(X_t = X_{t\wedge\tau} \text{ for all } t \geq 0) = 1 \tag{5.2}$$

to fix the values of (X_t) after stopping. If for all initial distributions $\nu \in \mathcal{M}_1(E)$ there is a unique solution of the stopped martingale problem with

$$P \circ X_0^{-1} = \nu,$$

then the stopped martingale problem is called well-posed.

Up to the stopping time τ_U the assumption on a solution for the stopped martingale problem and the original martingale problem coincide and after τ_U the values of the process of the stopped martingale problem are prescribed by (5.2). Therefore it is not hard to see:

Theorem 5.1. *If the martingale problem for A is well-posed, then the stopped martingale problem for A and U is well-posed, too.*

For a proof we refer to [17], 4.6.1, see also [82], Theo.3.4, for the case of continuous paths.
The important feature of the stopped martingale problem is that we can reverse this procedure and the well-posedness of the stopped martingale problem for a family of open sets, which cover the state space, implies the well-posedness of the original martingale problem:

Theorem 5.2. *Let $(U_k)_{k\in\mathbb{N}}$ be an open covering of E. Suppose for all initial distributions there exists a solution of the martingale problem for the operator A. If the stopped martingale problem for A and U_k is well-posed for all $k \in \mathbb{N}$, then also the martingale problem for A is well-posed.*

A detailed proof of this statement is given in [17], 4.6.2. The idea of the proof is based on the fact that it is possible to split a solution $P \in \mathcal{M}_1(D_E)$ of the martingale problem along a stopping time. More precisely, for an $(\mathcal{F}_t)$-stopping time τ we can find a regular conditional probability distribution $P_\omega(A|\mathcal{F}_\tau)$, i.e. a kernel of probabiltity measures on $\mathcal{M}_1(D_E)$, which for every $A \in \mathcal{F}$ is a version of the conditional probability $P(A|\mathcal{F}_\tau)(\omega)$ and which has the additional property that for all $\omega \in D_E$

$$P_\omega(X_{t\wedge\tau} = X_{t\wedge\tau}(\omega)) = 1.$$

Here as usual $\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \le t\} \in \mathcal{F}_t \text{ for all } t \ge 0\}$ is the σ-algebra of events up to τ. It turns out that up to an exceptional set of measure 0 for all $\omega \in D_E$ the measure $P_\omega(\cdot|\mathcal{F}_\tau)$ describes a solution of the martingale problem which starts at time $\tau(\omega)$ at the point $X_\tau(\omega)$.

Conversely, it is possible to compose a solution of the martingale problem by glueing together a solution up to the stopping time τ and solutions after τ in the following way: We can define a kernel of probabiltity measure on D_E, which for each $\omega \in D_E$ is a solution of the martingale problem starting at time $\tau(\omega)$ in $X_\tau(\omega)$, and which is $\mathcal{F}_\tau$-measurable with respect to ω. We then obtain a solution of the original martingale problem by integrating this kernel with respect to the solution of the martingale problem up to the stopping time τ. This feature of the martingale problem formulation is often called a hidden Markov property. We refer to [83] as a standard reference concerning this topic, in particular Theo. 1.2.10, see also [81], Theo.1.2 for the case of jump processes, as well as [17], section 4.6.
Note that for the argument above it is important that τ is a stopping time with respect to the cananical filtration.

In the situation of Theorem 5.2 we therefore can compose a solution of the martingale problem by piecewise joining solutions of stopped martingale problems along the stopping times τ_{U_k} in the above manner. Since the solutions of all stopped martingale problems are unique, this property carries over to the solution of the martingale problem itself.

We reformulate the localization technique for the case of pseudo differential operators in terms of the symbols

Theorem 5.3. *Let $p, p_k : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be continuous negative definite symbols, $k \in \mathbb{N}$, and suppose that the corresponding pseudo differential operators $p(x, D)$ and $p_k(x, D)$ map $C_0^\infty(\mathbb{R}^n)$ into $C_b(\mathbb{R}^n)$. Assume that for each initial distribution there is a solution of the martingale problem for $-p(x, D)$. Moreover assume that there is an open covering $(U_k)_{k\in\mathbb{N}}$ of $\mathbb{R}^n$ such that*

$$p(x, \xi) = p_k(x, \xi) \qquad \text{for all } x \in U_k,\ \xi \in \mathbb{R}^n.$$

If the martingale problem for $-p_k(x, D)$ is well-posed for all $k \in \mathbb{N}$, then the martingale problem for $-p(x, D)$ is well-posed, too.

Proof: Note that for $t < \tau_{U_k}$ we have $X_t \in U_k$ and therefore

$$\varphi(X_{t\wedge\tau_{U_k}}) - \int_0^{t\wedge\tau_{U_k}} (-p(x, D)\varphi)(X_s)\, ds \;=\; \varphi(X_{t\wedge\tau_{U_k}}) - \int_0^{t\wedge\tau_{U_k}} (-p_k(x, D)\varphi)(X_s)\, ds$$

for all $t \ge 0$ and all $\varphi \in C_0^\infty(\mathbb{R}^n)$. Thus $P \in \mathcal{M}_1(D_{\mathbb{R}^n})$ is a solution of the stopped martingale problem for $p(x, D)$ and U_k if and only if it is a solution to stopped martingale problem for $-p_k(x, D)$ and U_k. By Theorem 5.1 the stopped martingale problem for $-p_k(x, D)$ and U_k is well-posed, hence the stopped martingale problem for $-p(x, D)$ and U_k is well- posed for all $k \in \mathbb{N}$. Now the assertion follows from Theorem 5.2. □

5.2 A general uniqueness criterion

A probability measure $P \in \mathcal{M}_1(D_E)$ is determined by its finite-dimensional distributions, i.e. the distributions $P \circ (X_{t_1}, \dots, X_{t_k})^{-1}$, where $t_1, \dots, t_k \in [0, \infty)$. For solutions of the martingale problem the situation fortunately is more easy, since in this case uniqueness is implied by the uniqueness of the one-dimensional distributions:

Lemma 5.4. *Let $A : D(A) \to B(E)$ be linear operator with domain $D(A) \subset C_b(E)$. If for all $\mu \in \mathcal{M}_1(E)$ and solutions $P_1, P_2 \in \mathcal{M}_1(D_E)$ of the martingale problem for A with initial distribution μ we have*

$$P_1 \circ X_t^{-1} = P_2 \circ X_t^{-1} \qquad \text{for all } t \geq 0, \tag{5.3}$$

then for every initial distribution there is at most one solution of the martingale problem for A.

This result is again a consequence of the hidden Markov property as discussed in the previous section. Note that it is necessary to have uniqueness of the one-dimensional distribution for all initial distributions, even if one is interested only in the unique solvability for a particular initial distribution.

Proof: We follow the argument given in [17], 4.4.2. Let $P, P' \in \mathcal{M}_1(D_E)$ be solutions of the martingale problem for A with an initial distribution ν. It is sufficient to check that for all $m \in \mathbb{N}$ and all strictly positive functions $f_1, \dots, f_m \in B(E)$ and all $0 \leq t_1 < \dots < t_m$ we have

$$\mathbb{E}^P\left[\prod_{k=1}^m f_k(X_{t_k})\right] = \mathbb{E}^{P'}\left[\prod_{k=1}^m f_k(X_{t_k})\right]. \tag{5.4}$$

We prove (5.4) by induction. For $m = 1$ (5.4) follows from the assumption (5.3).
Taking for granted that (5.4) holds for some $m \in \mathbb{N}$ we define probability mesures P_1, $P_2 \in \mathcal{M}_1(D_E)$ by

$$P_1(B) = \mathbb{E}^P\left[(1_B \circ \theta_{t_m}) \cdot \prod_{k=1}^m f_k(X_{t_k})\right] \Big/ \mathbb{E}^P\left[\prod_{k=1}^m f_k(X_{t_k})\right]$$

and

$$P_2(B) = \mathbb{E}^{P'}\left[(1_B \circ \theta_{t_m}) \cdot \prod_{k=1}^m f_k(X_{t_k})\right] \Big/ \mathbb{E}^{P'}\left[\prod_{k=1}^m f_k(X_{t_k})\right]$$

for $B \in \mathcal{F}$, where $\theta_{t_m} : D_E \to D_E$ is the $\mathcal{F}$-$\mathcal{F}$-measurable shift operator given by: $\theta_{t_m}\omega(t) = \omega(t + t_m)$
Then for a Borel set $C \subset E$ the function $1_C = (1 + 1_C) - 1$ is the difference of strictly positive functions in $B(E)$. Therefore by (5.4)

$$\begin{aligned} P_1\{X_0 \in C\} &= \mathbb{E}^P\left[1_C(X_{t_m}) \cdot \prod_{k=1}^m f_k(X_{t_k})\right] \Big/ \mathbb{E}^P\left[\prod_{k=1}^m f_k(X_{t_k})\right] \\ &= \mathbb{E}^{P'}\left[1_C(X_{t_m}) \cdot \prod_{k=1}^m f_k(X_{t_k})\right] \Big/ \mathbb{E}^{P'}\left[\prod_{k=1}^m f_k(X_{t_k})\right] = P_2\{X_0 \in C\}. \end{aligned} \tag{5.5}$$

Furthermore for $r \in \mathbb{N}$, $0 \le s_1 \le \ldots \le s_r \le s_{r+1} \le s_{r+2}$, functions $h_l \in B(E)$, $l = 1, \ldots, r$, and $\varphi \in D(A)$ we have

$$\begin{aligned}
&\mathbb{E}^{P_1}\left[\left(\varphi(X_{s_{r+2}}) - \varphi(X_{s_{r+1}}) - \int_{s_{r+1}}^{s_{r+2}} A\varphi(X_u)\,du\right)\prod_{l=1}^{r} h_l(X_{s_l})\right] \\
&= \mathbb{E}^{P}\left[\left(\varphi(X_{s_{r+2}+t_m}) - \varphi(X_{s_{r+1}+t_m}) - \int_{s_{r+1}+t_m}^{s_{r+2}+t_m} A\varphi(X_u)\,du\right)\cdot\right. \\
&\qquad \left.\cdot\prod_{l=1}^{r} h_l(X_{s_l+t_m})\cdot\prod_{k=1}^{m} f_k(X_{t_k})\right] \Big/ \,\mathbb{E}^{P}\left[\prod_{k=1}^{m} f_k(X_{t_k})\right] = 0,
\end{aligned}$$

since $0 \le t_1 \le \ldots \le t_m \le s_1 + t_m \le \ldots \le s_r + t_m \le s_{r+1} + t_m \le s_{r+2} + t_m$ and P is a solution of the martingale problem for A. Analogously

$$\mathbb{E}^{P_2}\left[\left(\varphi(X_{s_{r+2}}) - \varphi(X_{s_{r+1}}) - \int_{s_{r+1}}^{s_{r+2}} A\varphi(X_u)\,du\right)\prod_{l=1}^{r} h_l(X_{s_l})\right] = 0.$$

Therefore P_1 and P_2 are solutions of the martingale problem for A , which have the same initial distribution by (5.5). Consequently, by assumption (5.3) and by (5.4) we obtain for all $t_{m+1} = t_m + t'$ with $t' \ge 0$ and all $f_{m+1} \in B(E)$

$$\begin{aligned}
\mathbb{E}^{P}\left[\prod_{k=1}^{m+1} f_k(X_{t_k})\right] &= \mathbb{E}^{P_1}\left[f_{m+1}(X_{t'})\right]\cdot\mathbb{E}^{P}\left[\prod_{k=1}^{m} f_k(X_{t_k})\right] \\
&= \mathbb{E}^{P_2}\left[f_{m+1}(X_{t'})\right]\cdot\mathbb{E}^{P'}\left[\prod_{k=1}^{m} f_k(X_{t_k})\right] \\
&= \mathbb{E}^{P'}\left[\prod_{k=1}^{m+1} f_k(X_{t_k})\right],
\end{aligned}$$

i.e. (5.4) holds with m replaced by $m + 1$. □

We want to apply Lemma 5.4 to prove well-posedness of the martingale problem for a pseudo differential operator $-p(x, D)$. The usual way to check the condition (5.3) is based on an analytic argument, namely the existence of sufficiently smooth solutions of the Cauchy problem for the operator $\frac{\partial}{\partial t} - p(x, D)$ (see [83], Theo. 6.3.2). To solve this kind of problem we again want to employ Hilbert space techniques in anisotropic Sobolev spaces $H^{s,\lambda}(\mathbb{R}^n)$ as developed in the previous chapter. We therefore formulate a criterion to verify (5.3) in terms of operators in spaces $H^{s,\lambda}(\mathbb{R}^n)$. Thus assume that $H^{s,\lambda}(\mathbb{R}^n)$ is defined as in (4.5), (4.6) and suppose that for some $s_0 > 0$ we have the Sobolev embedding

(5.6) $$H^{s_0,\lambda}(\mathbb{R}^n) \hookrightarrow C_\infty(\mathbb{R}^n) \subset C_b(\mathbb{R}^n)$$

For $T > 0$ by $C([0, T], H^{s,\lambda}(\mathbb{R}^n))$ and $C^1([0, T], H^{s,\lambda}(\mathbb{R}^n))$ we denote the spaces of strongly continuous and strongly continuously differentiable mappings $u : [0, T] \to H^{s,\lambda}(\mathbb{R}^n)$, respectively, equipped with the norms $\sup_{0\le t\le T} \|u(t)\|_{s,\lambda}$ and $\sup_{0\le t\le T} \|u(t)\|_{s,\lambda} + \sup_{0\le t\le T} \|u'(t)\|_{s,\lambda}$. In particular with the choice of s_0 as above

(5.7) $$C([0, T], H^{s_0,\lambda}(\mathbb{R}^n)) \hookrightarrow C_b([0, T] \times \mathbb{R}^n)$$

holds. Furthermore, since $C_0^\infty(\mathbb{R}^n) \subset H^{s_0,\lambda}(\mathbb{R}^n)$, it is clear that we can regard $C_0^\infty([0,T] \times \mathbb{R}^n)$ as a subspace of $C^1([0,T], H^{s_0,\lambda}(\mathbb{R}^n))$.

Theorem 5.5. *Let $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite symbol and $p(x,D)$ the corresponding pseudo differential operator defined on $C_0^\infty(\mathbb{R}^n)$. Assume that (5.6) holds for some $s_0 > 0$ and that $p(x,D)$ extends to a continuous operator*

$$p(x,D) : H^{s_0+2,\lambda}(\mathbb{R}^n) \to H^{s_0,\lambda}(\mathbb{R}^n).$$

If for all $T > 0$ and for all $\Psi \in C_0^\infty([0,T) \times \mathbb{R}^n) \subset C^1([0,T], H^{s_0,\lambda}(\mathbb{R}^n))$ the evolution equation in $H^{s_0,\lambda}(\mathbb{R}^n)$

$$\begin{aligned} u' - p(x,D)u &= -\Psi(\cdot) \qquad \text{on } [0,T), \\ u(T) &= 0 \end{aligned} \tag{5.8}$$

has a solution $u \in C([0,T], H^{s_0+2,\lambda}(\mathbb{R}^n)) \cap C^1([0,T], H^{s_0,\lambda}(\mathbb{R}^n))$, then for any initial distribution there is at most one solution of the martingale problem for $-p(x,D)$.

The proof uses the following lemma which concerns a time dependent formulation of the martingale problem as considered in [83], Theo.4.2.1, see also [81], Theo.1.1 for the non-local case. Here we consider the case of operators defined in spaces $H^{s,\lambda}(\mathbb{R}^n)$.

Lemma 5.6. *Let s_0 and $p(x,D)$ be as in Theorem 5.5 and let $P \in \mathcal{M}_1(D_{\mathbb{R}^n})$be a solution of the martingale problem for the operator $(-p(x,D), C_0^\infty(\mathbb{R}^n))$. Then for each $T > 0$ and each $u \in C([0,T], H^{s_0+2,\lambda}(\mathbb{R}^n)) \cap C^1([0,T], H^{s_0,\lambda}(\mathbb{R}^n))$ the process $(M_t)_{t\geq 0}$,*

$$M_t = u(t, X_t) - \int_0^t \left(\left(\frac{d}{ds} - p(x,D) \right) u \right)(s, X_s)\, ds, \qquad 0 \leq t \leq T, \tag{5.9}$$

is a P-martingale up to time T.

Proof: First note that for all $\varphi \in H^{s_0+2,\lambda}(\mathbb{R}^n)$ the process

$$\varphi(X_t) - \int_0^t (-p(x,D)\varphi)(X_s)\, ds, \qquad t \geq 0, \tag{5.10}$$

is a martingale under P, because for a sequence of testfunctions $(\varphi_k)_{k\in\mathbb{N}}$ that approximates φ in $H^{s_0+2,\lambda}(\mathbb{R}^n)$ we know by Sobolev embedding that (φ_k) tends to φ and $(p(x,D)\varphi_k)$ tends to $p(x,D)\varphi$ uniformly. Therefore the martingale property is preserved.

Next note that we may assume $u \in C^1([0,T], H^{s_0+2,\lambda}(\mathbb{R}^n))$. In fact, given u as in the assumption we extend u by reflexion to an open neighbourhood I of $[0,T]$ in such a way that $u \in C(I, H^{s_0+2,\lambda}(\mathbb{R}^n)) \cap C^1(I, H^{s_0,\lambda}(\mathbb{R}^n))$. For some $\varphi \in C_0^\infty(\mathbb{R})$, $\varphi \geq 0$, $\int_{\mathbb{R}} \varphi(s)\, ds = 1$, and small $\varepsilon > 0$ we define

$$u_\varepsilon(t) = \int_I \frac{1}{\varepsilon} \varphi\left(\frac{t-s}{\varepsilon}\right) u(s)\, ds, \qquad 0 \leq t \leq T,$$

as a Bochner integral. Then $u_\varepsilon \in C^1([0,T], H^{s_0+2,\lambda}(\mathbb{R}^n))$ and a straight forward calculation yields that $u_\varepsilon \to u$ as $\varepsilon \to 0$ in $C([0,T], H^{s_0+2,\lambda}(\mathbb{R}^n))$ as well as in $C^1([0,T], H^{s_0,\lambda}(\mathbb{R}^n))$.

Hence, by (5.7) u_ε, u'_ε and $p(x,D)u_\varepsilon$ converge boundedly and uniformly to u, u' and $p(x,D)u$, respectively, considered as functions in $C_b([0,T]\times\mathbb{R}^n)$. But the martingale property of (5.9) is preserved under this convergence.

Therefore suppose that $u \in C^1([0,T], H^{s_0+2,\lambda}(\mathbb{R}^n))$. We now adapt the argument given in the proof of [83], Theorem 4.2.1, taking into account that by (5.6) we can calculate u' considered as an element of $C_b([0,T]\times\mathbb{R}^n)$ as the usual partial time derivative, i.e. $u' = \frac{\partial}{\partial t}u(\cdot,\cdot)$. For all $0 \le t_1 \le t_2 \le T$ and $A \in \mathcal{F}_{t_1}$ we get

$$\begin{aligned}
&\mathbb{E}^P\left[(u(t_2,X_{t_2}) - u(t_1,X_{t_1}))\cdot 1_A\right]\\
&= \mathbb{E}^P\left[(u(t_2,X_{t_2}) - u(t_1,X_{t_2}))\cdot 1_A\right] + \mathbb{E}^P\left[(u(t_1,X_{t_2}) - u(t_1,X_{t_1}))\cdot 1_A\right].
\end{aligned}$$

Since $u(t_1) \in H^{s_0+2,\lambda}(\mathbb{R}^n)$, we obtain by the remark (5.10) above

$$\begin{aligned}
&\mathbb{E}^P\left[(u(t_2,X_{t_2}) - u(t_1,X_{t_1}))\cdot 1_A\right]\\
&= \mathbb{E}^P\left[\int_{t_1}^{t_2} u'(s,X_{t_2})\,ds\cdot 1_A\right] + \mathbb{E}^P\left[\int_{t_1}^{t_2}(-p(x,D)u)(t_1,X_v)\,dv\cdot 1_A\right]\\
&= \mathbb{E}^P\left[\int_{t_1}^{t_2}\left(\frac{d}{ds} - p(x,D)\right)u(s,X_s)\,ds\cdot 1_A\right]\\
&\quad+\mathbb{E}^P\left[\int_{t_1}^{t_2}(u'(s,X_{t_2}) - u'(s,X_s))\,ds\cdot 1_A\right.\\
&\qquad\left.+\int_{t_1}^{t_2}((-p(x,D)u)(t_1,X_v) - (-p(x,D)u)(v,X_v))\;dv\cdot 1_A\right].
\end{aligned}$$

Thus, to prove the lemma we have to show that the second expectation vanishes. But again by (5.10)

$$\mathbb{E}^P\left[\int_{t_1}^{t_2}(u'(s,X_{t_2}) - u'(s,X_s))\,ds\cdot 1_A\right] = \mathbb{E}^P\left[\int_{t_1}^{t_2}\left(\int_s^{t_2}(-p(x,D)u')(s,X_v)\,dv\right)\,ds\cdot 1_A\right].$$

On the other hand

$$\begin{aligned}
&\mathbb{E}^P\left[\int_{t_1}^{t_2}((-p(x,D)u)(t_1,X_v) - (-p(x,D)u)(v,X_v))\;dv\cdot 1_A\right]\\
&= -\mathbb{E}^P\left[\int_{t_1}^{t_2}\left(\int_{t_1}^{v}\frac{d}{ds}(-p(x,D)u)(s,X_v)\,ds\right)\,dv\cdot 1_A\right].
\end{aligned}$$

Note that $u' \in C([0,T], H^{s_0+2,\lambda}(\mathbb{R}^n))$, $p(x,D)u \in C^1([0,T], H^{s_0,\lambda}(\mathbb{R}^n))$ and that $p(x,D)u' = \frac{d}{dt}(p(x,D)u)$ as elements of $C([0,T], H^{s_0,\lambda}(\mathbb{R}^n))$ and hence by (5.7) also as elements of $C_b([0,T]\times\mathbb{R}^n)$. Therefore, since both iterated integrals are intergals over the same region, the integrals cancel. □

Proof of Theorem 5.5: Let $\Psi_1 \in C_0^\infty([0,T))$ and $\Psi_2 \in C_0^\infty(\mathbb{R}^n)$ and $u \in C([0,T], H^{s_0+2,\lambda}(\mathbb{R}^n)) \cap C^1([0,T], H^{s_0,\lambda}(\mathbb{R}^n))$ be a solution to (5.8) with inhomogenity $\Psi(t,x) = \Psi_1(t)\cdot\Psi_2(x)$. Then by Lemma 5.6

$$M_t = u(t,X_t) - \int_0^t\left(\frac{d}{ds} - p(x,D)\right)u(s,X_s)\,ds$$

is a P-martingale up to time T for any solution P to the martingale problem for $-p(x,D)$ and initial distribution μ. Hence

$$\begin{aligned}&\int_0^T \Psi_1(s)\cdot \mathbb{E}^P\left[\Psi_2(X_s)\right]\,ds\\ =\;& \mathbb{E}^P\left[\int_0^T \Psi(s,X_s)\,ds\right] = \mathbb{E}^P\left[M_T\right] = \mathbb{E}^P\left[M_0\right] = \mathbb{E}^P\left[u(0,X_0)\right] = \mathbb{E}^\mu\left[u(0,\cdot)\right]\end{aligned}$$

and the expectation on the left hand side is uniquely determined by the initial distribution μ. Since Ψ_1 is arbitrary und the paths of (X_t) are right-continuous, the same holds true for $\mathbb{E}^P\left[\Psi_2(X_s)\right]$ for all $0 \le s < T$. Now Ψ_2 and T are also arbitrary, so we conclude that the one-dimensional distributions $P \circ X_s^{-1}$ are uniquely determined by μ for all $s \ge 0$ and do not depend on the choice of a particular solution to the martingale problem. Now the theorem follow from Lemma 5.4 □

5.3 Well-posedness of the martingale problem for a class of pseudo differential operators

Assume again that

$$\psi : \mathbb{R}^n \to \mathbb{R}$$

is a continuous negative definite reference function, which satisfies

$$\psi(\xi) \ge c\,|\xi|^r \qquad \text{for all } |\xi| \ge 1 \tag{5.11}$$

for some $r > 0$ and $c > 0$. Define $\lambda(\xi) = (1+\psi(\xi))^{1/2}$ as ususal. Our aim is to prove the following uniqueness result for the martingale problem.

Theorem 5.7. *Let $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite symbol such that $p(x,0) = 0$ for all $x \in \mathbb{R}^n$. Let $M \in \mathbb{N}$ be the smallest integer such that*

$$M > \left(\frac{n}{r} \vee 2\right) + n \tag{5.12}$$

and suppose that

(i) $p(x,\xi)$ is $(2M+1-n)$-times continuously differentiable with respect to x and for all $\beta \in \mathbb{N}_0^n$, $|\beta| \le 2M+1-n$

$$\left|\partial_x^\beta p(x,\xi)\right| \le c\lambda^2(\xi), \qquad x, \in \mathbb{R}^n,\ \xi \in \mathbb{R}^n, \tag{5.13}$$

holds and

(ii) for some strictly positive function $\gamma : \mathbb{R}^n \to \mathbb{R}^+$

$$p(x,\xi) \ge \gamma(x)\cdot\lambda^2(\xi) \qquad \textit{for all } x \in \mathbb{R}^n,\ |\xi| \ge 1. \tag{5.14}$$

Then the martingale problem for the operator $-p(x, D)$ with domain $C_0^\infty(\mathbb{R}^n)$ is well-posed.

We want to apply Theorem 5.5 to prove this result. It is well-known (see [84]) how to solve the evolution equation (5.8) using a semigroup which is generated by the operator $-p(x, D)$ in $H^{s_0,\lambda}(\mathbb{R}^n)$. So we are again led to the investigation of the operator $-p(x, D)$ in the scale of anisotropic Sobolev spaces. In particular, by the Hille-Yosida theorem we are led to the problem of solving the equation (4.41). To attack this problem it is reasonable in view of our localization result, Theorem 5.3, and in view of the conditions we had to assume in the previous chapter, to modify the symbol $p(x, \xi)$ in the following way:

Fix $x_0 \in \mathbb{R}^n$ and $\Psi \in C_0^\infty(\mathbb{R}^n)$ such that $0 \le \Psi \le 1$, $\Psi = 1$ in $B_{1/2}(0)$, $\operatorname{supp}\Psi \subset B_1(0)$, and define for $R > 0$

$$\Psi_R(x) = \Psi\left(\frac{x - x_0}{R}\right). \tag{5.15}$$

Now consider

$$\begin{aligned} p_R(x, \xi) &= \Psi_R(x) \cdot p(x, \xi) + (1 - \Psi_R(x)) \cdot p(x_0, \xi) \\ &= p(x_0, \xi) + \Psi_R(x)(p(x, \xi) - p(x_0, \xi)). \end{aligned} \tag{5.16}$$

Then $p_R(x, \xi)$ coincides with $p(x, \xi)$ in the neighbourhood $B_{R/2}(x_0)$ of x_0. Moreover, p_R as a convex combination of continuous negative definite symbols is also a continuous negative definite symbol and has a decomposition as in (4.26) with a perturbation part

$$p_2(x, \xi) = \Psi_R(x)(p(x, \xi) - p(x_0, \xi)). \tag{5.17}$$

When R is chosen sufficiently small, we may hope that this perturbation also becomes small and the method of the previous chapter applies.

Unfortunately, the situation is more delicate, since condition (A.3.M) requires smallness not only for $p_2(x, \xi)$, but also for the derivatives with respect to x. But for $p_2(x, \xi)$ chosen as in (5.17) these derivatives in general explode as $R \to 0$. We therefore have to refine the decomposition. With the notation of Theorem 5.7 let N be the smallest integer such that

$$N > \left(\frac{n}{r} \vee 2\right),$$

i.e. $M = N + n$. We obtain by Taylor's formula

$$p(x, \xi) = p(x_0, \xi) + \sum_{0<|\beta|\le N} \frac{(x - x_0)^\beta}{\beta!} \partial_x^\beta p(x_0, \xi) + R_N(x, \xi),$$

where

$$R_N(x, \xi) = (N + 1) \int_0^1 (1 - t)^N \sum_{|\beta|=N+1} \frac{(x - x_0)^\beta}{\beta!} (\partial_x^\beta p)(tx + (1 - t)x_0, \xi)\, dt.$$

A comparison with (5.16) thus yields

$$\begin{aligned} p_R(x, \xi) &= p(x_0, \xi) + \sum_{0<|\beta|\le N} \Psi_R(x) \frac{(x - x_0)^\beta}{\beta!} \partial_x^\beta p(x_0, \xi) + \Psi_R(x) \cdot R_N(x, \xi) \\ &= p(x_0, \xi) + \sum_{0<|\beta|\le N} b_\beta(x) p_\beta(\xi) + \sum_{|\beta|=N+1} q_\beta(x, \xi), \end{aligned} \tag{5.18}$$

where for $\beta \in \mathbb{N}_0^n$ we introduced the notations

$$\begin{aligned}
b_\beta(x) &= \Psi_R(x)\frac{(x-x_0)^\beta}{\beta!}, \qquad 0<|\beta|\le N,\\
p_\beta(\xi) &= \partial_x^\beta p(x_0,\xi), \qquad 0<|\beta|\le N,\\
q_\beta(x,\xi) &= (N+1)\int_0^1 (1-t)^N \Psi_R(x)\frac{(x-x_0)^\beta}{\beta!}(\partial_x^\beta p)(tx+(1-t)x_0,\xi)\,dt, \qquad |\beta|=N+1.
\end{aligned}$$

Note that b_β and q_β depend on R, although it is not marked to avoid a too complex notation. Nevertheless we carefully have to keep track of the dependence of estimates for the corresponding operators on R.
The advantage of the decomposition (5.18) lies in the fact that we can force $q_\beta(x,\xi)$ to be small for $R\to 0$ because of the factor $\Psi_R(x)(x-x_0)^\beta$, $|\beta|=N+1$, which vanishes arbitrarily fast if N is sufficienty large. On the other hand also the terms $b_\beta(x)\cdot p_\beta(\xi)$ will be controllable due to their simpler product structure.

In order to prove Theorem 5.7 note first that (5.13) for $\beta=0$ implies by Theorem 3.15 that there is a solution to the martingale problem for $-p(x,D)$ for every initial distribution. Then for $x_0\in\mathbb{R}^n$ define the symbol $p_R(x,\xi)$ as in (5.16). Note that again by (5.13) the operators $p(x,D)$ and $p_R(x,D)$ map $C_0^\infty(\mathbb{R}^n)$ into $C_b(\mathbb{R}^n)$. If we can prove that for each $x_0\in\mathbb{R}^n$ there is an $R=R(x_0)>0$ such that the martingale problem for $-p_R(x,D)$ is well-posed, then we can choose a countable set of points x_0 in such way that the balls $B_{R(x_0)/2}(x_0)$ cover the whole space $\mathbb{R}^n$. But then by Theorem 5.3 also the martingale problem for $-p(x,D)$ is well-posed.
Thus the proof of Theorem 5.7 is reduced to verify

Theorem 5.8. *Let $p(x,\xi)$ be as in Theorem 5.7. Fix $x_0\in\mathbb{R}^n$ and define $p_R(x,\xi)$ as in (5.16). Then there is an $R>0$ such that the martingale problem for $-p_R(x,D)$ with domain $C_0^\infty(\mathbb{R}^n)$ is well-posed.*

As mentioned above we want to apply modified Hilbert space methods as in Chapter 4 to solve an evolution equation corresponding to $p_R(x,D)$ and then apply Theorem 5.5. To that end we need some preparations. We begin with commutator estimates for each term of the decomposition (5.18) of $p_R(x,\xi)$. Recall the defintion of M in (5.12).

Lemma 5.9. *Let $0\le s<M-n$. Then there is a constant $K_1(R)$ with $\lim_{R\to 0}K_1(R)=0$ such that for all $|\beta|=N+1$*

$$\|[\lambda^s(D),q_\beta(x,D)]\,u\|_0 \le K_1(R)\cdot\|u\|_{s+1,\lambda} \qquad \text{for all } u\in C_0^\infty(\mathbb{R}^n). \tag{5.19}$$

Proof: By (5.12) we know that $2+n<M$. Hence for all $0\le s<M-n$ we have $|s-1|+1+n<M$. Therefore Theorem 4.3 yields (5.19) provided

$$|\partial_x^\alpha q_\beta(x,\xi)|\le \Phi_\alpha(x)\cdot\lambda^2(\xi) \qquad \text{for all } \alpha\in\mathbb{N}_0^n,\ |\alpha|\le M, \tag{5.20}$$

for functions $\Phi_\alpha\in L^1(\mathbb{R}^n)$ and the constant $K_1(R)$ then is given by

$$K_1(R)=c_{M,s,\psi}\cdot\sum_{\substack{\alpha\in\mathbb{N}_0^n\\|\alpha|\le M}}\|\Phi_\alpha\|_{L^1(\mathbb{R}^n)}\,. \tag{5.21}$$

But by definition of q_β, Leibniz rule and assumption (5.13)

$$
\begin{aligned}
&|\partial_x^\alpha q_\beta(x,\xi)| \leq \\
&\leq \ (N+1)\int_0^1 \left|(1-t)^N \partial_x^\alpha \left[\Psi_R(x)\frac{(x-x_0)^\beta}{\beta!}(\partial_x^\beta p)(tx+(1-t)x_0,\xi)\right]\right| \, dt \\
&= \ (N+1)\cdot \int_0^1 \left|(1-t)^N \sum_{\gamma\leq\alpha}\binom{\alpha}{\gamma}\partial_x^{\alpha-\gamma}\left[\Psi_R(x)\frac{(x-x_0)^\beta}{\beta!}\right] t^{|\gamma|}(\partial_x^{\beta+\gamma}p)(tx+(1-t)x_0,\xi)\right| \, dt \\
&\leq \ (N+1)\int_0^1 (1-t)^N \sum_{\gamma\leq\alpha}\binom{\alpha}{\gamma}\left|\partial_x^{\alpha-\gamma}\left[\Psi_R(x)\frac{(x-x_0)^\beta}{\beta!}\right]\right| t^{|\gamma|}\cdot c\lambda^2(\xi)\, dt \\
&\leq \ \sum_{\gamma\leq\alpha}\binom{\alpha}{\gamma}\cdot\left|\partial_x^{\alpha-\gamma}\left[\Psi_R(x)\frac{(x-x_0)^\beta}{\beta!}\right]\right|\cdot c\,\lambda^2(\xi).
\end{aligned}
$$

Note that because of $|\beta+\gamma|\leq N+1+M = 2M+1-n$ all derivatives of p exist. Therefore (5.20) holds with L^1-functions

$$\Phi_\alpha(x) = c\cdot\sum_{\gamma\leq\alpha}\binom{\alpha}{\gamma}\cdot\left|\partial_x^{\alpha-\gamma}\left[\Psi_R(x)\frac{(x-x_0)^\beta}{\beta!}\right]\right|$$

and the assertion is proved as soon as we show

$$\lim_{R\to 0}\sup_{\gamma\leq\alpha}\int_{\mathbb{R}^n}\left|\partial_x^{\alpha-\gamma}\left[\Psi_R(x)\frac{(x-x_0)^\beta}{\beta!}\right]\right| dx = 0.$$

But again by Leibniz rule

$$
\begin{aligned}
&\sup_{\gamma\leq\alpha}\int_{\mathbb{R}^n}\left|\partial_x^{\alpha-\gamma}\left[\Psi_R(x)\frac{(x-x_0)^\beta}{\beta!}\right]\right| dx \\
&\quad\leq \ \sup_{|\gamma|\leq M}\int_{B_R(x_0)}\sum_{\delta\leq\gamma}\binom{\gamma}{\delta}\left|\partial_x^\delta\Psi_R(x)\right|\cdot\left|\partial_x^{\gamma-\delta}\frac{(x-x_0)^\beta}{\beta!}\right| dx \\
&\quad\leq \ \sup_{|\gamma|\leq M}\left\{\sum_{\delta\leq\gamma}\binom{\gamma}{\delta}\int_{B_R(x_0)}R^{-|\delta|}\left|(\partial_x^\delta\Psi)(\frac{x-x_0}{R})\right| dx\cdot\sup_{x\in B_R(x_0)}\left|\partial_x^{\gamma-\delta}\frac{(x-x_0)^\beta}{\beta!}\right|\right\} \\
&\quad\leq \ \sup_{|\gamma|\leq M}\left\{\sum_{\substack{\delta\leq\gamma\\ \gamma-\delta\leq\beta}}\binom{\gamma}{\delta}R^{-|\delta|+n}\int_{B_1(0)}\left|\partial_x^\delta\Psi(x)\right| dx\cdot c_{\beta,\gamma,\delta}\sup_{x\in B_R(x_0)}\left|(x-x_0)^{\beta-(\gamma-\delta)}\right|\right\},
\end{aligned}
$$

where it is sufficient to let the sum run over $\gamma-\delta\leq\beta$, since otherwise $\partial_x^{\gamma-\beta}(x-x_0)^\beta$ vanishes. In particular $|\beta-(\gamma-\delta)| = |\beta|-|\gamma|+|\delta|$ and so

$$\sup_{\gamma\leq\alpha}\int_{\mathbb{R}^n}\left|\partial_x^{\alpha-\gamma}\left[\Psi_R(x)\frac{(x-x_0)^\beta}{\beta!}\right]\right| dx$$

$$\leq \sup_{|\gamma|\leq M} \left\{ \sum_{\substack{\delta\leq\gamma\\ \gamma-\delta\leq\beta}} \binom{\gamma}{\delta} c_{\beta,\gamma,\delta} R^{-|\delta|+n} \int_{B_1(0)} \left|\partial_x^\delta \Psi(x)\right| \, dx \cdot R^{|\beta|-|\gamma|+|\delta|} \right\},$$

$$\leq c \sup_{|\gamma|\leq M} R^{n+|\beta|-|\gamma|} \to 0,$$

as $R \to 0$, because $n + |\beta| - |\gamma| \geq n + N + 1 - M = 1$ by the choice of N. □

Recall that $b_\beta(x) = \Psi_R(x)\frac{(x-x_0)^\beta}{\beta!}$, $0 < |\beta| \leq N$ is a smooth function with compact support. Hence for all $\alpha \in \mathbb{N}_0^n$

(5.22) $$|\partial_x^\alpha (b_\beta(x) \cdot p_\beta(\xi))| \leq |\partial_x^\alpha b_\beta(x)| \cdot c\,\lambda^2(\xi)$$

and

$$\|\partial_x^\alpha b_\beta\|_{L^1(\mathbb{R}^n)} \leq C(R).$$

Therefore we find as above

Lemma 5.10. *Let $s \geq 0$. Then for all $0 < |\beta| \leq N$*

$$\|[\lambda^s(D), b_\beta(\cdot)p_\beta(D)]\, u\|_0 \leq C(R)\, \|u\|_{s+1,\lambda} \quad \textit{for all } u \in C_0^\infty(\mathbb{R}^n).$$

We also have commutator estimates for the Friedrichs mollifier analogous to that of Theorem 4.4.

Lemma 5.11.

(i) Let $0 \leq s < M - n$. There is a constant $C(R)$ independent of $0 < \varepsilon \leq 1$ such that for all $|\beta| = N + 1$

$$\|[J_\varepsilon, q_\beta(x, D)]\, u\|_{s,\lambda} \leq C(R)\, \|u\|_{s+1,\lambda} \quad \textit{for all } u \in C_0^\infty(\mathbb{R}^n).$$

(ii) Let $s \geq 0$ .There is a constant $C(R)$ independent of $0 < \varepsilon \leq 1$ such that for all $0 < |\beta| \leq N$

$$\|[J_\varepsilon, b_\beta(\cdot)p_\beta(D)]\, u\|_{s,\lambda} \leq C(R)\, \|u\|_{s+1,\lambda} \quad \textit{for all } u \in C_0^\infty(\mathbb{R}^n).$$

For the proof it is enough to note that the estimates (5.20) in the proof of Lemma 5.9 and (5.22), respectively, are exactly the assumptions of Theorem 4.4

We now turn to the solution of the equation

(5.23) $$(p_R(x, D) + \tau)u = f.$$

For this purpose in principle we want to repeat the arguments of chapter 4. But the commutator estimates in Lemma 5.10 show that the operators $b_\beta(\cdot)p_\beta(D)$ in the decomposition (5.18) cannot be treated as in chapter 4, since the constant $C(R)$ may be very large for small R, and there is still a little more work to do. Since we are interested in a semigroup on a space $H^{s,\lambda}(\mathbb{R}^n)$, we will construct a weak solution of (5.23) already in this space. We begin with some continuity properties of the operators under consideration.

Lemma 5.12.

(i) Let $s \geq 0$. Then for all $0 < |\beta| \leq N$

$$\|p(x_0, D)u\|_{s,\lambda} \leq c\,\|u\|_{s+2,\lambda}\,, \tag{5.24}$$

$$\|p_\beta(D)u\|_{s,\lambda} \leq c\,\|u\|_{s+2,\lambda} \tag{5.25}$$

and

$$\|b_\beta(\cdot)p_\beta(D)u\|_{s,\lambda} \leq C(R)\,\|u\|_{s+2,\lambda} \qquad \text{for all } u \in C_0^\infty(\mathbb{R}^n). \tag{5.26}$$

(ii) Let $0 \leq s < M - n$. Then there is a constant $K_2(R)$ with $\lim\limits_{R\to 0} K_2(R) = 0$, such that

$$\|q_\beta(x, D)u\|_{s,\lambda} \leq K_2(R)\,\|u\|_{s+2,\lambda} \tag{5.27}$$

for all $|\beta| = N + 1$ and $u \in C_0^\infty(\mathbb{R}^n)$.

Proof: The estimates (5.24) and (5.25) follow by the same arguments as in (4.32). Moreover, by Lemma 5.10 and (5.25)

$$\begin{aligned}\|b_\beta(\cdot)p_\beta(D)u\|_{s,\lambda} &= \|\lambda^s(D)b_\beta(\cdot)p_\beta(D)u\|_0 \\ &\leq \|b_\beta(\cdot)p_\beta(D)\lambda^s(D)u\|_0 + \|[\lambda^s(D), b_\beta(\cdot)p_\beta(D)]\,u\|_0 \\ &\leq c\,\|b_\beta\|_{L^\infty(\mathbb{R}^n)}\,\|u\|_{s+2,\lambda} + C(R)\,\|u\|_{s+1,\lambda}\,,\end{aligned} \tag{5.28}$$

which gives (5.26).
For $|\beta| = N + 1$ we find as in the proof of Theorem 4.8, (4.33),

$$\|q_\beta(x, D)u\|_0 \leq C_M \cdot \left\|\langle\cdot\rangle^{-M}\right\|_{L^1(\mathbb{R}^n)} \cdot \sum_{\substack{\alpha\in\mathbb{N}_0^n \\ |\alpha|\leq M}} \|\Phi_\alpha\|_{L^1(\mathbb{R}^n)} \cdot \|u\|_{2,\lambda}\,,$$

where $\Phi_\alpha \in L^1(\mathbb{R}^n)$ are the functions in the estimate (5.20). But as shown in the proof of Lemma 5.9 we have $\lim_{R\to 0} \|\Phi_\alpha\|_{L^1(\mathbb{R}^n)} = 0$ and hence

$$\|q_\beta(x, D)u\|_0 \leq K_R \cdot \|u\|_{2,\lambda}$$

with a constant K_R such that $\lim_{R\to 0} K_R = 0$. Furthermore, by Lemma 5.9

$$\begin{aligned}\|q_\beta(x, D)u\|_{s,\lambda} &= \|\lambda^s(D)q_\beta(x, D)u\|_0 \\ &\leq \|q_\beta(x, D)\lambda^s(D)u\|_0 + \|[\lambda^s(D), q_\beta(x, D)]\,u\|_0 \\ &\leq K_R \cdot \|\lambda^s(D)u\|_{2,\lambda} + K_1(R)\,\|u\|_{s+1,\lambda} \\ &\leq (K_R + K_1(R))\,\|u\|_{s+2,\lambda}\,.\end{aligned}$$

□

In particular $p_R(x, D)$ maps $H^{s+2,\lambda}(\mathbb{R}^n)$ continuously into $H^{s,\lambda}(\mathbb{R}^n)$ for all $0 \leq s < M - n$.

We now fix a number $s_0 \geq 2$ such that

$$M - n > s_0 > \frac{n}{2r}. \tag{5.29}$$

This is possible by (5.12). Note that by Proposition 4.1 in particular the Sobolev embedding (5.6) and hence the assumption on s_0 in Theorem 5.5 hold with this choice of s_0. We investigate the equation

$$(p_R(x,D)+\lambda)u = f$$

in $H^{s_0,\lambda}(\mathbb{R}^n)$. First we look for a weak solution. Therefore define for $u,v \in C_0^\infty(\mathbb{R}^n)$ the bilinear form

$$B^R(u,v) = (p_R(x,D)u,v)_{s_0,\lambda}. \tag{5.30}$$

We decompose B^R in the following way.

$$B^R(u,v) = B_0^R(u,v) + \sum_{0<|\beta|\le N} B_\beta^R(u,v) + \sum_{|\beta|=N+1} B_\beta^R(u,v), \tag{5.31}$$

where

$$\begin{aligned} B_0^R(u,v) &= (p(x_0,D)u,v)_{s_0,\lambda}, \\ B_\beta^R(u,v) &= (b_\beta(\cdot)p_\beta(D)u,v)_{s_0,\lambda}, \quad 0<|\beta|\le N, \\ B_\beta^R(u,v) &= (q_\beta(x,D)u,v)_{s_0,\lambda}, \quad |\beta| = N+1. \end{aligned}$$

Proposition 5.13. *The bilinear form B^R extends continuously to $H^{s_0+1,\lambda}(\mathbb{R}^n)\times H^{s_0+1,\lambda}(\mathbb{R}^n)$ and*

$$\left|B^R(u,v)\right| \le C(R)\,\|u\|_{s_0+1,\lambda}\cdot\|u\|_{s_0+1,\lambda} \tag{5.32}$$

holds.

Proof: We estimate all terms of the decomposition (5.31) in turn. By Cauchy-Schwarz inequality and Lemma 5.12 we obtain for $u,v\in C_0^\infty(\mathbb{R}^n)$

$$\begin{aligned} \left|B_0^R(u,v)\right| &= |(\lambda^{s_0}(D)p(x_0,D)u,\lambda^{s_0}(D)v)_0| \\ &= \left|(\lambda^{s_0-1}(D)p(x_0,D)u,\lambda^{s_0+1}(D)v)_0\right| \\ &\le \|p(x_0,D)u\|_{s_0-1,\lambda}\cdot\|v\|_{s_0+1,\lambda} \\ &\le c\,\|u\|_{s_0+1,\lambda}\cdot\|v\|_{s_0+1,\lambda}\,. \end{aligned} \tag{5.33}$$

Using Lemma 5.10 and Lemma 5.12 we find analogously for $0<|\beta|\le N$

$$\begin{aligned} \left|B_\beta^R(u,v)\right| &= \left|(\lambda^{s_0-1}(D)b_\beta(\cdot)p_\beta(D)u,\lambda^{s_0+1}(D)v)_0\right| \\ &\le \left|(b_\beta(\cdot)p_\beta(D)\lambda^{s_0-1}(D)u,\lambda^{s_0+1}(D)v)_0\right| \\ &\quad + \left|(\left[\lambda^{s_0-1}(D),b_\beta(\cdot)p_\beta(D)\right]u,\lambda^{s_0+1}(D)v)_0\right| \\ &\le c\,\|b_\beta\|_{L^\infty(\mathbb{R}^n)}\cdot\|u\|_{s_0+1,\lambda}\cdot\|v\|_{s_0+1,\lambda} + C(R)\,\|u\|_{s_0,\lambda}\cdot\|v\|_{s_0+1,\lambda}\,. \end{aligned} \tag{5.34}$$

Finally we have for $|\beta| = N+1$ again by Lemma 5.12

$$\begin{aligned} \left|B_\beta^R(u,v)\right| &= \left|(\lambda^{s_0-1}(D)q_\beta(x,D)u,\lambda^{s_0+1}(D)v)_0\right| \\ &\le \|q_\beta(x,D)u\|_{s_0-1,\lambda}\cdot\|v\|_{s_0+1,\lambda} \le K_2(R)\,\|u\|_{s_0+1,\lambda}\cdot\|v\|_{s_0+1,\lambda}\,. \end{aligned} \tag{5.35}$$

Combining (5.33), (5.34) and (5.35) completes the proof. □

Condition (5.14) allows to prove a lower estimate for B^R given by the following Garding inequality.

Theorem 5.14. *If $R > 0$ is sufficiently small, then there is a constant $\tau(R)$ depending on R such that*

$$B^R(u,u) \geq \frac{\gamma(x_0)}{2} \|u\|^2_{s_0+1,\lambda} - \tau(R) \|u\|^2_{s_0,\lambda} \qquad \text{for all } u \in H^{s_0+1,\lambda}(\mathbb{R}^n). \tag{5.36}$$

Proof: By (5.14) there is a constant $c \geq 0$, such that

$$p(x_0, \xi) \geq \gamma(x_0)\lambda^2(\xi) - c$$

and therefore as in (4.38)

$$B^R_0(u,u) \geq \gamma(x_0) \|u\|^2_{s_0+1,\lambda} - c \|u\|^2_{s_0,\lambda}. \tag{5.37}$$

Moreover, by (5.34) for all $\varepsilon > 0$ we get for $0 < |\beta| \leq N$

$$\begin{aligned} \left|B^R_\beta(u,u)\right| &\leq c \|b_\beta\|_{L^\infty(\mathbb{R}^n)} \cdot \|u\|^2_{s_0+1,\lambda} + C(R) \|u\|_{s_0+1,\lambda} \cdot \|u\|_{s_0,\lambda} \\ &\leq (c \|b_\beta\|_{L^\infty(\mathbb{R}^n)} + \varepsilon) \|u\|^2_{s_0+1,\lambda} + C(R,\varepsilon) \|u\|^2_{s_0,\lambda}. \end{aligned}$$

Since

$$\|b_\beta\|_{L^\infty(\mathbb{R}^n)} = \sup_{x \in B_R(x_0)} \left| \Psi_R(x) \frac{(x-x_0)^\beta}{\beta!} \right| \to 0 \quad \text{as } R \to 0, \tag{5.38}$$

we find for R and ε sufficiently small

$$\sum_{0<|\beta|\leq N} \left|B^R_\beta(u,u)\right| \leq \frac{\gamma(x_0)}{4} \|u\|^2_{s_0+1,\lambda} + C(R) \|u\|^2_{s_0,\lambda} \tag{5.39}$$

and furthermore by (5.35)

$$\sum_{|\beta|=N+1} \left|B^R_\beta(u,u)\right| \leq \sum_{|\beta|=N+1} K_2(R) \|u\|^2_{s_0+1,\lambda} \leq \frac{\gamma(x_0)}{4} \|u\|^2_{s_0+1,\lambda}. \tag{5.40}$$

Summarizing (5.37), (5.39) and (5.40) we find

$$\begin{aligned} B^R(u,u) &\geq B^R_0(u,u) - \sum_{0<|\beta|\leq N} \left|B^R_\beta(u,u)\right| - \sum_{|\beta|=N+1} \left|B^R_\beta(u,u)\right| \\ &\geq \frac{\gamma(x_0)}{2} \|u\|^2_{s_0+1,\lambda} - \tau(R) \|u\|^2_{s_0,\lambda}. \end{aligned}$$

□

Let us fix R sufficiently small such that (5.36) holds true. Then for all $\tau \geq \tau(R)$ by Proposition 5.13 and Theorem 5.14 $B^R(\cdot,\cdot) + \tau(\cdot,\cdot)_{s_0,\lambda}$ is a continuous and coercive bilinear form on $H^{s_0+1,\lambda}(\mathbb{R}^n)$. Thus by the Lax-Milgram theorem ([96], p.92) for each $f \in H^{s_0,\lambda}(\mathbb{R}^n)$ we find a weak solution $u \in H^{s_0+1,\lambda}(\mathbb{R}^n)$ of the equation $(p_R(x,D) + \tau)u = f$, i.e. there is a unique $u \in H^{s_0+1,\lambda}(\mathbb{R}^n)$ such that

$$B^R(u,v) + \tau(u,v)_{s_0,\lambda} = (f,v)_{s_0,\lambda} \quad \text{for all } v \in H^{s_0+1,\lambda}(\mathbb{R}^n). \tag{5.41}$$

Our aim is again to prove that for this solution $u \in H^{s_0+2,\lambda}(\mathbb{R}^n)$ holds. We claim

Proposition 5.15. *If $R > 0$ is sufficiently small, then there is a constant $C(R)$ such that*

$$\|u\|_{s_0+2,\lambda} \leq C(R) \left(\|p_R(x,D)u\|_{s_0,\lambda} + \|u\|_{s_0,\lambda} \right) \tag{5.42}$$

for all $u \in H^{s_0+2,\lambda}(\mathbb{R}^n)$.

Proof: As in (4.40) we have by (5.14)

$$\|p(x_0,D)u\|^2_{s_0,\lambda} \geq \gamma(x_0)^2 \|u\|^2_{s_0+2,\lambda} - c\|u\|^2_{s_0,\lambda}\,. \tag{5.43}$$

Next for $0 < |\beta| \leq N$ we obtain by (5.28) and (4.11) for all $\varepsilon > 0$

$$\begin{aligned} \|b_\beta(\cdot)p_\beta(D)u\|_{s_0,\lambda} &= \|\lambda^{s_0}(D)b_\beta(\cdot)p_\beta(D)u\|_0 \\ &\leq c\|b_\beta\|_{L^\infty(\mathbb{R}^n)}\|u\|_{s_0+2,\lambda} + C(R)\|u\|_{s_0+1,\lambda} \\ &\leq (c\|b_\beta\|_{L^\infty(\mathbb{R}^n)} + \varepsilon)\|u\|_{s_0+2,\lambda} + C(R,\varepsilon)\|u\|_{s_0,\lambda}\,, \end{aligned}$$

which gives for R and ε sufficiently small by (5.38)

$$\sum_{0<|\beta|\leq N} \|b_\beta(\cdot)p_\beta(D)u\|_{s_0,\lambda} \leq \frac{\gamma(x_0)}{4}\|u\|_{s_0+2,\lambda} + C(R)\|u\|_{s_0,\lambda}\,. \tag{5.44}$$

Finally by Lemma 5.12 (ii) for small $R > 0$

$$\sum_{|\beta|=N+1} \|q_\beta(x,D)u\|_{s_0,\lambda} \leq \sum_{|\beta|=N+1} K_2(R)\|u\|_{s_0+2,\lambda} \leq \frac{\gamma(x_0)}{4}\|u\|_{s_0+2,\lambda} \tag{5.45}$$

holds. Hence, combining (5.43), (5.44) and (5.45) yields

$$\begin{aligned} \|p_R(x,D)u\|_{s_0,\lambda} &\geq \|p(x_0,D)u\|_{s_0,\lambda} - \sum_{0<|\beta|\leq N} \|b_\beta(\cdot)p_\beta(D)u\|_{s_0,\lambda} - \sum_{|\beta|=N+1} \|q_\beta(x,D)u\|_{s_0,\lambda} \\ &\geq \frac{\gamma(x_0)}{2}\|u\|_{s_0+2,\lambda} - C(R)\|u\|_{s_0,\lambda}\,, \end{aligned}$$

i.e. (5.42). □

Now we can prove

Theorem 5.16. *Let $R > 0$ be sufficiently small and $\tau(R)$ be the constant given in Theorem 5.14. Then for all $\tau \geq \tau(R)$ and all $f \in H^{s_0,\lambda}(\mathbb{R}^n)$ there is a unique solution $u \in H^{s_0+2,\lambda}(\mathbb{R}^n)$ of*

$$(p_R(x,D) + \tau)u = f.$$

Proof: Choose $R > 0$ sufficiently small, such that the statements of Theorem 5.14 and Proposition 5.15 hold and let $u \in H^{s_0+1,\lambda}(\mathbb{R}^n)$ be the unique weak solution of (5.41), note that $s_0 \geq 2$. Choose functions $u_k, v \in C_0^\infty(\mathbb{R}^n)$, $k \in \mathbb{N}$, such that (u_k) converges to u in $H^{s_0+1,\lambda}(\mathbb{R}^n)$. Then

$$B^R(u_k, v) = (p_R(x,D)u_k, v)_{s_0,\lambda} = (p_R(x,D)u_k, \lambda^{2s_0}(D)v)_0.$$

For $k \to \infty$ we get $p_R(x,D)u_k \to p_R(x,D)u$ in $L^2(\mathbb{R}^n)$ by Lemma 5.12 and thus

$$(p_R(x,D)u, \lambda^{2s_0}(D)v)_0 = B^R(u,v) = (f - \tau u, v)_{s_0,\lambda} = (f - \tau u, \lambda^{2s_0}(D)v)_0.$$

Since $\lambda^{2s_0}(D)(C_0^\infty(\mathbb{R}^n))$ is dense in $L^2(\mathbb{R}^n)$, this gives $(p_R(x,D)+\tau)u = f$ in $H^{s_0,\lambda}(\mathbb{R}^n)$.
We claim $u \in H^{s_0+2,\lambda}(\mathbb{R}^n)$. Recall the definition of $(J_\varepsilon)_{\varepsilon>0}$ in (4.22). We know that $J_\varepsilon u \in H^{s_0+2,\lambda}(\mathbb{R}^n)$ for all $0 < \varepsilon \leq 1$ and $J_\varepsilon u$ converges to u in $H^{s_0,\lambda}(\mathbb{R}^n)$ as $\varepsilon \to 0$. Consequently $(J_\varepsilon u)_{0<\varepsilon\leq 1}$ is bounded in $H^{s_0,\lambda}(\mathbb{R}^n)$. The same holds true for $(p_R(x,D)J_\varepsilon u)_{0<\varepsilon\leq 1}$, because by Lemma 5.11

$$\begin{aligned}
&\|p_R(x,D)J_\varepsilon u\|_{s_0,\lambda} \leq \|J_\varepsilon p_R(x,D)u\|_{s_0,\lambda} + \|[J_\varepsilon, p_R(x,D)]\,u\|_{s_0,\lambda} \\
&\quad\leq \|p_R(x,D)u\|_{s_0,\lambda} + \sum_{0<|\beta|\leq N} \|[J_\varepsilon, b_\beta(\cdot)p_\beta(D)]\,u\|_{s_0,\lambda} + \sum_{|\beta|=N+1} \|[J_\varepsilon, q_\beta(x,D)]\,u\|_{s_0,\lambda} \\
&\quad\leq \|f\|_{s_0,\lambda} + \|\tau u\|_{s_0,\lambda} + C(R)\,\|u\|_{s_0+1,\lambda} < \infty.
\end{aligned}$$

So, by Proposition 5.15 we find that $(J_\varepsilon u)_{0<\varepsilon\leq 1}$ is bounded in $H^{s_0+2,\lambda}(\mathbb{R}^n)$ and this implies $u \in H^{s_0+2,\lambda}(\mathbb{R}^n)$. □

Finally we are in the position to conclude the proof of our uniqueness result.

Proof of Theorem 5.8:
Choose $R > 0$ sufficiently small such that Theorem 5.14 and Proposition 5.15 hold. Then the operator $-(p_R(x,D)+\tau(R))$ with domain $H^{s_0+2,\lambda}(\mathbb{R}^n)$ is a densely defined operator on $H^{s_0,\lambda}(\mathbb{R}^n)$ which by Garding's inequality (5.36) is dissipative. Moreover, Theorem 5.16 yields that the range of $-(p_R(x,D)+\tau(R)) - \tau$ is $H^{s_0,\lambda}(\mathbb{R}^n)$ for all $\tau > 0$. Thus, by the Hille-Yosida theorem, Theorem 4.6, $-p_R(x,D)$ is the generator of a strongly continuous semigroup $(S_t)_{t\geq 0}$ of linear operators on $H^{s_0,\lambda}(\mathbb{R}^n)$. Hence given $T > 0$ by [84], Theorem 3.2.2, for $\Psi \in C^1([0,T], H^{s_0,\lambda}(\mathbb{R}^n))$

$$u(t) = \int_t^T S_{s-t}\Psi(s)\,ds$$

defines a solution $u : [0,T] \to H^{s_0+2,\lambda}(\mathbb{R}^n)$ satisfying $u \in C^1([0,T], H^{s_0,\lambda}(\mathbb{R}^n))$ of the evolution equation (5.8) with $p(x,D)$ replaced by $p_R(x,D)$. Moreover, by (5.8) we have $p_R(x,D)u \in C([0,T], H^{s_0,\lambda}(\mathbb{R}^n))$ and thus the regularity result (5.42) gives $u \in C([0,T], H^{s_0+2,\lambda}(\mathbb{R}^n))$. Therefore by Theorem 5.5 the martingale problem for $-p_R(x,D)$ is well-posed. □

5.4 The Feller property

The martingale problem formulation presents a comparatively weak relation between an operator and an associated process. But once well-posedness of the martingale problem for an operator $-p(x,D)$ is proved, we have much stronger results for the corresponding process. In particular for any point $x \in \mathbb{R}^n$ there is a unique solution $P^x \in \mathcal{M}_1(D_{\mathbb{R}^n})$ of the martingale problem starting at the point x, i.e. having the initial distribution ε_x. It is a general result that this family $(P^x)_{x\in\mathbb{R}^n}$ depends on x in a measurable way (see [17], 4.4.6).

Moreover in this case the family of processes $((X_t)_{t\geq 0}, (P^x)_{x\in\mathbb{R}^n})$ even satisfies the strong Markov property:

Theorem 5.17. *Let p be a continuous negative definite symbol and let the martingale problem for $-p(x,D)$ be well-posed. Then $((X_t)_{t\geq 0}, (P^x)_{x\in\mathbb{R}^n})$ defined as above is a strong Markov family with respect to the filtration $(\mathcal{F}_t)$, i.e. for every $(\mathcal{F}_{t+})$-stopping time τ and all $t \geq 0$, $f \in B(\mathbb{R}^n)$ we have*

$$\mathbb{E}^x \left[f(X_{t+\tau}) | \mathcal{F}_{\tau+} \right] = \mathbb{E}^{X_\tau} \left[f(X_t) \right] \qquad \text{on } \{\tau < \infty\} \ P^x\text{-a.s.},$$

where $\mathbb{E}^x$ denote the expectation with respect to P^x.

This theorem can be proved as in [41],p.205, Theorem 5.1, see also [17], 4.4.2. Thus in the case of well-posedness we have a complete characterazation of a Markov process in terms of it generator.

But furthermore in [91] J. van Casteren as shown that the well- posedness of the martingale problem is almost equivalent to the statement that the operator extends to the generator of a Feller semigroup. See also [71] concerning this subject. In particular Proposition 2.6 in [91] implies

Proposition 5.18. *Assume that $p(x,D)$ maps $C_0^\infty(\mathbb{R}^n)$ into $C_\infty(\mathbb{R}^n)$. If the martingale problem for $-p(x,D)$ is well-posed, then $-p(x,D)$ has an extension which is the generator of a Feller semigroup.*

In general the range of a pseudo differential operator $p(x,D)$ as considered in the previous chapters is not contained in $C_\infty(\mathbb{R}^n)$ due to the non-local behaviour. Therefore they can't be extended to the generator of a Feller semigroup, which by definition is a semigroup in $C_\infty(\mathbb{R}^n)$, and the result is not applicable directly.

Nevertheless using the tightness results for solutions of the martingale problem as developed in Chapter 3, we can prove under an additional weak assumption that the solution P^x depend on x in a continuous way. Using this result we will show that the associated process actually defines a Feller semigroup.

Let $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ be a continuous negative definite symbol such that $p(x,0) = 0$ for all $x \in \mathbb{R}^n$ and

$$|p(x,\xi)| \leq c(1 + |\xi|^2) \tag{5.46}$$

and hence the existence result Theorem 3.15 applies. The additional assumption on the symbol will be the equicontinuity at $\xi = 0$, i.e.

$$\sup_{x\in\mathbb{R}^n} |p(x,\xi)| \underset{\xi\to 0}{\longrightarrow} 0. \tag{5.47}$$

We need the following auxiliary result.

Lemma 5.19. *Let $p(x,\xi)$ be as above, in particular assume (5.46) and (5.47). Then for any $\varphi \in C_0^\infty(\mathbb{R}^n)$ we have with $\varphi_R(x) = \varphi(\frac{x}{R})$, $R > 0$*

$$\lim_{R\to\infty} \sup_{x\in\mathbb{R}^n} |p(x,D)\varphi_R(x)| = 0 \tag{5.48}$$

Proof: It is enough to repeat the argument in (3.30) in the proof of Lemma 3.16 (replace k by R) and to note that by assumption $\sup_{|\xi|\leq\frac{1}{\sqrt{R}}}|p(x,\xi)|$ tends to 0 uniformly with respect to x as $R\to\infty$ □

We now can prove

Theorem 5.20. *Assume that $p(x,\xi)$ is a continuous negative definite symbol that satisfies (5.46), (5.47) and let $(\mu_k)_{k\in\mathbb{N}}$ be a tight set of probability measures in $\mathbb{R}^n$. Let $P_k\in\mathcal{M}_1(D_{\mathbb{R}^n})$ be a solution of the martingale problem for $-p(x,D)$ and the initial distribution μ_k, then the set $\{P_k\}_{k\in\mathbb{N}}$ is tight in $\mathcal{M}_1(D_{\mathbb{R}^n})$.*

Proof: Since $(-p(x,D),C_0^\infty(\mathbb{R}^n))$ is a linear operator in $C_b(\mathbb{R}^n)$, the result follows from Theorem 3.10 (with a family (A_α) chosen such that all A_α are equal to $-p(x,D)$) provided the compact containment condition holds for $(P_k)_{k\in\mathbb{N}}$. This means for all $T>0$ and all $\varepsilon>0$ there is a compact set $K\subset\mathbb{R}^n$ such that

$$\sup_{k\in\mathbb{N}}P_k(X_t\notin K\text{ for some }0\leq t\leq T)\leq\varepsilon.$$

Choose $\varphi\in C_0^\infty(\mathbb{R}^n)$ such that $0\leq\varphi\leq1$, $\varphi(x)=1$ for $|x|\leq\frac{1}{2}$, $\operatorname{supp}\varphi\subset B_1(0)$ and let $\varphi_R(x)=\varphi(\frac{x}{R})$ for $R>0$.
Given $\varepsilon,T>0$ the tightness of $\{\mu_k\}_{k\in\mathbb{N}}$ implies

$$\sup_{k\in\mathbb{N}}\mu_k\left(|x|\geq\frac{R}{2}\right)\leq\frac{\varepsilon}{2}$$

for R sufficiently large. We define the $\mathcal{F}_{t+}$-stopping time

$$\tau=\inf\{t\geq0:|X_t|>R\}.$$

Then optional sampling gives for the right-continuous martingale

$$\varphi_R(X_t)-\int_0^t(-p(x,D)\varphi_R)(X_u)\,du$$

that

$$\begin{aligned}&\mathbb{E}^{P_k}\left[1-\varphi_R(X_{\tau\wedge T})+\int_0^{\tau\wedge T}(-p(x,D)\varphi_R)(X_u)\,du\right]\\&=\ \mathbb{E}^{P_k}\left[1-\varphi_R(X_0)\right]=\int_{\mathbb{R}^n}(1-\varphi_R)\,d\mu_k\leq\frac{\varepsilon}{2}.\end{aligned}$$

Moreover, again for R sufficiently large we have by Lemma 5.19

$$\sup_{x\in\mathbb{R}^n}|p(x,D)\varphi_R(x)|\leq\frac{\varepsilon}{2T}$$

and therefore

$$\begin{aligned}P_k\left(\sup_{0\leq t\leq T}|X_t|>R\right)&\leq P_k\left(|X_{\tau\wedge T}|\geq R\right)\leq\mathbb{E}^{P_k}\left[1-\varphi_R(X_{\tau\wedge T})\right]\\&\leq\ \frac{\varepsilon}{2}+\mathbb{E}^{P_k}\left[\int_0^{\tau\wedge T}|(p(x,D)\varphi_R)(X_u)|\,du\right]\leq\frac{\varepsilon}{2}+T\frac{\varepsilon}{2T}=\varepsilon,\end{aligned}$$

which proves the compact containment condition. □

Theorem 5.20 now immediately implies

Corollary 5.21. *Let $p(x,\xi)$ be a continuous negative definite symbol such that (5.46) and (5.47) hold. If the martingale problem for $-p(x,D)$ is well-posed and P^x denotes the solution starting at $x \in \mathbb{R}^n$, the the map $x \mapsto P^x$ is a continuous map on $\mathbb{R}^n$ into $\mathcal{M}_1(\mathbb{R}^n)$ equipped with the weak toplogy.*

Proof: Let $(x_k)_{k\in\mathbb{N}}$ be a sequence of points in $\mathbb{R}^n$ which converges to $x \in \mathbb{R}^n$. Then the initial distributions $\mu_k = \varepsilon_{x_k}$ converge weakly to ε_x and form a tight subset. Thus, the corresponding solutions P^{x_n} also form a tight set and have at least one accumulation point $P \in \mathcal{M}_1(D_{\mathbb{R}^n})$. Since $X_0 : D_{\mathbb{R}^n} \to \mathbb{R}^n$ is continuous, this implies that the one dimensional distributions $P^{x_k} \circ X_0^{-1} = \varepsilon_{x_k}$ have $P \circ X_0^{-1}$ as a cluster point for the weak topology, hence $P \circ X_0^{-1} = \varepsilon_x$. Moreover, a look to the argument in proof of Proposition 3.14 (choose A_θ, A_k all equal $-p(x,D)$) shows that P is also a solution of the martingale problem for $-p(x,D)$. Thus well-posedness implies that the cluster point $P = P^x$ is uniquely determined and hence P^{x_k} converges to P^x. □

Since in the case of well-posedness the family $(P^x)_{x\in\mathbb{R}^n}$ of solutions of the martingale problem defines a Markov process, we obtain a Markovian semigroup $(P_t)_{t\geq 0}$ acting on the bounded Borel measurable functions by

$$P_t f(x) = \mathbb{E}^x\left[f(X_t)\right].$$

If we would know that the distributions of X_t, $t \geq 0$ depend on the starting point x in a continuous way like the measures P^x do , then P_t even maps $C_b(\mathbb{R}^n)$ into itself, because in this case

$$P_t f(x) = \mathbb{E}^{P^x \circ X_t^{-1}}\left[f\right] \to \mathbb{E}^{P^{x_0} \circ X_t^{-1}}\left[f\right] = P_t f(x_0) \text{ as } x \to x_0 \in \mathbb{R}^n.$$

But the random variables $X_t : D_{\mathbb{R}^n} \to \mathbb{R}^n$ are not continuous for $t > 0$ as it is the case for the continuous path space. Hence we cannot immediately conclude that the continuous dependence of P^x on x implies the same for the one-dimensional distributions. Nevertheless $P^x \to P^{x_0}$ weakly implies $P^x \circ {X_t}^{-1} \to P^{x_0} \circ {X_t}^{-1}$ weakly, if we know that X_t is continuous at least at P^{x_0}-almost all points $\omega \in D_{\mathbb{R}^n}$ (see [5], p.30, Theorem 5.1). But the continuity points ω of X_t are given by the points which satisfy $X_t(\omega) = X_{t-}(\omega)$. Therefore we need the following

Proposition 5.22. *Let $p(x,\xi)$ be as in Corollary 5.21 and let P be a solution of the martingale problem for $-p(x,\xi)$. Then the process $(X_t)_{t\geq 0}$ has no fixed times of discontinuity, i.e.*

$$P(X_t \neq X_{t-}) = 0$$

for all $x \in \mathbb{R}^n$, $t > 0$.

Proof: Choose a sequence $(\varphi_k)_{k\in\mathbb{N}}$ in $C_0^\infty(\mathbb{R}^n)$, $0 \leq \varphi_k \leq 1$, that separates the points of $\mathbb{R}^n$ in such a way that for any $x, y \in \mathbb{R}^n$, $x \neq y$, there is a φ_k such that $\varphi_k(x) = 1$ and $\varphi_k(y)$=0.

Let $0 \leq s_0 \leq s < t$. Since P solves the martingale problem, we have for all $A \in \mathcal{F}_{s_0}$ and all $k \in \mathbb{N}$

$$\begin{aligned} \left| \int_A [\varphi_k(X_t) - \varphi_k(X_s)] \, dP \right| &= \left| \int_A \int_s^t (p(x,D)\varphi_k)(X_u) \, du \, dP \right| \\ &\leq |t-s| \sup_{\xi \in \mathbb{R}^n} |p(x,D)\varphi_k(x)| \end{aligned}$$

and for $s \uparrow t$ we find

$$\int_A [\varphi_k(X_t) - \varphi_k(X_{t-})] \, dP = 0 \tag{5.49}$$

for all $A \in \mathcal{F}_{s_0}$ and, since $s_0 < t$ is arbitrary, also for all $A \in \mathcal{F}_{t-} = \sigma\{X_s : s < t\}$. In particular, for $A = \{\varphi_k(X_{t-}) = 0\}$ equation (5.49) yields for all $k \in \mathbb{N}$

$$\begin{aligned} P\left(\varphi_k(X_t) = 1, \varphi_k(X_{t-}) = 0\right) &= \int_{\{\varphi_k(X_{t-})=0\}} 1_{\{\varphi_k(X_t)=1\}} \, dP \\ \leq \int_{\{\varphi_k(X_{t-})=0\}} \varphi_k(X_t) \, dP &= \int_{\{\varphi_k(X_{t-})=0\}} \varphi_k(X_{t-}) \, dP = 0, \end{aligned}$$

which finally gives

$$P(X_t \neq X_{t-}) = P\left(\bigcup_{k \in \mathbb{N}} \{\varphi_k(X_t) = 1, \varphi_k(X_{t-}) = 0\} \right) = 0.$$

□

To prove that $(P_t)_{t \geq 0}$ is a Feller semigroup it remains to show that P_t leaves $C_\infty(\mathbb{R}^n)$ invariant.

Theorem 5.23. *Assume that $p(x,\xi)$ is a continuous negative definite symbol such that (5.46) and (5.47) hold. If the martingale problem for $-p(x,D)$ is well-posed then $(P_t)_{t\geq 0}$ defines a Feller semigroup, whose generator is an extension of $-p(x,D)$. In particular, $p(x,D)$ maps $C_0^\infty(\mathbb{R}^n)$ into $C_\infty(\mathbb{R}^n)$.*

Proof: By Proposition 5.22 and the remarks preceding it we know that $(P_t)_{t\geq 0}$ is a semigroup of Markovian operators on $C_b(\mathbb{R}^n)$. Let $\varphi \in C_0^\infty(\mathbb{R}^n)$, $0 \leq \varphi \leq 1$, $\varphi(0) = 1$, $\operatorname{supp} \varphi \subset B_1(0)$ and put $\varphi_{x_0,R}(x) = \varphi\left(\frac{x-x_0}{R}\right)$, $x_0 \in \mathbb{R}^n$, $R > 0$. Then by Lemma 5.19

$$\sup_{x \in \mathbb{R}^n} |p(x,D)\varphi_{x_0,R}(x)| \to 0 \text{ as } R \to \infty$$

and a look to the proof immediately shows that this limit holds uniformly with respect to x_0, i.e. $\lim_{R\to\infty} \varrho(R) = 0$, where

$$\varrho(R) = \sup_{x_0 \in \mathbb{R}^n} \sup_{x \in \mathbb{R}^n} |p(x,D)\varphi_{x_0,R}(x)| \, .$$

Thus, for any $\psi \in C_0^\infty(\mathbb{R}^n)$ and $x_0 \notin \operatorname{supp}\psi$ we have with $R = \operatorname{dist}(x_0, \operatorname{supp}\psi) > 0$

$$\begin{aligned} |P_t\psi(x_0)| = |\mathbb{E}^{x_0}[\psi(X_t)]| &\leq \|\psi\|_\infty P^{x_0}(X_t \in \operatorname{supp}\psi) \leq \|\psi\|_\infty \mathbb{E}^{x_0}[1 - \varphi_{x_0,R}(X_t)] \\ &= \|\psi\|_\infty \mathbb{E}^{x_0}[\varphi_{x_0,R}(X_0) - \varphi_{x_0,R}(X_t)] = \|\psi\|_\infty \mathbb{E}^{x_0}\left[\int_0^t (p(x,D)\varphi_{x_0,R})(X_u) \, du \right] \\ &\leq \|\psi\|_\infty \cdot t \cdot \varrho(R) \to 0 \text{ as } |x_0| \to \infty. \end{aligned}$$

Therefore, since $C_0^\infty(\mathbb{R}^n)$ is dense in $C_\infty(\mathbb{R}^n)$, P_t maps $C_\infty(\mathbb{R}^n)$ into itself. Moreover, for fixed $x \in \mathbb{R}^n$ we find for $f \in C_\infty(\mathbb{R}^n)$ by Lebesgue's theorem

$$\lim_{t\to 0} P_t f(x) = \lim_{t\to 0} \mathbb{E}^x\left[f(X_t)\right] = \mathbb{E}^x\left[f(X_0)\right] = f(x).$$

Recall that bounded pointwise convergence implies weak convergence in $C_\infty(\mathbb{R}^n)$, since the dual space of $C_\infty(\mathbb{R}^n)$ consists of signed measures of bounded variation. Therefore $(P_t)_{t\geq 0}$ is a weakly and by [96], p.233, even a strongly continuous semigroup in $C_\infty(\mathbb{R}^n)$, i.e. a Feller semigroup. For a testfunction $\varphi \in C_0^\infty(\mathbb{R}^n)$ the generator is given by

$$\begin{aligned}
&\lim_{t\to 0} \frac{P_t f(x) - f(x)}{t} = \lim_{t\to 0} \mathbb{E}^x\left[\frac{1}{t}\left(\varphi(X_t) - \varphi(X_0)\right)\right] \\
&= \lim_{t\to 0} \mathbb{E}^x\left[\frac{1}{t}\int_0^t (-p(x,D)\varphi)(X_u)\,du\right] = \mathbb{E}^x\left[-p(x,D)\varphi(X_0)\right] = -p(x,D)\varphi(x),
\end{aligned}$$

where the pointwise convergence again implies weak convergence and therefore strong convergence in $C_\infty(\mathbb{R}^n)$ by [84], Theorem 3.1.2. □

We combine this result with the the well-posedness criterion of the last section. Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be the continuous negative definite reference function, which for some $r > 0$, $c > 0$ satisfies $\psi(\xi) \geq |\xi|^r$ for $|\xi| \geq 1$ and let $\lambda(\xi) = (1 + \psi(\xi))^{1/2}$ as usual. Let M be the smallest integer such that $M > \left(\frac{n}{r} \vee 2\right) + n$ and define $k = 2M + 1 - n$.

Theorem 5.24. *Let $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite symbol. Moreover assume that*

(i) the map $x \mapsto p(x,\xi)$ is k-times continuously differentiable and

$$\left|\partial_x^\beta p(x,\xi)\right| \leq c\lambda^2(\xi), \qquad \beta \in \mathbb{N}_0^n, |\beta| \leq k$$

holds,

(ii) for some strictly positive function $\gamma : \mathbb{R}^n \to \mathbb{R}^+$

$$p(x,\xi) \geq \gamma(x)\cdot\lambda^2(\xi) \qquad \text{for } |\xi| \geq 1, x \in \mathbb{R}^n$$

and

(iii)

$$\sup_{x\in\mathbb{R}^n} p(x,\xi) \underset{\xi\to 0}{\longrightarrow} 0.$$

Then $-p(x,D) : C_0^\infty(\mathbb{R}^n) \to C_\infty(\mathbb{R}^n)$ has an extension that generates a Feller semigroup given by

$$P_t f(x) = \mathbb{E}^x\left[f(X_t)\right].$$

Here $\mathbb{E}^x$ denotes the expectation taken with respect to the solution of the associated well-posed martingale problem starting at x.

Chapter 6

A symbolic calculus

6.1 General remarks

Up to now we used the notion of a pseudo differential operator simply for an operator which has a representation

$$p(x,D)\varphi(x) = \int_{\mathbb{R}^n} e^{i(x,\xi)} p(x,\xi)\hat{\varphi}(\xi)\, đ\xi$$

and we assumed that the symbol is negative definite that relates this type of operator to generators of Markov processes and Feller semigroups. In a narrower sense pseudo differential operators are classes of operators which can be handled by a symbolic calculus. This means that operations on the level of the operators have an equivalent on the level of the symbols. For example the composition of two operators corresponds to the product of their symbols plus terms of lower order.
In this context an important notion is the asymptotic expansion of a symbol into a series of symbols of decreasing order (Here the order of a symbol has to be specified either by the mapping properties of the corresponding pseudo differential operator or by growth properties of the symbol $p(x,\xi)$ with respect to ξ). This series does not converge in the usual sense, but adding the first terms of such expansion describes the given symbol up to a remainder part, whose order decreases when more and more terms of the expansion are taken into account.
The terms of this expansion, for example for the composition of two operators, are computable in a more or less algebraic way. Therefore a symbolic calculus for pseudo differential operators is often useful to justify intuitive ideas in a well founded framework. This explains the importance of pseudo differential operators as an auxiliary tool in the theory of partial differential equations. As standard references in this field we refer to the books of Hörmander [28], Kumano-go [54] and Taylor [85].

The standard class of symbols which are considered in this context is the Hörmander class $S^m_{\varrho,\delta}$ of all C^∞ functions $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ such that

$$\left|\partial_\xi^\alpha \partial_x^\beta p(x,\xi)\right| \leq c_{\alpha,\beta}\langle\xi\rangle^{m-\varrho|\alpha|+\delta|\beta|} \qquad \text{for all } \alpha,\, \beta \in \mathbb{N}_0^n. \tag{6.1}$$

Here $m \in \mathbb{R}$ is the order of the symbol and ϱ, δ are parameters that satisfy $0 \leq \delta \leq \varrho \leq 1$. In the classical case $\varrho = 1$, $\delta = 0$ this class generalizes the behaviour of the symbols of linear partial differential operators, which are polynomial with respect to ξ.

For us it is important to note that in the standard case $\varrho > 0$ the power $m - \varrho|\alpha| + \delta|\beta|$ in (6.1) decreases to $-\infty$ as $|\alpha| \to \infty$. Since terms of the asymptotic expansion are determined by derivatives of the symbol, it is this property of the symbol, which permits a reasonable asymptotic expansion with remainder terms of arbitrary low order.

Unfortunately, continuous negative definite symbols do not fit into this framework. First of all they are in general not even differentiable with respect to ξ. But also in the case of differentiable negative definite functions the derivatives do not have a behaviour as it is needed for Hörmander type symbols. Therefore the classical theory of pseudo differential operators is not applicable and this was the reason to develop new techniques for pseudo differential operators with continuous negative definite symbols as presented in the previous chapters.

The direct construction of the Feller semigroup by solving the corresponding equation and using the Hille-Yosida theorem in Chapter 4 yields also L^2-estimates for the generator. These estimates provide important informations for the semigroup and the corresponding process (see Chapter 8), but in order to make the approach work we had to restrict ourselves to small perturbations of the x- independent case. On the other hand by the martingale problem we can treat a reasonable class of continuous negative definite symbols which are defined in terms of a continuous negative definite reference function without smallness assumptions. But by the localization technique we loose the L^2-type estimates in the anisotropic Sobolev spaces. It is therefore desirable to possess a symbolic calculus for this kind of pseudo differential operators, since such calculus provides us with good L^2-estimates for the operators, but will not be restricted to the case of small perturbations.

The main idea to overcome the lack of differentiability of continuous negative definite symbols with respect to ξ is to restrict to the case, where the Lévy measures of the continuous negative definite functions have a bounded support. In Proposition 3.11 and Theorem 3.12 we have seen that we can split the Lévy measures of a continuous negative definite symbol into a part supported in some bounded neighbourhood of the origin and in a remainder part. This remainder part consists of bounded Lévy measures and we can hope to treat this part as a perturbation in the spirit of Proposition 3.6.

Now the first part turns out to be differentiable with respect to ξ. But in view of the properties of symbols in Hörmander class it is also important to have estimates for the derivatives. Therefore let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite function with Lévy–Khinchin representation

$$\psi(\xi) = q(\xi) + c + \int_{\mathbb{R}^n\setminus\{0\}} (1 - \cos(y, \xi))\mu(dy), \tag{6.2}$$

where $q \geq 0$ is the quadratic form, $c \geq 0$ is a constant and μ is a symmetric Lévy measure, see Corollary 2.14. Then we have the following theorem.

Theorem 6.1. *Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite function with Lévy-Khinchin representation (6.2). If for some ball $B_R(0)$, $R > 0$ the Lévy measure satisfies*

$$supp\, \mu \subset B_R(0),$$

then ψ is infinitely often differentiable and we have the estimates

$$\left|\partial_\xi^\alpha \psi(\xi)\right| \leq c_{|\alpha|} \begin{cases} \psi(\xi) & \alpha = 0 \\ \psi^{1/2}(\xi) \quad \textit{if} & |\alpha| = 1 \\ 1 & |\alpha| \geq 2 \end{cases}, \alpha \in \mathbb{N}_0^n. \tag{6.3}$$

The constants $c_{|\alpha|}$ depend only on $|\alpha|$, n, R and the constant c_ψ in the upper bound

$$\psi(\xi) \le c_\psi(1 + |\xi|^2).$$

Proof: Define the absolute moments

$$M_l = \int_{\mathbb{R}^n \setminus \{0\}} |y|^l \ \mu(dy)$$

and let Λ be the maximal eigenvalue of the quadratic form q. Then by the assumption M_l is finite for all $l \ge 2$.
For $\alpha = 0$ there is nothing to prove. Let $|\alpha| \ge 1$. We may consider all terms in the representation (6.2) of ψ separately. The constant term is trivial and the estimate (6.3) is well-known for the quadratic form with constants $c_1 = 2\Lambda^{1/2}$, $c_2 = 2\Lambda$ and $c_l = 0$ for $l > 2$. So we may restrict to the integral part in (6.2) and assume that

$$\psi(\xi) = \int_{\mathbb{R}^n \setminus \{0\}} (1 - \cos(y, \xi))\mu(dy).$$

Since the moments M_l, $l \ge 2$ are finite, we may interchange differentiation and integration and find

$$\partial_\xi^\alpha \psi(\xi) = -\int_{\mathbb{R}^n \setminus \{0\}} y^\alpha \cos^{(|\alpha|)}(y, \xi)\mu(dy),$$

which gives for $|\alpha| = 1$ by Cauchy-Schwarz inequality

$$\begin{aligned}
\left|\partial_{\xi_i}\psi(\xi)\right| &\le \left(\int_{\mathbb{R}^n \setminus \{0\}} |y_i|^2 \mu(dy)\right)^{1/2} \cdot \left(\int_{\mathbb{R}^n \setminus \{0\}} \sin^2(y, \xi)\mu(dy)\right)^{1/2} \\
&\le \left(\int_{\mathbb{R}^n \setminus \{0\}} |y|^2 \mu(dy)\right)^{1/2} \cdot \left(2\int_{\mathbb{R}^n \setminus \{0\}} (1 - \cos(y, \xi))\mu(dy)\right)^{1/2} \\
&= (2M_2)^{1/2} \cdot \psi^{1/2}(\xi)
\end{aligned}$$

and for $|\alpha| \ge 2$

$$\begin{aligned}
\left|\partial_\xi^\alpha \psi(\xi)\right| &\le \int_{\mathbb{R}^n \setminus \{0\}} |y^\alpha| \left|\cos^{(|\alpha|)}(\xi, y)\right| \mu(dy) \\
&\le \int_{\mathbb{R}^n} |y|^{|\alpha|} \mu(dy) = M_{|\alpha|}.
\end{aligned}$$

This proves (6.3) for $c_0 = 1$, $c_1 = (2M_2)^{1/2} + 2\Lambda^{1/2}$, $c_2 = M_2 + 2\Lambda$ and $c_l = M_l$ for $l > 2$.
To complete the proof it is sufficient to remark that

$$\Lambda = \sup_{|\xi| \le 1} q(\xi) \le \sup_{|\xi| \le 1} \psi(\xi) \le c_\psi \sup_{|\xi| \le 1} (1 + |\xi|^2) = 2c_\psi$$

and by Lemma 2.15 for $l \geq 2$

$$\begin{aligned} M_l &= \int_{B_R(0)\setminus\{0\}} |y|^l \, \mu(dy) \leq c_{R,l} \int_{B_R(0)\setminus\{0\}} \frac{|y|^2}{(1+|y|^2)} \mu(dy) \\ &= c_{R,l} \int_{\mathbb{R}^n\setminus\{0\}} \int_{\mathbb{R}^n\setminus\{0\}} (1-\cos(y,\xi)) \, \mu(dy) \, \nu(d\xi) \leq c_{R,l} \int_{\mathbb{R}^n\setminus\{0\}} \psi(\xi) \, \nu(d\xi) \\ &\leq c_{R,l} \cdot c_\psi \int_{\mathbb{R}^n} (1+|\xi|^2) \, \nu(d\xi). \end{aligned}$$

□

6.2 The symbol classes $\mathbf{S}_\varrho^{m,\lambda}$ and $\mathbf{S}_0^{m,\lambda}$

As before we again fix a continuous negative definite reference function

$$\psi : \mathbb{R}^n \to \mathbb{R}.$$

We assume that ψ has a Lévy measure which has support in a bounded set. This is no restriction, because we are only interested in the growth behaviour of ψ for $|\xi| \to \infty$ and this will not change if we cut off the Lévy measure outside some neighbourhood of the origin. As usual we assume that

$$\psi(\xi) \geq c\,|\xi|^r \qquad \text{for all } |\xi| \geq 1 \tag{6.4}$$

for some $r > 0$ and $c > 0$ and we use the notation

$$\lambda(\xi) = (1+\psi(\xi))^{1/2}. \tag{6.5}$$

Consider a real-valued continuous negative definite symbol $\tilde{p} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ that can be estimated by the reference function:

$$\tilde{p}(x,\xi) \leq c\lambda^2(\xi). \tag{6.6}$$

Then the Lévy-Khinchin formula yields

$$\tilde{p}(x,\xi) = q(x,\xi) + c(x) + \int_{\mathbb{R}^n\setminus\{0\}} (1-\cos(y,\xi))\mu(x,dy),$$

where c, q, and μ satisfy for each $x \in \mathbb{R}^n$ the same conditions as the corresponding terms in (6.2). We now decompose $\tilde{p}$ as described in the previous section. For that purpose let $\theta \in C_0^\infty(\mathbb{R}^n)$, $0 \leq \theta \leq 1$, be a some even cut-off function such that $\theta(x) = 1$ in a neighbourhood of the origin. Then we obtain the decomposition

$$\tilde{p}(x,\xi) = p(x,\xi) + p_r(x,\xi) \tag{6.7}$$

by splitting the Lévy-measures into a leading term

$$p(x,\xi) = q(x,\xi) + c(x) + \int_{\mathbb{R}^n\setminus\{0\}} (1-\cos(y,\xi))\,\theta(y)\,\mu(x,dy)$$

and a remainder term

$$p_r(x,\xi) = \int_{\mathbb{R}^n\setminus\{0\}} (1-\cos(y,\xi))\,(1-\theta(y))\,\mu(x,dy)$$

By Proposition 3.11 and Theorem 3.12 p, $p_r : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ are continuous negative definite symbols and the Lévy-type representation of $p_r(x,D)$ is given by

$$p_r(x,D)\varphi(x) = -\int_{\mathbb{R}^n\setminus\{0\}} (\varphi(x+y)-\varphi(x))\,(1-\theta(y))\,\mu(x,dy).$$

Moreover, Theorem 3.12 implies because of (6.6) that the Lévy-kernel $(1-\theta(y))\,\mu(x,dy)$ of $p_r(x,D)$ consists of finite measures with uniformly bounded mass. This shows that $p_r(x,D)$ is bounded as an operator on the space of bounded Borel measurable functions. Moreover, under a mild additional conditions $C_\infty(\mathbb{R}^n)$ is invariant and in typical examples the operator is bounded on $L^2(\mathbb{R}^n)$, see Section 6.6. Therefore we regard $p_r(x,D)$ as a perturbation of $p(x,D)$ and we will look in following to the part $p(x,\xi)$ which contains the typically dominating part of the Lévy-measure concentrated around the origin.

To simplify the notation we introduce with regard to (6.3)

$$\varrho(k) = k \wedge 2, \qquad k \in \mathbb{N}_0. \tag{6.8}$$

Since all Lévy-measures of the continuous negative definite symbol $p(x,\xi)$ are supported in $\operatorname{supp}\theta$, Theorem 6.1 applies to $p(x,\xi)$ and we have

$$\begin{aligned} \left|\partial_\xi^\alpha p(x,\xi)\right| &\le c_\alpha p(x,\xi)^{\frac{1}{2}(2-\varrho(|\alpha|))} \\ &= c_\alpha \lambda(\xi)^{(2-\varrho(|\alpha|))} \end{aligned} \tag{6.9}$$

with a constant c_α not depending on x. The estimate (6.9) reflects the typical behaviour of negative definite symbols and in order to define a proper symbol class it is quite natural to assume the same estimate for the derivates $\partial_x^\beta p(x,\xi)$ of the symbol. Then real-valued continuous negative definite symbols are symbols of order 2 in the sense of the following definition.

Definition 6.2. *Let $\lambda(\xi)$ and ϱ be defined as in (6.5) and (6.8) and let $m \in \mathbb{R}$. A C^∞-function $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ is called a symbol in class $S_\varrho^{m,\lambda}$ if for all α, $\beta \in \mathbb{N}_0^n$ there are constants $c_{\alpha,\beta} \ge 0$ such that*

$$\left|\partial_\xi^\alpha \partial_x^\beta p(x,\xi)\right| \le c_{\alpha,\beta}\lambda(\xi)^{m-\varrho(|\alpha|)}, \qquad x \in \mathbb{R}^n,\ \xi \in \mathbb{R}^n. \tag{6.10}$$

The number $m \in \mathbb{R}$ is called the order of the symbol.

Let us also define the following enlarged class of symbols.

Definition 6.3. *A C^∞-function $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ is called a symbol in class $S_0^{m,\lambda}$ if for all α, $\beta \in \mathbb{N}_0^n$ there are constants $c_{\alpha,\beta} \ge 0$ such that*

$$\left|\partial_\xi^\alpha \partial_x^\beta p(x,\xi)\right| \le c_{\alpha,\beta}\lambda(\xi)^{m}, \qquad x \in \mathbb{R}^n,\ \xi \in \mathbb{R}^n. \tag{6.11}$$

In the definition of $S_\varrho^{m,\lambda}$ and $S_0^{m,\lambda}$ we have replaced the weight function $\langle\xi\rangle$ of Hörmander type symbols by the more general function $\lambda(\xi)$. Symbol classes defined by general so-called basic weight functions $\lambda(\xi)$ had been considered before by H. Kumano-go (see [54]), but his assumptions on λ are not satisfied by continuous negative definite functions in general. Therefore the major part of the work that has to be done is to show that arguments similar to those in [54] can be applied in the situation here. For that purpose we have to exploit again estimates for continuous negative definite functions that replace estimates for the basic weight functions used in [54]. See also the paper [67] of M. Nagase where he also considers basic weight functions as in [54]. In his paper Nagase also lines out how the technique of Friedrichs symmetrization applies to his class of symbols. We adapt this procedure to our situation proving also in our case a Friedrichs symmetrization and a sharp Garding inequality.

Let us consider some typical situations, where continuous negative definite symbols in class $S_\varrho^{2,\lambda}$ appear.

Example 1: Let $\tilde{\mu}$ be a symmetric Lévy-measure on $\mathbb{R}^n \setminus \{0\}$ and define a continuous negative definite reference function

$$\psi(\xi) = \int_{0<|y|\le 1} (1-\cos(y,\xi))\,\tilde{\mu}(dy)$$

and let $\lambda(\xi)$ as in (6.5). Suppose that $\mu(x,dy)$ is a Lévy-kernel with Lévy-measures absolutely continuous with respect to $\tilde{\mu}$, i.e. there is a density f defined on $\mathbb{R}^n \times \mathbb{R}^n \setminus \{0\}$ such that

$$\mu(x,dy) = f(x,y)\,\tilde{\mu}(dy).$$

Assume that $f(x,y)$ has bounded derivatives with respect to x of all orders. Then the symbol

$$p(x,\xi) = \int_{\mathbb{R}^n\setminus\{0\}} (1-\cos(y,\xi))\,\mu(x,dy)$$

has a decomposition $p(x,\xi) = p_1(x,\xi) + p_2(x,\xi)$, where

$$p_1(x,\xi) = \int_{0<|y|\le 1} (1-\cos(y,\xi))\,\mu(x,dy)$$

and

$$p_2(x,\xi) = \int_{|y|>1} (1-\cos(y,\xi))\,\mu(x,dy)$$

Then we have $p_1 \in S_\varrho^{2,\lambda}$. In fact

$$\begin{aligned}
\left|\partial_x^\beta p_1(x,\xi)\right| &\le \left|\int_{0<|y|\le 1} (1-\cos(y,\xi))\partial_x^\beta f(x,y)\,\tilde{\mu}(dy)\right| \\
&\le \sup\left|\partial_x^\beta f\right| \cdot \int_{0<|y|\le 1} (1-\cos(y,\xi))\,\tilde{\mu}(dy) \\
&\le c_\beta\,\lambda^2(\xi)
\end{aligned}$$

and for $|\alpha| \geq 1$

$$\begin{aligned}
\left|\partial_\xi^\alpha \partial_x^\beta p_1(x,\xi)\right| &\leq \left|\partial_\xi^\alpha \partial_x^\beta \int_{0<|y|\leq 1} (1-\cos(y,\xi)) f(x,y)\, \tilde{\mu}(dy)\right| \\
&\leq \sup\left|\partial_x^\beta f\right| \cdot \int_{0<|y|\leq 1} \left|y^\alpha \cdot \cos^{(\alpha)}(y,\xi)\right| \tilde{\mu}(dy) \\
&\leq c_\beta \cdot c_a\, \lambda^{2-\varrho(|\alpha|)},
\end{aligned}$$

where in the last step we used the argument of the proof of Theorem 6.1.
Moreover, we will see in Section 6.6 that $p_2(x, D)$ is a bounded operator in $C_\infty(\mathbb{R}^n)$ as well as in $L^2(\mathbb{R}^n)$.

Example 2: In particular, continuous negative definite symbols of the following sum structure are covered:

$$p(x,\xi) = \sum_{j=1}^N b_j(x)\, \psi_j(\xi)$$

for some $N \in \mathbb{N}$. Symbols of this type are considered in [42], [29], [30], [39], [31] and [38], where associated Feller semigroups, Dirichlet forms and solutions to the martingale problem are constructed. We can regard these examples as special cases of Example 1. Thus if $\psi_j : \mathbb{R}^n \to \mathbb{R}$ are continuous negative definite functions having Lévy-measures with bounded support and if $b_j : \mathbb{R}^n \to \mathbb{R}^+$ are C^∞-functions with bounded derivatives, then $p \in \mathrm{S}_\varrho^{2,\lambda}$ for $\lambda(\xi) = (1 + \sum_{j=1}^N \psi_j(\xi))^{1/2}$.

Many more examples can be obtained by subordination of the symbol using Bernstein functions. A Bernstein function is a C^∞-function

$$f : (0,\infty) \to \mathbb{R}^+,$$

which satisfies

$$(-1)^k \frac{d^k f}{ds^k} \leq 0 \qquad \text{for all } k \in \mathbb{N}.$$

They admit a unique representation

$$f(s) = a + bs + \int_0^\infty (1 - e^{-sr})\, \mu(dr), \tag{6.12}$$

where a, $b \geq 0$ are constants and μ is a measure on $(0,\infty)$ such that $\int_0^\infty \frac{r}{1+r}\, \mu(dr) < \infty$.
Typical examples of Bernstein functions are the fractional powers $s \mapsto s^\alpha$ for $0 < \alpha \leq 1$. Bernstein functions are considered in more detail in the monograph [4].
Bernstein functions have a unique continuous extension to the complex half-plane $\mathrm{Re}\, z \geq 0$ which is holomorphic in the open half-plane $\mathrm{Re}\, z > 0$. In particular the composition of a Bernstein function and a continuous negative definite function is well-defined. The important feature of Bernstein functions is that this composition is again a continuous negative definite function. It can be shown that Bernstein functions are the only functions with this property, see [23].

Note that there is a close connection to the subordination of stochastic processes (see [18]), that is a time change of a process by a subordinator, i.e. an process with independent and stationary increments and almost surely increasing paths on $\mathbb{R}^+$. A subordinator is described by a Bernstein function f in the sense that the Laplace transforms of the transition functions are given by e^{-tf}. In particular, for a continuous negative definite function ψ the Lévy-process with characteristic exponent ψ subordinated in this sense is again a Lévy-process with characteristic exponent $f \circ \psi$.

For the derivatives we have the following result, which in the case of so-called complete Bernstein functions is contained in [47], Lemma 2.10.

Proposition 6.4. *For the derivatives of a Bernstein function $f : (0, \infty) \to \mathbb{R}^+$ we have*

$$\left|f^{(k)}(s)\right| \leq \frac{k!}{s^k} f(s) \qquad \text{for all } s > 0. \tag{6.13}$$

Proof: From the estimate $1 + \frac{x^k}{k!} \leq e^x$, $x > 0$, we obtain

$$x^k e^{-x} \leq k!\,(1 - e^{-x}), \qquad x > 0.$$

We may assume that $a + bs = 0$ in the representation (6.12), since the estimate for these terms is trivial. Therefore (6.12) and interchanging differentiation and integration yield

$$\begin{aligned}
\left|f^{(k)}(s)\right| &= \left|\frac{d^k}{ds^k} \int_0^\infty (1 - e^{-rs})\,\mu(dr)\right| \\
&= \left|\int_0^\infty r^k \cdot e^{-rs}\,\mu(dr)\right| \\
&\leq \frac{k!}{s^k} \int_0^\infty (1 - e^{-rs})\,\mu(dr) \\
&= \frac{k!}{s^k} f(s).
\end{aligned}$$

□

We return to the symbol classes $\mathrm{S}_\varrho^{m,\lambda}$

Example 3: Let $f : (0, \infty) \to \mathbb{R}^+$ be a Bernstein function that satisfies $f(s) \leq cs^r$ for $s \geq 1$ and some $0 < r \leq 1$ (Note that this is always true for $r = 1$). Let $\lambda(\xi) = (1 + \psi(\xi))^{1/2}$ as above and let $p \in \mathrm{S}_\varrho^{m,\lambda}$, $m \geq 2$, be a real-valued elliptic symbol, i.e.

$$p(x, \xi) \geq c\,\lambda^m(\xi)$$

holds for some $c > 0$. Let $\lambda'(\xi) = (1 + \psi^r(\xi))^{1/2}$ the reference function obtain from the continuous negative definite function ψ^r.
Then $f \circ p \in \mathrm{S}_\varrho^{m,\lambda'}$. In particular for an elliptic continuous negative definite symbol $p \in \mathrm{S}_\varrho^{2,\lambda}$, also $f \circ p$ is a continuous negative definite symbol in $\mathrm{S}_\varrho^{2,\lambda'}$.

Proof: If g is a differentiable function on $\mathbb{R}^{2n}$ then for $\gamma \in \mathbb{N}_0^{2n}$ we find by induction on $|\gamma|$ using Leibniz rule

$$\partial^\gamma (f \circ g) = \sum_{j=1}^{|\gamma|} f^{(j)} \circ g \ \cdot \sum_{\substack{\gamma_1+\ldots,+\gamma_j=\gamma \\ \gamma_1,\ldots,\gamma_j \in \mathbb{N}_0^{2n}}} c(\gamma_1,\ldots,\gamma_j) \cdot \prod_{i=1}^{j} \partial^{\gamma_i} g.$$

For α, $\beta \in \mathbb{N}_0^n$ let $\gamma = (\alpha, \beta) \in \mathbb{N}_0^{2n}$ and $g = p(x,\xi)$. This gives by (6.13)

$$\begin{aligned}
\left|\partial_\xi^\alpha \partial_x^\beta (f \circ p)(x,\xi)\right| &\le c \sum_{j=1}^{|\gamma|} \left|f^{(j)}(p(x,\xi))\right| \cdot \sum_{\substack{\alpha_1+\ldots,+\alpha_j=\alpha \\ \beta_1+\ldots+\beta_j=\beta}} \cdot \prod_{i=1}^{j} \left|\partial_\xi^{\alpha_i} \partial_x^{\beta_i} p(x,\xi)\right| \\
&\le c \sum_{j=1}^{|\gamma|} j! \cdot f(p(x,\xi)) \cdot \sum_{\substack{\alpha_1+\ldots,+\alpha_j=\alpha \\ \beta_1+\ldots+\beta_j=\beta}} \cdot \prod_{i=1}^{j} \left|\frac{\partial_\xi^{\alpha_i} \partial_x^{\beta_i} p(x,\xi)}{p(x,\xi)}\right| \\
&\le c \cdot f(p(x,\xi)) \sum_{j=1}^{|\gamma|} \sum_{\substack{\alpha_1+\ldots,+\alpha_j=\alpha \\ \beta_1+\ldots+\beta_j=\beta}} \cdot \prod_{i=1}^{j} \lambda^{-\varrho(|\alpha_i|)}(\xi) \\
&\le c_{\alpha,\beta}\ p(x,\xi)^r \cdot \lambda^{-\varrho(|\alpha|)}(\xi)
\end{aligned}$$

by subadditivity of ϱ. Therefore

$$\left|\partial_\xi^\alpha \partial_x^\beta (f \circ p)(x,\xi)\right| \le c_{\alpha,\beta}\ \lambda^{rm}(\xi) \cdot \lambda^{-\varrho(|\alpha|)}(\xi) \le c_{\alpha,\beta} \lambda^{r(m-\varrho(|\alpha|))}(\xi) \le c_{\alpha,\beta} (\lambda')^{m-\varrho(|\alpha|)}(\xi)$$

□

6.3 A calculus for $\mathrm{S}_\varrho^{m,\lambda}$ and $\mathrm{S}_0^{m,\lambda}$

We now start with the investigation of the symbol classes $\mathrm{S}_\varrho^{m,\lambda}$ and $\mathrm{S}_0^{m,\lambda}$ and develop a symbolic calculus for the corresponding pseudo differential operators.
Since in the following the symbols are not assumed to be negative definite in general, it is reasonable to consider the case of complex-valued functions.

First we remark that $\lambda^m(\xi)$ gives a generic example of symbols in $\mathrm{S}_\varrho^{m,\lambda}$.

Lemma 6.5. *For $m \in \mathbb{R}$ and $\alpha \in \mathbb{N}_0^n$ we have*

$$\left|\partial_\xi^\alpha \lambda^m(\xi)\right| \le c_\alpha \lambda(\xi)^{m-\varrho(|\alpha|)}. \tag{6.14}$$

In particular $\lambda^m \in S_\varrho^{m,\lambda}$.

Proof: By Theorem 6.1 we know

$$\left|\partial_\xi^\alpha (1+\psi(\xi))\right| \le c_\alpha (1+\psi(\xi))^{1/2(2-\varrho(|\alpha|))}$$

and therefore

$$\left|\frac{\partial_\xi^\alpha(1+\psi(\xi))}{1+\psi(\xi)}\right| \leq c_\alpha(1+\psi(\xi))^{-\frac{1}{2}\varrho(|\alpha|)} = c_\alpha\lambda(\xi)^{-\varrho(|\alpha|)} \tag{6.15}$$

holds. Next note that by induction on $|\alpha|$ using Leibniz rule we have

$$\partial_\xi^\alpha\lambda^m(\xi) = \partial_\xi^\alpha[(1+\psi(\xi))^{m/2}] = (1+\psi(\xi))^{m/2}\sum_{\alpha_1+\ldots+\alpha_{|\alpha|}=\alpha} c(\alpha_1,\ldots,\alpha_{|\alpha|},m)\prod_{i=1}^{|\alpha|}\frac{\partial_\xi^{\alpha_i}(1+\psi(\xi))}{1+\psi(\xi)},$$

where $a_1,\ldots,\alpha_{|\alpha|}\in\mathbb{N}_0^n$, and therefore

$$\begin{aligned}\left|\partial_\xi^\alpha\lambda^m(\xi)\right| &\leq c_\alpha\lambda^m(\xi)\sum_{\alpha_1+\ldots+\alpha_{|\alpha|}=\alpha}\prod_{i=1}^{|\alpha|}\lambda^{-\varrho(|\alpha_i|)} \leq c_\alpha\lambda^m(\xi)\cdot\sum_{\alpha_1+\ldots+\alpha_{|\alpha|}=\alpha}\lambda(\xi)^{-\sum_{i=1}^{|\alpha|}\varrho(|\alpha_i|)}\\ &\leq c_\alpha\lambda^{m-\varrho(|\alpha|)}\end{aligned}$$

again by subadditivity of ϱ. □

Clearly for two symbols $p_i\in S_0^{m_i,\lambda}$, $i=1,2$, by Leibniz rule we have

$$\left|\partial_\xi^\alpha\partial_x^\beta(p_1\cdot p_2)(x,\xi)\right| \leq c\sum_{\substack{\alpha'+\alpha''=\alpha\\ \beta'+\beta''=\beta}}\left|\partial_\xi^{\alpha'}\partial_x^{\beta'}p_1(x,\xi)\right|\cdot\left|\partial_\xi^{\alpha''}\partial_x^{\beta''}p_2(x,\xi)\right| \leq c\lambda^{m_1+m_2}(\xi), \tag{6.16}$$

i.e. $p_1\cdot p_2\in S_0^{m_1+m_2,\lambda}$ and $(S_0^{m,\lambda})_{m\in\mathbb{R}}$ forms an algebra of symbols, which repects the order of the symbols.

Definition 6.6. *For symbols in $S_\varrho^{m,\lambda}$ and $S_0^{m,\lambda}$ we denote the corresponding classes of operators defined by*

$$p(x,D)\varphi(x) = \int_{\mathbb{R}^n} e^{i(x,\xi)}p(x,\xi)\cdot\hat{\varphi}(\xi)\,đ\xi$$

by $\Psi_\varrho^{m,\lambda}$ and $\Psi_0^{m,\lambda}$, respectively.

By Theorem 2.7 the operators in $\Psi_\varrho^{s,\lambda}$ and $\Psi_0^{s,\lambda}$ are well defined on $\mathcal{S}(\mathbb{R}^n,\mathbb{C})$ and moreover for $u\in\mathcal{S}(\mathbb{R}^n,\mathbb{C})$, $\alpha,\beta\in\mathbb{N}_0^n$ and $N>|\beta|+m+n$

$$\begin{aligned}\left|\partial_x^\beta\left(x^\alpha\int_{\mathbb{R}^n} e^{i(x,\xi)}p(x,\xi)\hat{u}(\xi)đ\xi\right)\right| &= \left|\partial_x^\beta\left(\int_{\mathbb{R}^n} e^{i(x,\xi)}D_\xi^\alpha(p(x,\xi)\hat{u}(\xi))đ\xi\right)\right|\\ &= \left|\int_{\mathbb{R}^n}\sum_{\substack{\alpha_1+\alpha_2=\alpha\\ \beta_1+\beta_2=\beta}}\binom{\alpha}{\alpha_1}\binom{\beta}{\beta_1}(i\xi)^{\beta_1}\cdot e^{i(x,\xi)}\partial_x^{\beta_2}D_\xi^{\alpha_1}p(x,\xi)D_\xi^{\alpha_2}\hat{u}(\xi)đ\xi\right|\\ &\leq c\int_{\mathbb{R}^n}\langle\xi\rangle^{|\beta|}\cdot\lambda^m(\xi)\cdot\sum_{|\gamma|\leq|\alpha|}\left|\partial_\xi^\gamma\hat{u}(\xi)\right|đ\xi\\ &\leq c\int_{\mathbb{R}^n}\langle\xi\rangle^{|\beta|+m-N}đ\xi\cdot\sup_{\xi\in\mathbb{R}^n}\left[\langle\xi\rangle^N\sum_{|\gamma|\leq|\alpha|}\left|\partial_\xi^\gamma\hat{u}(\xi)\right|\right].\end{aligned}$$

Since the Fourier transform is continuous on $\mathcal{S}(\mathbb{R}^n,\mathbb{C})$ this gives

Proposition 6.7. *An operator $p(x,D) \in \Psi_0^{m,\lambda}$ maps $\mathcal{S}(\mathbb{R}^n,\mathbb{C})$ continuously into itself.*

Let us recall the definition of **oscillatory integrals** (see [54], Chapt.1.6). A C^∞-function g on $\mathbb{R}^n \times \mathbb{R}^n$ is called of class $\mathcal{A}$ if the estimates

$$(6.17) \qquad \left|\partial_\eta^\alpha \partial_y^\beta g(\eta,y)\right| \le c_{\alpha\beta}\langle\eta\rangle^{m+\delta|\beta|}\langle y\rangle^\tau, \qquad \alpha,\beta \in \mathbb{N}_0^n,$$

hold for suitable $m \in \mathbb{R}$, $0 \le \delta < 1$ and $\tau \ge 0$. In this case the oscillatory integral is defined by

$$(6.18) \qquad O_S - \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(y,\eta)}\, g(\eta,y)\, dy đ\eta = \lim_{\varepsilon\to 0}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(y,\eta)}\,\chi(\varepsilon\eta,\varepsilon y)\, g(\eta,y)\, dy đ\eta.$$

where $\chi \in \mathcal{S}(\mathbb{R}^n \times \mathbb{R}^n)$ having the property $\chi(0) = 1$. The oscillatory integral is well-defined for any g of class $\mathcal{A}$ and independent of the particular choice of the function χ.
If we choose $l, l' \in \mathbb{N}_0$ sufficiently large (depending on m, δ and τ) the oscillatory integral coincides with the ordinary integral

$$O_S - \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(y,\eta)}\, g(\eta,y)\, dy đ\eta \quad = \quad \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(y,\eta)}\,\langle y\rangle^{-2l'}\langle D_\eta\rangle^{2l'}\left\{\langle\eta\rangle^{-2l}\langle D_y\rangle^{2l} g(\eta,y)\right\} dy đ\eta,$$
(6.19)

where we use the standard notation $D_x = (D_{x_1},\ldots,D_{x_n}) = (-i\partial_{x_1},\ldots,-i\partial_{x_n})$. Moreover the following partial integration rule holds

$$(6.20)\, O_S - \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(y,\eta)}\,\eta^\alpha\, g(\eta,y)\, dy đ\eta = O_S - \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(y,\eta)}\, D_y^\alpha g(\eta,y)\, dy đ\eta, \qquad \alpha \in \mathbb{N}_0^n.$$

We introduce the class of **double symbols** in terms of the reference or weight function λ.

Definition 6.8. *Let m_1, $m_2 \in \mathbb{R}$. The class $S_0^{m_1,m_2,\lambda}$ of double symbols of order m_1 and m_2 denotes all C^∞-functions $p : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ satisfying*

$$(6.21) \qquad \left|\partial_\xi^\alpha \partial_x^\beta \partial_{\xi'}^{\alpha'} \partial_{x'}^{\beta'} p(x,\xi,x',\xi')\right| \le c_{\alpha,\beta,\alpha',\beta'}\lambda(\xi)^{m_1}\cdot\lambda(\xi')^{m_2}, \quad \alpha,\beta,\alpha',\beta' \in \mathbb{N}_0^n.$$

For $p \in S_0^{m,m',\lambda}$ we define the corresponding operator

$$(6.22) \qquad p(x,D_x,x',D_{x'})u(x) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{i((x-x'),\xi)+i(x',\xi')} p(x,\xi,x',\xi')\hat{u}(\xi')đ\xi' dx' đ\xi.$$

for all $u \in \mathcal{S}(\mathbb{R}^n,\mathbb{C})$.

As in the classical situation it turns out that double symbols determine the same classes of operators $\Psi_0^{m,\lambda}$ as simple symbols, but they are a very useful tool for their investigation. More precisely we have

Theorem 6.9. *Let $p \in S_0^{m,m',\lambda}$ and $u \in \mathcal{S}(\mathbb{R}^n,\mathbb{C})$. Then the iterated integral in (6.22) exists and defines a pseudo differential operator in the class $\Psi_0^{m+m',\lambda}$. Moreover*

$$(6.23) \qquad p_L(x,\xi) = O_S - \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(y,\eta)}\, p(x,\xi+\eta,x+y,\xi) dy đ\eta$$

is a symbol in $S_0^{m+m',\lambda}$ and defines the same operator, i.e.

$$p(x,D_x,x',D_{x'})u = p_L(x,D)u$$

for all $u \in \mathcal{S}(\mathbb{R}^n,\mathbb{C})$.

Definition 6.10. *In the situation of Theorem 6.9* $p_L(x,\xi)$ *is called the* **simplified symbol** *of* $p(x,\xi,x',\xi')$.

Recall that by Lemma 2.6

$$\frac{\lambda^s(\xi)}{\lambda^s(\eta)} \leq 2^{|s|/2}\,\lambda^{|s|}(\xi-\eta) \tag{6.24}$$

for all $s\in\mathbb{R}$ and all $\xi,\eta\in\mathbb{R}^n$. Thus by (6.24) and (2.8)

$$\left|\partial_\eta^\alpha\partial_y^\beta p(x,\xi+\eta,x+y,\xi)\right| \leq c\lambda^m(\xi+\eta)\lambda^{m'}(\xi) \leq c\lambda^{m+m'}(\xi)\cdot\lambda^{|m|}(\eta) \leq c_\xi\langle\eta\rangle^{|m|}.$$

Therefore the integrand in (6.23) is of class $\mathcal{A}$ and the integral is well defined. Moreover note

Remark 6.11. The oscillatory integral in (6.23) actually defines a symbol p_L in $\mathrm{S}_0^{m+m',\lambda}$. To see this we use the representation (6.19) for the oscillatory integral. For l,l' sufficiently large we get by exchanging differentiation and integration, (6.21) and (6.24)

$$\begin{aligned}\left|\partial_\xi^\alpha\partial_x^\beta p_L(x,\xi)\right| &\leq c_{\alpha\beta}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\langle\eta\rangle^{-2l}\langle y\rangle^{-2l'}\lambda^m(\xi+\eta)\lambda^{m'}(\xi)\,dy đ\eta \\ &\leq c_{\alpha\beta}\lambda^{m+m'}(\xi).\end{aligned}\tag{6.25}$$

Moreover note that the constants $c_{\alpha\beta}$ are expressed in terms of the constants $c_{\alpha\beta\alpha'\beta}$ for the double symbol in (6.21). In particular, if a family of double symbols satisfies (6.21) uniformly for each $\alpha,\beta,\alpha',\beta'$, then also the simplified symbols satisfy an estimate (6.25) with uniform constants $c_{\alpha\beta}$.

Proof of Theorem 6.9: We adapt the consideration in [54], Chapter 2, to our situation. Choose $\chi\in\mathcal{S}(\mathbb{R}^n\times\mathbb{R}^n)$ such that $\chi(0)=1$ and note that (see [54], Lemma 1.6.3)

$$\partial_\eta^\alpha\partial_y^\beta\,|\chi(\varepsilon\eta,\varepsilon y)| \leq c_{\alpha\beta}\langle\eta\rangle^{-|\alpha|}\langle y\rangle^{-|\beta|} \quad \text{uniformly for } 0\leq\varepsilon\leq 1. \tag{6.26}$$

For $0\leq\varepsilon\leq 1$ let $p_\varepsilon(x,\xi,x',\xi')=\chi(\varepsilon(\xi-\xi'),\varepsilon(x'-x))p(x,\xi,x',\xi')$. Then by Leibniz rule and (6.26) have

$$\left|\partial_\xi^\alpha\partial_x^\beta\partial_{\xi'}^{\alpha'}\partial_{x'}^{\beta'}p_\varepsilon(x,\xi,x',\xi')\right| \leq c_{\alpha,\beta,\alpha',\beta'}\lambda^m(\xi)\lambda^{m'}(\xi') \tag{6.27}$$

with constants $c_{\alpha,\beta,\alpha',\beta'}$ independent of ε. Define

$$\begin{aligned} p_{u,\varepsilon}(x,\xi,x',\xi') &= p_\varepsilon(x,\xi,x',\xi')\hat{u}(\xi') \\ q_{u,\varepsilon}(x,\xi,x') &= \int_{\mathbb{R}^n} e^{i(x',\xi')}p_{u,\varepsilon}(x,\xi,x',\xi')đ\xi' \\ r_{u,\varepsilon}(x,\xi) &= \int_{\mathbb{R}^n} e^{-i(x',\xi)}q_{u,\varepsilon}(x,\xi,x')dx' \end{aligned}$$

and fix $l,n_0\in\mathbb{N}$ such that $2l>n+m^+$ and $2n_0>n$. Note that $e^{i(x',\xi')}=\langle x'\rangle^{-2n_0}\langle D_{\xi'}\rangle^{2n_0}e^{i(x',\xi')}$. Thus for all $|\beta'|\leq 2l$ by partial integration and Leibniz rule

$$\begin{aligned}\left|\partial_{x'}^{\beta'}q_{u,\varepsilon}(x,\xi,x')\right| &\leq \left|\partial_{x'}^{\beta'}\int_{\mathbb{R}^n}\langle x'\rangle^{-2n_0}e^{i(x',\xi')}\langle D_{\xi'}\rangle^{2n_0}p_{u,\varepsilon}(x,\xi,x',\xi')đ\xi'\right| \\ &\leq c_{p,u,l,n_0}\lambda^m(\xi)\langle x'\rangle^{-2n_0},\end{aligned}\tag{6.28}$$

where the estimate is again uniform in ε. Therefore $r_{u,\varepsilon}$ is well defined and as above

$$\begin{aligned} |r_{u,\varepsilon}(x,\xi)| &\leq \left| \langle\xi\rangle^{-2l} \int_{\mathbb{R}^n} e^{-i(x',\xi)} \langle D_{x'}\rangle^{2l} q_{u,\varepsilon}(x,\xi,x')dx' \right| \\ &\leq c_{p,u,l,n_0} \lambda^m(\xi) \cdot \langle\xi\rangle^{-2l} \leq c_{p,u,l,n_0,\lambda} \langle\xi\rangle^{-2l+m^+} \end{aligned} \tag{6.29}$$

uniformly in ε, where the last inequality follows from (2.8). Thus the integral

$$p_\varepsilon(x, D_x, x', D_{x'})u(x) = \int_{\mathbb{R}^n} e^{i(x,\xi)} r_{u,\varepsilon}(x,\xi) đ\xi$$

exists. In particular for $\varepsilon = 0$ we see that the iterated integral in (6.22) is well defined. Moreover, since the estimates (6.28) and (6.29) are uniform with respect to $0 \leq \varepsilon \leq 1$, we find by a successive application of Lebesgue's theorem

$$\begin{aligned} p(x, D_x, x', D_{x'})u(x) &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i((x-x'),\xi)+i(x',\xi')} \lim_{\varepsilon\to 0} p_{u,\varepsilon}(x,\xi,x',\xi') đ\xi' dx đ\xi \\ &= \lim_{\varepsilon\to 0} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i((x-x'),\xi)+i(x',\xi')} p_{u,\varepsilon}(x,\xi,x',\xi') đ\xi' dx đ\xi \\ &= \lim_{\varepsilon\to 0} p_\varepsilon(x, D_x, x', D_{x'})u(x). \end{aligned} \tag{6.30}$$

For $\varepsilon > 0$ define

$$p_{L,\varepsilon}(x,\xi) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-i(y,\eta)} \chi(\varepsilon\eta, \varepsilon y) p(x, \xi+\eta, x+y, \xi) dy đ\eta, \tag{6.31}$$

Then by definition of the oscillatory integral

$$\lim_{\varepsilon\to 0} p_{L,\varepsilon}(x,\xi) = p_L(x,\xi) \tag{6.32}$$

and moreover by partial integration for $l_1, l_1' \in \mathbb{N}_0$ such that $2l_1 > |m| + n$, $2l_1' > n$

$$\begin{aligned} |p_{L,\varepsilon}(x,\xi)| &= \left| \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-i(y,\eta)} \langle\eta\rangle^{-2l_1} \langle D_y\rangle^{2l_1} \left\{ \langle y\rangle^{-2l_1'} \langle D_\eta\rangle^{2l_1'} \chi(\varepsilon\eta,\varepsilon y) p(x,\xi+\eta,x+y,\xi) \right\} dy đ\eta \right| \\ &\leq c \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \langle\eta\rangle^{-2l_1} \langle y\rangle^{-2l_1'} \lambda^m(\xi+\eta) \lambda^{m'}(\xi) dy đ\eta \\ &\leq c \int_{\mathbb{R}^n} \langle\eta\rangle^{-2l_1+|m|} \lambda^{m+m'}(\xi) đ\eta \\ &= c\lambda^{m+m'}(\xi) \end{aligned}$$

uniformly in $0 < \varepsilon \leq 1$. Therefore by (6.32)

$$\lim_{\varepsilon\to 0} p_{L,\varepsilon}(x,D)u(x) = p_L(x,D)u(x), \qquad u \in \mathcal{S}(\mathbb{R}^n, \mathbb{C}). \tag{6.33}$$

On the other hand substituting $x' = x + y$ and $\xi = \xi' + \eta$ shows

$$\begin{aligned} p_\varepsilon(x, D_x, x', D_{x'})u(x) &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i((x-x'),\xi)+i(x',\xi')} p_\varepsilon(x,\xi,x',\xi') \hat{u}(\xi') đ\xi' dx' đ\xi \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i(x,\xi')} e^{-i(y,\eta)} \chi(\varepsilon\eta,\varepsilon y) p(x,\xi'+\eta,x+y,\xi') \hat{u}(\xi') đ\xi' dy đ\eta \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i(x,\xi')} p_{L,\varepsilon}(x,\xi') \hat{u}(\xi') đ\xi' \\ &= p_{L,\varepsilon}(x,D)u(x). \end{aligned} \tag{6.34}$$

Thus combining (6.30), (6.33) and (6.34) gives

$$p(x, D_x, x', D_{x'})u(x) = p_L(x, D)u(x).$$

□

Theorem 6.9 has a series of useful corollaries. First we consider the composition of two operators.

Corollary 6.12. *Let $p_i \in S_0^{m_i,\lambda}$, $m_i \in \mathbb{R}$, $i = 1, 2$. Then $p_1(x, D) \circ p_2(x, D) \in \Psi_0^{m_1+m_2,\lambda}$.*

Proof: Put $p(x, \xi, x', \xi') = p_1(x, \xi) \cdot p_2(x', \xi')$. Then $p \in \mathrm{S}_0^{m_1,m_2,\lambda}$. Therefore $p_L(x, D) \in \Psi_0^{m_1+m_2,\lambda}$ and for $u \in \mathcal{S}(\mathbb{R}^n, \mathbb{C})$

$$\begin{aligned} p_1(x, D) \circ p_2(x, D)u(x) &= \int_{\mathbb{R}^n} e^{i(x,\xi)} p_1(x, \xi) \int_{\mathbb{R}^n} e^{-i(x',\xi)} \int_{\mathbb{R}^n} e^{i(x',\xi')} p_2(x', \xi')\hat{u}(\xi') đ\xi' dx' đ\xi \\ &= p(x, D_x, x', D_{x'})u(x) = p_L(x, D)u(x). \end{aligned}$$

□

We also can handle the formally adjoint operator in $L^2(\mathbb{R}^n, \mathbb{C})$.

Corollary 6.13. *Let $p \in S_0^{m,\lambda}$. Then there is a $p^* \in S_0^{m,\lambda}$ such that*

$$(p(x, D)u, v)_0 = (u, p^*(x, D)v)_0$$

for all $u, v \in \mathcal{S}(\mathbb{R}^n, \mathbb{C})$.

Proof: Define $\tilde{p}(x, \xi, x', \xi') = \overline{p(x', \xi)}$. Then $\tilde{p} \in \mathrm{S}_0^{m,0,\lambda}$ and as in the proof of Corollary 2.2.5 in [54]

$$\begin{aligned} (p(x, D)u, v)_0 &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i(x',\xi)} p(x', \xi)\hat{u}(\xi) đ\xi \cdot \overline{v(x')} dx' \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-i(x,\xi)} u(x) \left\{ \int_{\mathbb{R}^n} e^{i(x',\xi)} p(x', \xi)\overline{v(x')} dx' \right\} dx đ\xi \\ &= \int_{\mathbb{R}^n} u(x) \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-i(x-x',\xi)} p(x', \xi)\overline{v(x')} dx' đ\xi dx \\ &= \int_{\mathbb{R}^n} u(x) \overline{\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i(x-x',\xi)+i(x',\xi')} \overline{p(x', \xi)} \hat{v}(\xi') đ\xi' dx' đ\xi} dx \\ &= (u, \tilde{p}(x, D_x, x', D_{x'})v)_0, \end{aligned}$$

which proves the corollary with $p^*(x, D) = \tilde{p}_L(x, D)$. Here we applied Fubini's theorem several times. This is possible in particular since

$$\left| \int_{\mathbb{R}^n} e^{i(x',\xi)} p(x', \xi)\overline{v(x')} dx' \right| = \langle \xi \rangle^{-2n_0} \left| \int_{\mathbb{R}^n} e^{i(x',\xi)} \langle D_{x'} \rangle^{2n_0} (p(x', \xi)\overline{v(x')}) dx' \right| \leq c \langle \xi \rangle^{-2n_0} \lambda^m(\xi)$$

is integrable w.r.t. ξ for $n_0 \in \mathbb{N}$ sufficiently large. □

Summerizing we find that $\bigcup_{m\in\mathbb{R}} \Psi_0^{m,\lambda}$ is an algebra of pseudo differential operators with multiplication $\circ$ and involution $*$ that respects the graded structure given by $(\mathrm{S}_0^{m,\lambda})_{m\in\mathbb{R}}$, i.e.

$$\begin{aligned} \Psi_0^{m,\lambda} + \Psi_0^{m,\lambda} &\subset \Psi_0^{m,\lambda} \\ (\Psi_0^{m,\lambda})^* &\subset \Psi_0^{m,\lambda} \\ \Psi_0^{m,\lambda} \circ \Psi_0^{m',\lambda} &\subset \Psi_0^{m+m',\lambda} \end{aligned}$$

Next we extend the domain of the operators. Corollary 6.13 immediately implies by duality that $p(x,D) \in \Psi_0^{m,\lambda}$ has a continuous extension $p(x,D) : \mathcal{S}'(\mathbb{R}^n,\mathbb{C}) \to \mathcal{S}'(\mathbb{R}^n,\mathbb{C})$ defined by

$$\langle p(x,D)u, v\rangle = \langle u, p^*(x,D)v\rangle, \qquad u \in \mathcal{S}'(\mathbb{R}^n,\mathbb{C}),\ v \in \mathcal{S}(\mathbb{R}^n,\mathbb{C}).$$

We show that the order m of an operator $p(x,D) \in \Psi_0^{m,\lambda}$ has a natural interpretation in terms of mapping properties between the anisotropic Sobolev spaces introduced in Chapter 4, see (4.5), (4.6); the definition extends immediately to the complex-valued case.

Theorem 6.14. *A pseudo differential operator with symbol $p \in S_0^{m,\lambda}$ is a continuous operator*

$$p(x,D) : H^{s+m,\lambda}(\mathbb{R}^n,\mathbb{C}) \to H^{s,\lambda}(\mathbb{R}^n,\mathbb{C})$$

for all $s \in \mathbb{R}$ and we have

$$\|p(x,D)u\|_{s,\lambda} \le c\,\|u\|_{s+m,\lambda} \quad \textit{for all } u \in H^{s+m,\lambda}(\mathbb{R}^n,\mathbb{C}). \tag{6.35}$$

Proof: It is sufficient to prove (6.35) for $u \in \mathcal{S}(\mathbb{R}^n,\mathbb{C})$. First suppose $s = m = 0$. Then $p \in \mathrm{S}_0^{0,\lambda}$ has bounded derivatives and by the well-known L^2-continuity result of Calderón and Vaillancourt [9] we find

$$\|p(x,D)u\|_0 \le c\,\|u\|_0$$

with a constant c depending only on the constants $c_{\alpha\beta}$ in (6.11) for $|\alpha|, |\beta| \le 3$. Next suppose $s = 0$ and m arbitrary. Then

$$p(x,D)u(x) = \int_{\mathbb{R}^n} e^{i(x,\xi)} p(x,\xi)\lambda^{-m}(\xi)\lambda^m(\xi)\hat{u}(\xi)đ\xi$$

and $p(x,\xi)\lambda^{-m}(\xi)$ is a symbol in $\mathrm{S}_0^{0,\lambda}$. Therefore

$$\|p(x,D)u\|_0 \le c\,\|\lambda^m(D)u\|_0 = c\,\|u\|_{m,\lambda}.$$

Finally for the general case observe that $\lambda^s(D) \circ p(x,D) \in \Psi_0^{s+m,\lambda}$ by Corollary 6.12 and thus

$$\|p(x,D)u\|_{s,\lambda} = \|\lambda^s(D)p(x,D)u\|_0 \le c\,\|u\|_{s+m,\lambda}.$$

□

Remark 6.15. Observe that from the above proof, Corollary 6.12 and Remark 6.11 it is clear that the same constant c in (6.35) may be chosen for a family of pseudo differential operators which satisfy (6.11) uniformly.

The symbol classes $\mathrm{S}_0^{m,\lambda}$ lead to a reasonable algebra of pseudo differential operators, but are bad symbol classes in the sense that all derivatives of the symbols are estimated by the same power m of $\lambda(\xi)$ as in the case of Hörmander class $S^m_{0,0}$ and not by a smaller power. Therefore we cannot expect asymptotic expansion formulas for this type of symbols. On the other hand the symbols of class $\mathrm{S}_\varrho^{m,\lambda}$ have a somewhat better behaviour of their derivatives with respect to ξ. This will yield expansion formulas including terms up to order 2. We consider the expansion of the simplified symbol.

Theorem 6.16. *Given a double symbol $p \in S_0^{m,m',\lambda}$ such that*

$$\partial_\xi^\alpha p(x,\xi,x',\xi') \in S_0^{m-\varrho(|\alpha|),m',\lambda} \tag{6.36}$$

holds for all $\alpha \in \mathbb{N}_0^n$. Then for all $N \in \mathbb{N}$ the simplified symbol p_L satiesfies

$$p_L(x,\xi) - \sum_{|\alpha|<N} \frac{1}{\alpha!} p_\alpha(x,\xi) \in S_0^{m+m'-\varrho(N),\lambda}, \tag{6.37}$$

where

$$p_\alpha(x,\xi) = D_{x'}^\alpha \partial_\xi^\alpha p(x,\xi,x',\xi')\big|_{\substack{x'=x\\ \xi'=\xi}} \in S_0^{m+m'-\varrho(|\alpha|),\lambda}. \tag{6.38}$$

Proof: We modify the argument given in [67]. By Taylor's formula we have

$$\begin{aligned} p(x,\xi+\eta,x+z,\xi) &= \sum_{|\alpha|<N} \frac{\eta^\alpha}{\alpha!} \partial_\xi^\alpha\, p(x,\xi,x+z,\xi')|_{\xi'=\xi} \\ &\quad + N \sum_{|\gamma|=N} \frac{\eta^\gamma}{\gamma!} p_\gamma(x,z,\xi,\eta) \end{aligned}$$

with

$$p_\gamma(x,z,\xi,\eta) = \int_0^1 (1-t)^{N-1} \partial_\xi^\gamma\, p(x,\xi+t\eta,x+z,\xi')|_{\xi'=\xi}\, dt$$

and therefore by (6.23)

$$\begin{aligned} p_L(x,\xi) &= \sum_{|\alpha|<N} \frac{1}{\alpha!} O_S - \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(z,\eta)} \eta^\alpha \partial_\xi^\alpha\, p(x,\xi,x+z,\xi')|_{\xi'=\xi}\, dz đ\eta \\ &\quad + \sum_{|\gamma|=N} \frac{N}{\gamma!} O_S - \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(z,\eta)} \eta^\gamma p_\gamma(x,z,\xi,\eta)\, dz đ\eta \\ &= \sum_{|\alpha|<N} \frac{1}{\alpha!} I_\alpha(x,\xi) + \sum_{|\gamma|=N} \frac{N}{\gamma!} J_\gamma(x,\xi). \end{aligned}$$

We have to show that

$$I_\alpha = p_\alpha \in \mathrm{S}_0^{m+m'-\varrho(|\alpha|),\lambda}, \quad |\alpha| < N \tag{6.39}$$

and

$$J_\gamma \in \mathrm{S}_0^{m+m'-\varrho(N),\lambda}. \tag{6.40}$$

Let $|\alpha| < N$ and choose $\chi_1, \chi_2 \in C_0^\infty(\mathbb{R}^n)$ such that χ_1 and χ_2 equal 1 in a neighbourhood of the origin. Then by definition of I_α and (6.20)

$$\begin{aligned}
I_\alpha &= O_S - \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(z,\eta)}\eta^\alpha \partial_\xi^\alpha\, p(x,\xi,x+z,\xi')|_{\xi'=\xi}\, dz đ\eta \\
&= O_S - \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(z,\eta)} D_z^\alpha \partial_\xi^\alpha\, p(x,\xi,x+z,\xi')|_{\xi'=\xi}\, dz đ\eta \\
&= \lim_{\varepsilon\to 0}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(z,\eta)}\chi_1(\varepsilon\eta)\chi_2(\varepsilon z) D_{x'}^\alpha \partial_\xi^\alpha\, p(x,\xi,x',\xi')|_{\substack{x'=x+z\\ \xi'=\xi}}\, dz đ\eta \\
&= \lim_{\varepsilon\to 0}\int_{\mathbb{R}^n} \chi_2(\varepsilon z)\varepsilon^{-n}\hat{\chi_1}(\frac{z}{\varepsilon}) D_{x'}^\alpha \partial_\xi^\alpha\, p(x,\xi,x',\xi')|_{\substack{x'=x+z\\ \xi'=\xi}}\, dz \\
&= D_{x'}^\alpha \partial_\xi^\alpha\, p(x,\xi,x',\xi')|_{\substack{x'=x\\ \xi'=\xi}} = p_\alpha(x,\xi),
\end{aligned}$$

because $\chi_2(\varepsilon z)\varepsilon^{-n}\hat{\chi_1}(\frac{z}{\varepsilon})$ converges to the unit mass at 0 as $\varepsilon \to 0$, and $p_\alpha \in \mathrm{S}_0^{m+m'-\varrho(|\alpha|),\lambda}$ by (6.36).
Moreover for $|\gamma| = N$ we have by (6.36), (6.24) and (2.8)

$$\begin{aligned}
&\left|\partial_x^\alpha \partial_z^{\alpha'} \partial_\xi^\beta \partial_\eta^{\beta'} p_\gamma(x,z,\xi,\eta)\right| \\
&= \left|\int_0^1 (1-t)^{N-1}\partial_x^\alpha \partial_z^{\alpha'} \partial_\xi^\beta \partial_\eta^{\beta'} \left(\partial_\xi^\gamma\, p(x,\xi+t\eta,x+z,\xi')|_{\xi'=\xi}\right) dt\right| \\
&\le c_{\alpha,\alpha',\beta,\beta',\gamma}\int_0^1 \lambda^{m-\varrho(N)}(\xi+t\eta)\cdot\lambda^{m'}(\xi) dt \\
&\le c_{\alpha,\alpha',\beta,\beta',\gamma}\lambda^{m-\varrho(N)}(\xi)\lambda^{m'}(\xi)\int_0^1 \lambda^{|m-\varrho(N)|}(t\eta)dt \\
&\le c_{\alpha,\alpha',\beta,\beta',\gamma,\lambda}\lambda^{m+m'-\varrho(N)}(\xi)\ \cdot\langle\eta\rangle^{|m-\varrho(N)|}.
\end{aligned}$$

Hence again by (6.19) for $l, n_0 \in \mathbb{N}$, $2l > N + |m-\varrho(N)| + n$, $2n_0 > n$,

$$\begin{aligned}
&\left|\partial_\xi^\alpha \partial_x^\beta J_\gamma(x,\xi)\right| \\
&= \left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(z,\eta)}\langle\eta\rangle^{-2l}\langle D_z\rangle^{2l}\left\{\langle z\rangle^{-2n_0}\langle D_\eta\rangle^{2n_0}\left[\eta^\gamma \partial_\xi^\alpha \partial_x^\beta p_\gamma(x,z,\xi,\eta)\right]\right\} dz đ\eta\right| \\
&\le c_{l,\alpha,\beta}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\langle\eta\rangle^{-2l+N+|m-\varrho(N)|}\cdot\langle z\rangle^{-2n_0}\lambda(\xi)^{m+m'-\varrho(N)} dz đ\eta \\
&\le c_{l,\alpha,\beta}\lambda(\xi)^{m+m'-\varrho(N)}.
\end{aligned}$$

which gives (6.40). □

Remark 6.17 . The proof shows that p_α and the remainder term $p_L - \sum_{|\alpha|<N}\frac{1}{\alpha!}p_\alpha$ are in the class $\mathrm{S}_0^{m+m'-\varrho(|\alpha|),\lambda}$ and $\mathrm{S}_0^{m+m'-\varrho(N),\lambda}$, respectively, and moreover satisfy estimates (6.11) with constants $c_{\alpha\beta}$ that depend only on the constants $c_{\alpha,\beta,\alpha',\beta'}$ in (6.21) of the double symbol $p(x,\xi,x',\xi')$ itself.

We apply Theorem 6.16 to the double symbols of the composition of two pseudo differential operators and of the formally adjoint operator, see Corollaries 6.12, 6.13 and their proofs, and obtain the following expansion formulas.

Corollary 6.18. *Let $p_1 \in S_\varrho^{m_1,\lambda}$, $p_2 \in S_\varrho^{m_2,\lambda}$ and $p \in S_\varrho^{m,\lambda}$. Then the symbols p_c and p^* of the composition $p_c(x,D) = p_1(x,D) \circ p_2(x,D)$ and the formally adjoint $p^*(x,D) = p(x,D)^*$ satisfy*

$$p_c(x,\xi) = p_1(x,\xi) \cdot p_2(x,\xi) + \sum_{j=1}^{n} \partial_{\xi_j} p_1(x,\xi) \cdot D_{x_j} p_2(x,\xi) + p_{r_1}(x,\xi)$$

and

$$p^*(x,\xi) = \overline{p(x,\xi)} + \sum_{j=1}^{n} \partial_{\xi_j} D_{x_j} \overline{p(x,\xi)} + p_{r_2}(x,\xi),$$

where $p_{r_1} \in S_0^{m_1+m_2-2,\lambda}$ and $p_{r_2} \in S_0^{m-2,\lambda}$.

Note that in particular the highest order terms are given by the product and the conjugate of the symbols.

Remark 6.19 .

(i) Since $\varrho(k) \leq 2$, the expansions given by formula (6.37) do not improve for $N > 2$ towards terms of lower order. In this sense we obtain expansion formulas with terms up to order two. Obviously this effect is due to the choice of the function $\varrho(k) = k \wedge 2$, which is determined by the behaviour of negative definite symbols.

(ii) Of course the statement itself does not depend on the specific choice of ϱ and choosing another increasing subadditive function $\varrho : \mathbb{N}_0 \to \mathbb{R}_+$ will not affect the proof.

6.4 Friedrichs symmetrization

It is well-known that a pseudo differential operator with real-valued symbol is in general no symmetric operator if the symbol depends on x, but there is a modification that is symmetric and differs from the original operator only by a lower order perturbation. This modification can be constructed explictly by the so-called Friedrichs symmetrization. The purpose of this section is to show by the results obtained in the previous section that also for symbols in $\mathrm{S}_\varrho^{m,\lambda}$ a Friedrichs symmetrization is available. For that end fix a function $q \in C_0^\infty(\mathbb{R}^n)$ such that q is even, non-negative, supported in the unit ball $B_1(0)$ and $\int_{\mathbb{R}^n} q^2(\sigma) d\sigma = 1$ and define

$$F(\xi,\zeta) = \lambda(\xi)^{-n/4} \cdot q((\zeta - \xi) \cdot \lambda^{-1/2}(\xi)). \tag{6.41}$$

For a symbol $p \in \mathrm{S}_0^{m,\lambda}$ let us define its **Friedrichs symmetrization** to be the double symbol p_F not depending on x given by

$$p_F(\xi, x', \xi') = \int_{\mathbb{R}^n} F(\xi,\zeta) p(x',\zeta,) F(\xi',\zeta) d\zeta.$$

Then we have

Theorem 6.20. *Let $p \in S_0^{m,\lambda}$. Then*

$$\left|\partial_{x'}^{\alpha'}\partial_\xi^\beta\partial_{\xi'}^{\beta'} p_F(\xi,x',\xi')\right| \le c_{\alpha',\beta,\beta'}\lambda^{m-\frac{1}{2}\varrho(|\beta|)}(\xi)\cdot\lambda^{-\frac{1}{2}\varrho(|\beta'|)}(\xi'). \tag{6.42}$$

In particular $p_F \in S_0^{m,0,\lambda}$ and the simplified symbol $p_{F,L} \in S_0^{m,\lambda}$. Moreover, if $p \in S_\varrho^{m,\lambda}$ we have

$$p - p_{F,L} \in S_0^{m-1,\lambda}. \tag{6.43}$$

First we prove

Lemma 6.21. *For all $\beta \in \mathbb{N}_0^n$ we have*

$$\partial_\xi^\beta F(\xi,\zeta) = \lambda(\xi)^{-\frac{n}{4}} \sum_{\substack{|\gamma|\le|\beta|\\ \gamma_1\le\gamma}} \varphi_{\beta,\gamma,\gamma_1}(\xi)\cdot\Big((\xi-\zeta)\cdot\lambda^{-\frac{1}{2}}(\xi)\Big)^{\gamma_1}\cdot\Big(\partial^\gamma q\Big)((\xi-\zeta)\cdot\lambda^{-\frac{1}{2}}(\xi)), \tag{6.44}$$

where $\varphi_{\beta.\gamma,\gamma_1} \in S_0^{-\frac{1}{2}\varrho(|\beta|),\lambda}$.

Proof: Obviously (6.44) holds true for $\beta = 0$ with $\varphi_{0,0,0} = 1$. Note that

$$\partial_{\xi_i}\lambda^m(\xi) = m\lambda^m(\xi)\lambda^{-1}(\xi)\cdot\partial_{\xi_i}\lambda(\xi).$$

Proceeding by induction we differentiate (6.44)

$$\begin{aligned}
&\partial_{\xi_j}\partial_\xi^\beta F(\xi,\zeta) = \\
&= \lambda^{-n/4}(\xi)\cdot\sum_{\substack{|\gamma|\le|\beta|\\ \gamma_1\le\gamma}} \Big\{ \Big[\psi^{(1)}_{\beta,\gamma,\gamma_1}(\xi)+\psi^{(2)}_{\beta,\gamma,\gamma_1}(\xi)\Big]((\xi-\zeta)\cdot\lambda^{-1/2}(\xi))^{\gamma_1}\cdot(\partial^\gamma q)((\xi-\zeta)\cdot\lambda^{-1/2}(\xi)) \\
&\quad + \psi^{(3)}_{\beta,\gamma,\gamma_1}(\xi)((\xi-\zeta)\cdot\lambda^{-1/2}(\xi))^{\gamma_1-\varepsilon_j}\cdot(\partial^\gamma q)((\xi-\zeta)\cdot\lambda^{-1/2}(\xi)) \\
&\quad + \psi^{(4)}_{\beta,\gamma,\gamma_1}(\xi)((\xi-\zeta)\cdot\lambda^{-1/2}(\xi))^{\gamma_1}\cdot(\partial^{\gamma+\varepsilon_j} q)((\xi-\zeta)\cdot\lambda^{-1/2}(\xi)) \\
&\quad + \psi^{(5)}_{\beta,\gamma,\gamma_1}(\xi)\sum_{k=1}^n((\xi-\zeta)\cdot\lambda^{-1/2}(\xi))^{\gamma_1+\varepsilon_k}\cdot(\partial^{\gamma+\varepsilon_k} q)((\xi-\zeta)\cdot\lambda^{-1/2}(\xi))\Big\}
\end{aligned}$$

with

$$\begin{aligned}
\psi^{(1)}_{\beta,\gamma,\gamma_1}(\xi) &= -\varphi_{\beta,\gamma,\gamma_1}(\xi)\cdot\left(\frac{n}{4}+\frac{|\gamma_1|}{2}\right)\lambda^{-1}(\xi)\partial_{\xi_j}\lambda(\xi),\\
\psi^{(2)}_{\beta,\gamma,\gamma_1}(\xi) &= \partial_{\xi_j}\varphi_{\beta,\gamma,\gamma_1}(\xi),\\
\psi^{(3)}_{\beta,\gamma,\gamma_1}(\xi) &= \gamma_{1,j}\lambda^{-1/2}(\xi)\varphi_{\beta,\gamma,\gamma_1}(\xi),\\
\psi^{(4)}_{\beta,\gamma,\gamma_1}(\xi) &= \lambda^{-1/2}(\xi)\varphi_{\beta,\gamma,\gamma_1}(\xi),\\
\psi^{(5)}_{\beta,\gamma,\gamma_1}(\xi) &= -\frac{1}{2}\lambda^{-1}(\xi)\partial_{\xi_j}\lambda(\xi)\varphi_{\beta,\gamma,\gamma_1}(\xi),
\end{aligned}$$

which is of the form claimed in (6.44) and we have to check that $\psi^{(l)}_{\beta,\gamma,\gamma_1} \in \mathrm{S}_0^{-\frac{1}{2}(\varrho(|\beta|+1)),\lambda}$, $l = 1,\ldots,5$. Note that $\lambda^{-1/2} \in \mathrm{S}_0^{-1/2,\lambda}$ and $\lambda^{-1}\partial_{\xi_j}\lambda \in \mathrm{S}_0^{-1,\lambda}$, see Lemma 6.5. Since $\varphi_{0,0,0} = 1$ we see for $\beta = 0$ that

$$\psi^{(l)}_{\beta,\gamma,\gamma_1} \in \ \mathrm{lin}\{\lambda^{-1/2}, \lambda^{-1}\partial_{\xi_j}\lambda\} \subset \mathrm{S}_0^{-1/2,\lambda},$$

which also implies $\varphi_{\beta,\gamma,\gamma_1} \in \mathrm{S}_0^{-1/2,\lambda}$ for $|\beta| = 1$. Next note that $\partial_{\xi_k}\lambda^{-1/2} \in \mathrm{S}_0^{-3/2,\lambda}$ and $\partial_{\xi_k}(\lambda^{-1}\partial_{\xi_i}\lambda) \in \mathrm{S}_0^{-2,\lambda}$, which yields $\partial_{\xi_k}\varphi_{\beta,\gamma,\gamma_1} \in \mathrm{S}_0^{-3/2,\lambda}$ for $|\beta| = 1$. Thus by the algebra property (6.16) of the symbols we find for $|\beta| = 1$ that $\psi^{(l)}_{\beta,\gamma,\gamma_1} \in \mathrm{S}_0^{-1,\lambda}$.
But $\mathrm{S}_0^{-1,\lambda}$ is stable under taking derivatives and therefore again (6.16) yields $\psi^{(l)}_{\beta,\gamma,\gamma_1} \in \mathrm{S}_0^{-1,\lambda}$ for all $|\beta| \geq 2$ by induction. □

Proof of Theorem 6.20, Estimate (6.42): By Lemma 6.21 and the support properties of q we have

$$\begin{aligned}
&\left|\partial^{\alpha'}_{x'}\partial^{\beta}_{\xi}\partial^{\beta'}_{\xi'} p_F(\xi,x',\xi')\right| \\
&= \left|\int_{\mathbb{R}^n} \partial^{\beta}_{\xi}F(\xi,\zeta)\partial^{\alpha'}_{x'}p(x',\zeta,)\partial^{\beta'}_{\xi'}F(\xi',\zeta)d\zeta\right| \\
&\leq \lambda(\xi)^{-n/4}\lambda(\xi')^{-n/4}\sum_{\substack{|\gamma|\leq|\beta|\\ \gamma_1\leq\gamma}}\sum_{\substack{|\gamma'|\leq|\beta'|\\ \gamma_1'\leq\gamma'}}\left|\varphi_{\beta,\gamma,\gamma_1}(\xi)\cdot\varphi_{\beta',\gamma',\gamma_1'}(\xi')\right|\cdot \\
&\quad \cdot\Bigg|\int_{\substack{|\xi-\zeta|\leq\lambda^{1/2}(\xi)\\ |\xi'-\zeta|\leq\lambda^{1/2}(\xi')}}((\xi-\zeta)\cdot\lambda^{-1/2}(\xi))^{\gamma_1}\ ((\xi'-\zeta)\cdot\lambda^{-1/2}(\xi'))^{\gamma_1'} \\
&\qquad\qquad \cdot(\partial^{\gamma}q)((\xi-\zeta)\cdot\lambda^{-1/2}(\xi))\ (\partial^{\gamma'}q)((\xi'-\zeta)\cdot\lambda^{-1/2}(\xi'))\cdot\partial^{\alpha'}_{x'}p(x',\zeta)d\zeta\Bigg| \\
&\leq c_{\alpha',\beta,\beta'}\lambda(\xi)^{-n/4}\lambda(\xi')^{-n/4}\lambda(\xi)^{-\frac{1}{2}\varrho(|\beta|)}\lambda(\xi')^{-\frac{1}{2}\varrho(|\beta'|)}\cdot I,
\end{aligned} \tag{6.45}$$

where

$$I = \int_{\substack{|\xi-\zeta|\leq\lambda^{1/2}(\xi)\\ |\xi'-\zeta|\leq\lambda^{1/2}(\xi')}}\left|\partial^{\alpha'}_{x'}p(x',\zeta)\right|d\zeta.$$

Observe that by (4.18) and (2.8) for $|\sigma| \leq 1$

$$\begin{aligned}
\lambda(\xi+\lambda^{1/2}(\xi)\sigma) &\leq \lambda(\xi)+\lambda(\lambda^{1/2}(\xi)\sigma) \\
&\leq \lambda(\xi)+c(1+\lambda^{1/2}(\xi)\,|\sigma|) \\
&\leq c\lambda(\xi).
\end{aligned} \tag{6.46}$$

Hence using the substitution $\zeta = \xi + \lambda^{1/2}(\xi)\cdot\sigma$ we find by Cauchy-Schwarz inequality

$$\begin{aligned}
|I| &\leq \left(\int_{|\xi-\zeta|\leq\lambda^{1/2}(\xi)}\left|\partial^{\alpha'}_{x'}p(x',\zeta)\right|^2 d\zeta\right)^{1/2}\left(\int_{|\xi'-\zeta|\leq\lambda^{1/2}(\xi')} 1\ d\zeta\right)^{1/2} \\
&= \lambda^{n/4}(\xi)\left(\int_{|\sigma|\leq1}\left|\partial^{\alpha'}_{x'}p(x',\xi+\lambda^{1/2}(\xi)\cdot\sigma)\right|^2 d\sigma\right)^{1/2}\lambda^{n/4}(\xi')\left(\int_{|\sigma|\leq1}d\sigma\right)^{1/2} \\
&\leq c\lambda^{n/4}(\xi)\lambda^{n/4}(\xi')\lambda^{m}(\xi),
\end{aligned}$$

which together with (6.45) gives (6.42).

In order to prove (6.43) we need the following

Lemma 6.22. *Let $p \in S_0^{m,\lambda}$, $t \in \mathbb{R}$ and $\sigma \in \mathbb{R}^n$. Then*

$$\partial_x^\alpha \partial_\xi^\beta (p(x, \xi + t\lambda^{1/2}(\xi) \cdot \sigma)) = \sum_{\substack{|\gamma| \leq |\beta| \\ \gamma_1 \leq \gamma}} \psi_{\beta,\gamma,\gamma_1}(\xi)(\partial_x^\alpha \partial_\xi^\gamma p)(x, \xi + t\lambda^{1/2}(\xi) \cdot \sigma) \cdot (t\sigma)^{\gamma_1}, \tag{6.47}$$

where $\psi_{\beta,\gamma,\gamma_1} \in S_0^{0,\lambda}$.

Proof: Since also $\partial_x^\alpha p(x, \xi) \in \mathrm{S}_0^{m,\lambda}$ for all $\alpha \in \mathbb{N}_0^n$ as well, we may replace p by $\partial_x^\alpha p$ and assume $\alpha = 0$. With $\psi_{0,0,0} = 1$ there is nothing to prove for $\beta = 0$. Let $\tilde{\xi} = \xi + t\lambda^{1/2}(\xi) \cdot \sigma$. Then by induction

$$\begin{aligned} \partial_{\xi_i} \partial_\xi^\beta p(x, \tilde{\xi}) = \sum_{\substack{|\gamma| \leq |\beta| \\ \gamma_1 \leq \gamma}} & \Big\{ \partial_{\xi_i} \psi_{\beta,\gamma,\gamma_1}(\xi)(\partial_\xi^\gamma p)(x, \tilde{\xi}) \cdot (t\sigma)^{\gamma_1} \\ & + \psi_{\beta,\gamma,\gamma_1}(\xi)(\partial_\xi^{\gamma+\varepsilon_i} p)(x, \tilde{\xi}) \cdot (t\sigma)^{\gamma_1} \\ & + \sum_{k=1}^n \psi_{\beta,\gamma,\gamma_1}(\xi)(\partial_\xi^{\gamma+\varepsilon_k} p)(x, \tilde{\xi}) \cdot \partial_{\xi_i} \lambda^{1/2}(\xi) \cdot (t\sigma)^{\gamma_1+\varepsilon_k} \Big\}, \end{aligned}$$

which proves the lemma, since $\partial_{\xi_i} \lambda^{1/2} \in \mathrm{S}_0^{0,\lambda}$. □

We now finish the proof of Theorem 6.20.
Proof of (6.43): By the expansion formula (6.37) (replace $\varrho(\cdot)$ by $\frac{1}{2}\varrho(\cdot)$, which does not affect the proof) and (6.42) we know that

$$p_{F,L} - p_{F,0} - \sum_{|\alpha|=1} p_{F,\alpha} \in \mathrm{S}_0^{m-1,\lambda}.$$

Thus it is enough to prove

$$p_{F,\alpha} \in \mathrm{S}_0^{m-1,\lambda} \quad \text{for } |\alpha| = 1 \tag{6.48}$$

and

$$p_{F,0} - p \in \mathrm{S}_0^{m-1,\lambda}. \tag{6.49}$$

Let $|\alpha| = 1$. Then

$$\begin{aligned} \partial_\xi^\alpha F(\xi, \eta) &= \partial_\xi^\alpha \left(\lambda^{-n/4}(\xi) q((\eta - \xi)\lambda^{-1/2}(\xi)) \right) \\ = \lambda^{-n/4}(\xi) &\cdot \Big[-\frac{n}{4} q((\eta - \xi)\lambda^{-1/2}(\xi)) \cdot \lambda^{-1}(\xi) \partial_\xi^\alpha \lambda(\xi) \\ &+ \sum_{k=1}^n (\partial_k q)((\eta - \xi)\lambda^{-1/2}(\xi)) \cdot (\eta_k - \xi_k) \cdot \lambda^{-1/2}(\xi) \lambda^{1/2}(\xi) \partial_\xi^\alpha \lambda^{-1/2}(\xi) \\ &- (\partial^\alpha q)((\eta - \xi)\lambda^{-1/2}(\xi)) \cdot \lambda^{-1/2}(\xi) \Big] \end{aligned}$$

and consequently with $\sigma = (\eta - \xi)\lambda^{-1/2}(\xi)$

$$\begin{aligned}
p_{F,\alpha} &= D^\alpha_{x'}\partial^\alpha_\xi \, p_F(\xi, x', \xi')|_{\substack{x'=x \\ \xi'=\xi}} = \int_{\mathbb{R}^n} \partial^\alpha_\xi F(\xi,\eta) \cdot D^\alpha_x p(x,\eta) F(\xi,\eta) d\eta \\
&= -\frac{n}{4}\lambda^{-1}(\xi)\partial^\alpha_\xi \lambda(\xi) \cdot \int_{\mathbb{R}^n} q^2(\sigma) D^\alpha_x p(x, \xi + \lambda^{1/2}(\xi)\sigma) d\sigma \\
&+ \sum_{k=1}^n \lambda^{1/2}(\xi)\partial^\alpha_\xi \lambda^{-1/2}(\xi) \cdot \int_{\mathbb{R}^n} \sigma_k \partial_k q(\sigma) \cdot q(\sigma) D^\alpha_x p(x, \xi + \lambda^{1/2}(\xi)\sigma) d\sigma \\
&- \lambda^{-1/2}(\xi) \cdot \int_{\mathbb{R}^n} (\partial^\alpha q)(\sigma) q(\sigma) D^\alpha_x p(x, \xi + \lambda^{1/2}(\xi)\sigma) d\sigma \\
&= I_1 + I_2 + I_3.
\end{aligned}$$

We consider each term separately. Observe that $\int_{\mathbb{R}^n} q^2(\sigma) D^\alpha_x p(x, \xi + \lambda^{1/2}(\xi)\sigma) d\sigma$ is a symbol in $\mathrm{S}_0^{m,\lambda}$, since using Lemma 6.22 and (6.46)

$$\begin{aligned}
&\left| \partial^\delta_x \partial^\beta_\xi \int_{\mathbb{R}^n} q^2(\sigma) D^\alpha_x p(x, \xi + \lambda^{1/2}(\xi)\sigma) d\sigma \right| \\
&\leq \left| \int_{\mathbb{R}^n} q^2(\sigma) \partial^\delta_x \partial^\beta_\xi D^\alpha_x p(x, \xi + \lambda^{1/2}(\xi)\sigma) d\sigma \right| \\
&\leq \sum_{\substack{|\gamma| \leq |\beta| \\ \gamma_1 \leq \gamma}} |\psi_{\beta,\gamma,\gamma_1}| \int_{\mathbb{R}^n} q^2(\sigma) \left| \left(\partial^\delta_x \partial^\beta_\xi D^\alpha_x p \right) (x, \xi + \lambda^{1/2}(\xi)\sigma) \right| |\sigma^{\gamma_1}| \, d\sigma \\
&\leq c \int_{\mathbb{R}^n} q^2(\sigma) \lambda(\xi + \lambda^{1/2}(\xi)\sigma)^m d\sigma \leq c \int_{\mathbb{R}^n} q^2(\sigma) d\sigma \cdot \lambda^m(\xi) = c\lambda^m(\xi)
\end{aligned}$$

and $\lambda^{-1}\partial^\alpha_\xi \lambda \in \mathrm{S}_0^{-1,\lambda}$ gives $I_1 \in \mathrm{S}_0^{m-1,\lambda}$.
Analogously

$$\int_{\mathbb{R}^n} \sigma_k \partial_k q(\sigma) \cdot q(\sigma) D^\alpha_x p(x, \xi + \lambda^{1/2}(\xi)\sigma) d\sigma \in \mathrm{S}_0^{m,\lambda}$$

and thus by $\lambda^{1/2}\partial^\alpha_\xi \lambda^{-1/2} \in \mathrm{S}_0^{-1,\lambda}$ we have $I_2 \in \mathrm{S}_0^{m-1,\lambda}$.
Moreover concerning I_3 we have by Taylor's formula

$$\begin{aligned}
&\int_{\mathbb{R}^n} \partial^\alpha q(\sigma) q(\sigma) D^\alpha_x p(x, \xi + \lambda^{1/2}(\xi)\sigma) d\sigma \\
&= \int_{\mathbb{R}^n} \partial^\alpha q(\sigma) q(\sigma) d\sigma \cdot D^\alpha_x p(x,\xi) \\
&+ \lambda^{1/2}(\xi) \int_{\mathbb{R}^n} \partial^\alpha q(\sigma) q(\sigma) \sum_{k=1}^n \sigma_k \cdot \int_0^1 \left(\partial_{\xi_k} D^\alpha_x p \right) (x, \xi + \lambda^{1/2}(\xi) t\sigma) dt d\sigma.
\end{aligned}$$

By the symmetry of q the first term vanishes and we find for the derivatives of the second integral using again Lemma 6.22

$$\left| \partial^\delta_x \partial^\beta_\xi \int_{\mathbb{R}^n} \partial^\alpha q(\sigma) q(\sigma) \sigma_k \cdot \int_0^1 \left(\partial_{\xi_k} D^\alpha_x p \right) (x, \xi + \lambda^{1/2}(\xi) t\sigma) dt d\sigma \right|$$

$$\leq \int_{\mathbb{R}^n} |\partial^\alpha q(\sigma)q(\sigma)\sigma_k| \cdot \int_0^1 \left| \partial_x^\delta \partial_\xi^\beta \left(D_x^\alpha \partial_{\xi_k} p \right) (x, \xi + \lambda^{1/2}(\xi)t\sigma) \right| dt d\sigma$$
$$\leq c \int_{\mathbb{R}^n} |\partial^\alpha q(\sigma)q(\sigma)\sigma_k| \cdot \int_0^1 \lambda^{m-1}(\xi + \lambda^{1/2}(\xi)t\sigma) dt d\sigma \leq c \cdot \lambda^{m-1}(\xi).$$

Hence $\int_{\mathbb{R}^n} \partial^\alpha q(\sigma)q(\sigma)D_x^\alpha p(x, \xi + \lambda^{1/2}(\xi)\sigma)d\sigma$ is in $\mathrm{S}_0^{m-1/2,\lambda}$, which means $I_3 \in \mathrm{S}_0^{m-1,\lambda}$ and we have (6.48). Let us turn to (6.49). By Taylor's formula we find

$$p_{F,0}(x,\xi) = \int_{\mathbb{R}^n} q^2(\sigma)p(x, \xi + \lambda^{1/2}(\xi)\sigma)d\sigma$$
$$= \int_{\mathbb{R}^n} q^2(\sigma) \Big\{ p(x,\xi) + \sum_{k=1}^n \lambda^{1/2}(\xi) \cdot \sigma_k \cdot \partial_{\xi_k} p(x,\xi)$$
$$+ \int_0^1 (1-t) \sum_{|\gamma|=2} \frac{2}{\gamma!} \lambda(\xi) \sigma^\gamma (\partial_\xi^\gamma p)(x, \xi + t\lambda^{1/2}(\xi)\sigma)dt \Big\} d\sigma.$$

By the symmetry of q the integral over the first order term again vanishes and therefore

$$p_{F,0}(x,\xi) - p(x,\xi) = \sum_{|\gamma|=2} \frac{2}{\gamma!} \lambda(\xi) \int_{\mathbb{R}^n} \int_0^1 (1-t) q^2(\sigma) \sigma^\gamma (\partial_\xi^\gamma p)(x, \xi + t\lambda^{1/2}(\xi)\sigma) dt d\sigma.$$

Using again Lemma 6.22 and (6.46) we see as above that the integral defines a symbol in $\mathrm{S}_0^{m-2,\lambda}$, which gives (6.49) □

The next theorem summerizes the important properties of the Friedrichs symmetrization.

Theorem 6.23. *Assume $p \in S_0^{m,\lambda}$ is real-valued. Then $p_F(D_x, x', D_{x'})$ with domain $\mathcal{S}(\mathbb{R}^n, \mathbb{C})$ is a symmetric operator in $L^2(\mathbb{R}^n, \mathbb{C})$. If moreover $p(x,\xi)$ is non-negative, then $p_F(D_x, x', D_{x'})$ is non-negative.*

Proof: This is clear, because for u, $v \in \mathcal{S}(\mathbb{R}^n, \mathbb{C})$ and if $\mathcal{F}^{-1}_{\xi\to x}$ denotes the inverse Fourier transform

$$(p_F(D_x, x', D_{x'})u, v)_0$$
$$= \int_{\mathbb{R}^n} \mathcal{F}^{-1}_{\xi\to x} \left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-i(x',\xi)+i(x',\xi')} p_F(\xi, x', \xi') \hat{u}(\xi') đ\xi' dx' \right)(x) \overline{v(x)} dx$$
$$= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-i(x',\xi)+i(x',\xi')} \int_{\mathbb{R}^n} F(\xi,\eta)p(x',\eta)F(\xi',\eta)d\eta \, \hat{u}(\xi') đ\xi' dx' \, \overline{\hat{v}(\xi)} đ\xi$$
$$= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} p(x',\eta) \int_{\mathbb{R}^n} e^{i(x',\xi')} F(\xi',\eta)\hat{u}(\xi')đ\xi' \cdot \overline{\int_{\mathbb{R}^n} e^{i(x',\xi)} F(\xi,\eta)\hat{v}(\xi)đ\xi} d\eta dx'.$$

□

6.5 Application to generators of Feller semigroups

In this section we want to apply the results of the previous sections to pseudo differential operators with negative definite symbols. In particular we assume the symbols to be real-valued. As we have seen it is a natural condition to assume that the symbols are of class $S_\varrho^{2,\lambda}$ for some convenient $\lambda(\xi)$. To prove that a pseudo differential operator fulfills the assumptions of the Hille-Yosida theorem and therefore is the generator of an operator semigroup to most extent amounts to solve the equation

$$p(x,D)u + \tau u = f, \tag{6.50}$$

as we have seen in Chapter 4. We will solve this problem for elliptic elements in $S_\varrho^{m,\lambda}$. In order to apply modified Hilbert space methods we need again some estimates for the operator and the corresponding bilinear form. It turns out that by the symbolic calculus, once it was proven, these estimates are obtained in a considerably more elegant way.
As an application of the Friedrichs symmetrization we first prove the sharp Garding inequality which gives a first non-trivial lower bound for the corresponding sesquilinear form.

Theorem 6.24. *Let $p \in S_\varrho^{m,\lambda}$ be nonnegative. There is a $K \geq 0$ such that*

$$Re(p(x,D)u,u)_0 \geq -K \|u\|^2_{\frac{m-1}{2},\lambda} \qquad \text{for all } u \in \mathcal{S}(\mathbb{R}^n,\mathbb{C}).$$

Proof: By Theorem 6.20 we know that $p(x,D) - p_F(D_x,x',D_{x'})$ is of order $m-1$. Since $p(x,\xi) \geq 0$ we have by Theorem 6.23

$$\begin{aligned}
\mathrm{Re}(p(x,D)u,u)_0 &= \mathrm{Re}(p_F(D_x,x',D_{x'})u,u)_0 + \mathrm{Re}((p(x,D)-p_F(D_x,x',D_{x'}))u,u)_0 \\
&\geq \mathrm{Re}\left(\lambda^{-\frac{m-1}{2}}(D)(p(x,D)-p_F(D_x,x',D_{x'}))u, \lambda^{\frac{m-1}{2}}(D)u\right)_0 \\
&\geq -K\|u\|^2_{\frac{m-1}{2},\lambda}.
\end{aligned}$$

□

We are interested in further bounds for the sesquilinear form, in particular in the elliptic case.

Theorem 6.25. *Let $p \in S_\varrho^{m,\lambda}$ be real-valued, $m > 0$. Then*

$$|(p(x,D)u,v)_0| \leq c\|u\|_{m/2,\lambda} \cdot \|v\|_{m/2,\lambda}, \qquad u,v \in \mathcal{S}(\mathbb{R}^n,\mathbb{C}) \tag{6.51}$$

and the sesquilinear form extends continuously to $H^{m/2,\lambda}(\mathbb{R}^n,\mathbb{C})$. If moreover

$$p(x,\xi) \geq \delta\lambda^m(\xi), \qquad |\xi| > R, \tag{6.52}$$

for some $\delta > 0$ and some $R > 0$, then the Garding inequality

$$Re(p(x,D)u,u)_0 \geq \frac{\delta}{2}\|u\|^2_{\frac{m}{2},\lambda} - c\|u\|^2_{\frac{m-1}{2},\lambda}, \qquad u \in H^{m/2,\lambda}(\mathbb{R}^n,\mathbb{C}), \tag{6.53}$$

holds.

Proof: We know that

$$|(p(x,D)u,u)_0| = \left|(\lambda^{-\frac{m}{2}}(D)p(x,D)u, \lambda^{\frac{m}{2}}(D)u)_0\right| \le c\,\|u\|_{m/2,\lambda} \cdot \|v\|_{m/2,\lambda}\,,$$

since $\lambda^{-m/2}(D) \circ p(x,D)$ is of order $\frac{m}{2}$.
Now assume (6.52). Let $p_\tau(x,\xi) = p(x,\xi) + \tau$. Then for τ sufficiently large

$$p_\tau(x,\xi) \ge \delta\lambda^m(\xi)$$

holds for all $\xi \in \mathbb{R}^n$. We put $q(x,\xi) = p_\tau(x,\xi) - \delta\lambda^m(\xi) \ge 0$. Theorem 6.24 implies

$$\operatorname{Re}(p(x,D)u,u)_0 - \delta\,\|u\|^2_{m/2,\lambda} + \tau\,\|u\|^2_0 = \operatorname{Re}(q(x,D)u,u)_0 \ge -K\,\|u\|^2_{\frac{m-1}{2},\lambda}\,,$$

which gives (6.53) by (4.11). □

Let us turn next to estimates for the operator itself. The operator $p(x,D) \in \Psi^{m,\lambda}_0$ is a continuous operator between the Sobolev spaces $H^{s,\lambda}(\mathbb{R}^n,\mathbb{C})$, see Theorem 6.14, i.e. $\|p(x,D)u\|_{s,\lambda} \le c\,\|u\|_{s+m,\lambda}$. Again, if moreover (6.52) holds, we even have a converse inequality.

Theorem 6.26. *Let $p \in S^{m,\lambda}_\varrho$ be real-valued and assume the ellipticity condition (6.52). Then for $s > -m$*

$$\frac{\delta^2}{2}\,\|u\|^2_{s+m,\lambda} \le \|p(x,D)u\|^2_{s,\lambda} + c\,\|u\|^2_{s+m-\frac{1}{2},\lambda}\,. \tag{6.54}$$

Proof: Let $q_s(x,\xi) = p(x,\xi)^2\lambda^{2s}(\xi) \ge \delta^2\lambda^{2(m+s)}(\xi)$ for $|\xi|$ large. By Corollary 6.18 we know that the highest order term in the expansion of the symbol of $p^*(x,\xi)$ is given by $\overline{p(x,\xi)} = p(x,\xi)$. Thus

$$\begin{aligned}\|p(x,D)u\|^2_{s,\lambda} &= (\lambda^s(D)p(x,D)u, \lambda^s(D)p(x,D)u)_0 \\ &= (p^*(x,D)\lambda^{2s}(D)p(x,D)u,u)_0 = \operatorname{Re}(q_s(x,D)u,u)_0 + \operatorname{Re}(q(x,D)u,u)_0,\end{aligned}$$

where $q(x,D) \in \mathrm{S}^{2(m+s)-1,\lambda}_0$. Hence Theorem 6.25 implies

$$\|p(x,D)u\|^2_{s,\lambda} \ge \delta^2\,\|u\|^2_{m+s,\lambda} - c\,\|u\|^2_{m+s-\frac{1}{2},\lambda} - c'\,\|u\|^2_{m+s-\frac{1}{2},\lambda}\,.$$

□

The proof of regularity results for solutions of (6.50) again involves commutators for the Friedrichs mollifiers as defined in (4.22). Obviously J_ε is a pseudo differential operator with symbol $\hat{\jmath}(\varepsilon\xi)$ in $\mathrm{S}^{0,\lambda}_\varrho$ and the constants $c_{\alpha,\beta}$ in the corresponding estimate (6.10) are uniformly bounded for $0 < \varepsilon \le 1$, cf. [54], Lemma 1.6.3. Let $p \in \mathrm{S}^{m,\lambda}_\varrho$. We consider the commutator

$$[p(x,D), J_\varepsilon] = p(x,D)J_\varepsilon - J_\varepsilon p(x,D).$$

Recall that the commutator is described by the difference of the double symbols $p(x,\xi)\cdot\hat{\jmath}(\varepsilon\xi')$ and $\hat{\jmath}(\varepsilon\xi)\cdot p(x',\xi')$. Therefore it is now obvious to see that the commutator has an order reducing effect, since the highest order terms in the expansion series (6.37) cancel, and $[p(x,D),J_\varepsilon]$ is an operator of order $m-1$. Moreover the remaining terms of the expansion are controlled uniformly with repect to ε, see Remark 6.15 and Remark 6.19. Therefore we get

Proposition 6.27. *Let $p \in S_\varrho^{m,\lambda}$ and $s \in \mathbb{R}$. There is a constant $c \geq 0$ not depending on $0 < \varepsilon \leq 1$ such that*

$$\|[p(x,D), J_\varepsilon]u\|_{s,\lambda} \leq c\,\|u\|_{m+s-1,\lambda}\,.$$

We summerize the results obtained so far and solve equation (6.50).

Theorem 6.28. *Let $p \in S_\varrho^{m,\lambda}$, $m \geq 2$, be a real-valued symbol, $s \geq 0$ and assume that (6.52) holds. If $\tau > 0$ is sufficiently large, then for $f \in H^{s,\lambda}(\mathbb{R}^n, \mathbb{C})$ there is a unique solution $u \in H^{s+m,\lambda}(\mathbb{R}^n, \mathbb{C})$ of the equation*

$$p(x,D)u + \tau u = f.$$

Proof: By Theorem 6.25 we know that

$$(u,v) \mapsto ((p(x,D)+\tau)u, v)_0$$

is a continuous coercive form on $H^{m/2,\lambda}(\mathbb{R}^n, \mathbb{C})$ for τ large enough. Thus, see [20], Theo. I.14.1, there is a unique weak solution $u \in H^{m/2,\lambda}(\mathbb{R}^n, \mathbb{C})$ of

$$((p(x,D)+\tau)u, v)_0 = (f,v)_0 \qquad \text{for all } v \in H^{m/2,\lambda}(\mathbb{R}^n, \mathbb{C})$$

and the proof is complete, if we show that $u \in H^{s+m,\lambda}(\mathbb{R}^n, \mathbb{C})$. Let $u_\varepsilon = J_\varepsilon u$. Then $u_\varepsilon \in H^{t+m,\lambda}(\mathbb{R}^n, \mathbb{C})$ for all $t \leq s$, $0 < \varepsilon \leq 1$ and by Therorem 6.26 and Proposition 6.27 we have

$$\begin{aligned}
\|u_\varepsilon\|_{t+m,\lambda} &\leq c\,\|p(x,D)J_\varepsilon u\|_{t,\lambda} + c\,\|J_\varepsilon u\|_{t+m-\frac{1}{2},\lambda} \\
&\leq c\,\|J_\varepsilon(p(x,D)+\tau)u\|_{t,\lambda} + c\,\|J_\varepsilon u\|_{t,\lambda} + c\,\|[p(x,D),J_\varepsilon]u\|_{t,\lambda} + c\,\|J_\varepsilon u\|_{t+m-\frac{1}{2},\lambda} \\
&\leq c\,\|J_\varepsilon f\|_{t,\lambda} + c\,\|u\|_{t,\lambda} + c\,\|u\|_{t+m-1,\lambda} + c\,\|u\|_{t+m-\frac{1}{2},\lambda} \\
&\leq c\,\|f\|_{s,\lambda} + c\,\|u\|_{t+m-\frac{1}{2},\lambda}\,.
\end{aligned}$$

So $u \in H^{t+m-\frac{1}{2},\lambda}(\mathbb{R}^n, \mathbb{C})$ implies that $(u_\varepsilon)_{0<\varepsilon\leq 1}$ is bounded in $H^{t+m,\lambda}(\mathbb{R}^n, \mathbb{C})$. Since $u_\varepsilon \to u$ in $H^{t+m-\frac{1}{2},\lambda}(\mathbb{R}^n, \mathbb{C})$ as $\varepsilon \to 0$, this implies $u \in H^{t+m,\lambda}(\mathbb{R}^n, \mathbb{C})$. A recursive application of this conclusion starting with $t = \frac{1-m}{2}$ proves the theorem. □

We finally state our result about generators of Feller semigroups. Recall that we have assumed a lower bound (6.4) on the reference function, that is for some $r > 0$

$$\lambda(\xi) \geq c\,|\xi|^{r/2} \tag{6.55}$$

for some $c > 0$ and $|\xi|$ large. Under this condition for $s > \frac{n}{r}$ the dense and continuous embedding, see Proposition 4.1,

$$H^{s,\lambda}(\mathbb{R}^n) \hookrightarrow C_\infty(\mathbb{R}^n)$$

holds. Now we have

Theorem 6.29. *Assume that (6.55) holds. If $p(x,\xi)$ is a negative definite symbol of class $S_\varrho^{2,\lambda}$ and moreover*

$$p(x,\xi) \geq \delta\lambda^2(\xi)$$

for some $\delta > 0$ and $|\xi|$ large, then $-p(x,D)$ defined on $C_0^\infty(\mathbb{R}^n)$ is closable in $C_\infty(\mathbb{R}^n)$ and the closure generates a Feller semigroup.

Proof: It is enough to repeat the arguments in the proof of Theorem 4.13. Recall that for a continuous negative definite symbol the operator $p(x,D)$ preserves real-valued functions. Choose $s > \frac{n}{r}$. Then the operator $A = -p(x,D) : H^{s+2,\lambda}(\mathbb{R}^n) \to H^{s,\lambda}(\mathbb{R}^n) \subset C_\infty(\mathbb{R}^n)$ is a densely defined operator in $C_\infty(\mathbb{R}^n)$ with domain $H^{s+2,\lambda}(\mathbb{R}^n)$ und thus A fulfills condition (i) of Theorem 4.7. Since $C_0^\infty(\mathbb{R}^n)$ is dense in $H^{s+2,\lambda}(\mathbb{R}^n)$, we see again from the continuity and the regularity estimate (6.35), (6.54) and the Sobolev embedding that $C_0^\infty(A)$ is a core for A. Moreover by Theorem 2.16 A satisfies the positive maximum principle on $C_0^\infty(\mathbb{R}^n)$ and therefore also on $H^{s+2,\lambda}(\mathbb{R}^n)$, see Proposition 2.20. This is (ii) of Theorem 4.7 and finally (iii) is the claim of Theorem 6.28. □

6.6 Perturbation results

In section 6.2 we used the decomposition of a continuous negative definite symbol into a part with Lévy-measures supported in a bounded set and a remainder part. The first part was studied using the symbolic calculus. In this section we consider the remainder symbol, which has a Lévy-kernel consisting of finite measures. In particular we are interested in the perturbation effect of this remainder part in terms of estimates in L^∞- and L^2-norms.

Let us assume that $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a continuous negative definite symbol that satisfies

$$p(x,\xi) \le c(1+|\xi|^2). \tag{6.56}$$

For simplicity we assume that $p(x,0) = 0$. This is no restriction since an operator of the type $\varphi \mapsto c(\cdot)\cdot\varphi$ with a bounded continuous function $c(x)$ is a bounded operator in $C_\infty(\mathbb{R}^n)$ as well as in $L^2(\mathbb{R}^n)$.
We decompose the symbol according to Proposition 3.11 and Theorem 3.12 by using an even cut-off function $\theta \in C_0^\infty(\mathbb{R}^n)$, $0 \le \theta \le 1$, $\theta(0) = 1$:

$$p(x,\xi) = p_1^\theta(x,\xi) + p_2^\theta(x,\xi) = p_1(x,\xi) + p_2(x,\xi). \tag{6.57}$$

In particular

$$p_2(x,\xi) = \int_{\mathbb{R}^n} (p(x,\xi) - p(x,\xi+\eta) + p(x,\eta))\,\hat\theta(\eta)\,đ\eta.$$

Then $\mu_2(x,dy) = (1-\theta(y))\,\mu(x,dy)$ is the Lévy-kernel of $p_2(x,\xi)$, where $\mu(x,dy)$ is the Lévy-kernel for $p(x,\xi)$, and we have

$$-p_2(x,D)\varphi(x) = \int_{\mathbb{R}^n\setminus\{0\}} (\varphi(x+y)-\varphi(x))\,\mu_2(x,dy), \qquad \varphi \in C_0^\infty(\mathbb{R}^n). \tag{6.58}$$

Moreover

$$\mu_2(x,\mathbb{R}^n\setminus\{0\}) \le c_\theta \int_{\mathbb{R}^n} p(x,\xi)\,\nu(d\xi) \le c\int_{\mathbb{R}^n} (1+|\xi|^2)\,\nu(d\xi) < \infty \tag{6.59}$$

by Theorem 3.12 and (6.56). The central question of this section is, whether the original operator $p(x,D)$ which we get from $p_1(x,D)$ by the perturbation $p_2(x,D)$ has similar properties as $p_1(x,D)$. For the readers' convenience we recall a standard perturbation result for generators of contraction semigroups (see for example [17], 1.7.1).

Theorem 6.30. *Let $(A, D(A))$ be the generator of a strongly continuous contraction semigroup on a Banach space $(X, \|\cdot\|)$ and $(B, D(B))$ a linear dissipative operator in X such that $D(A) \subset D(B)$ and*

$$\|Bu\| \leq \alpha \|Au\| + \beta \|u\|, \qquad u \in D(A),$$

for some $0 \leq \alpha < 1$ and $\beta \geq 0$. Then $(A+B, D(A))$ generates a strongly continuous contraction semigroup.

The theorem in particular applies to bounded perturbations, that is if B is a bounded operator, then $B - \|B\| \cdot \mathrm{Id}$ is a dissipative operator. Therefore $A + B$ generates a strongly continuous semigroup, but not necessarily a contraction semigroup.
In the case of Feller semigroups however the submarkovian property is characterized by the positive maximum principle. Therefore if A is the generator of a Feller semigroup, then for any bounded operator B such that $A + B$ satisfies the positive maximum principle, also $A + B$ is the generator of a Feller semigroup.

In order to see that $p_2(x, D)$ defines a perturbation of this type first note that by (6.58) and (6.59) $p_2(x, D)$ has an extension to the bounded Borel measurable functions $B(\mathbb{R}^n)$

$$p_2(x, D) : B(\mathbb{R}^n) \to B(\mathbb{R}^n)$$

which is continuous with a bound given by $2 \cdot \sup_{x\in\mathbb{R}^n} \|\mu_2(x, \cdot)\|_\infty$. In order to apply Theorem 6.30 to the perturbation $-p_2(x, D)$ in the case of Feller semigroups we have to show that $C_\infty(\mathbb{R}^n)$ is invariant under $p_2(x, D)$. In general this is not true, since the non-local character of $p_2(x, D)$ may destroy the behaviour at infinity. A reasonable condition to control the non-locality is the tightness of the Lévy-measures $\mu_2(x, dy)$. We give a complete characterization of the tightness in terms of the symbol.

Theorem 6.31. *Let $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite symbol such that $p(x, \xi) \leq c(1 + |\xi|^2)$ and with Lévy-Khinchin representation*

$$p(x, \xi) = q(x, \xi) + c(x) + \int_{\mathbb{R}^n\setminus\{0\}} (1 - \cos(y, \xi))\, \mu(x, dy).$$

Then the following are equivalent:

(i) for every $\varepsilon > 0$ there is a ball $B_R(0) \subset \mathbb{R}^n$ such that $\sup\limits_{x\in\mathbb{R}^n} \mu(x, B_R(0)^{\complement}) \leq \varepsilon$,

(ii) $\sup\limits_{x\in\mathbb{R}^n} (p(x, \xi) - p(x, 0)) \to 0 \quad$ as $\xi \to 0$.

In this case $p(x, D)$ maps $C_0^\infty(\mathbb{R}^n)$ into $C_\infty(\mathbb{R}^n)$.

Note that typically a condition on the symbol for small ξ, here the equicontinuity at $\xi = 0$, implies properties of the Lévy-measures are infinity.

Proof: Note that by the assumptions $q(x, \xi) + c(x) \leq c(1 + |\xi|^2)$ therefore $c(x)$ and the coefficients of $q(x, \cdot)$ are bounded. Thus they are equicontinuous at $\xi = 0$ and we may assume that

$$p(x, \xi) = \int_{\mathbb{R}^n\setminus\{0\}} (1 - \cos(y, \xi))\, \mu(x, dy).$$

Assume that (ii) holds true. Let $\varphi \in C_0^\infty(\mathbb{R}^n)$, $0 \le \varphi \le 1$, $\varphi(0) = 1$, $\operatorname{supp}\varphi \subset B_1(0)$ and $\varphi_R(x) = \varphi(x/R)$, $R \ge 1$. Then

$$
\begin{aligned}
(6.60) \qquad \mu(x, B_R(0)^{\complement}) &\le \int_{\mathbb{R}^n\setminus\{0\}} (\varphi_R(0) - \varphi_R(y))\,\mu(x,dy) \\
&= \int_{\mathbb{R}^n\setminus\{0\}} \int_{\mathbb{R}^n} (1-\cos(y,\xi))\hat{\varphi_R}(\xi)\,đ\xi\,\mu(x,dy) \le \int_{\mathbb{R}^n} p(x,\xi)\,|\hat{\varphi_R}(\xi)|\,đ\xi \\
&\le \int_{|\xi|\le\frac{1}{\sqrt{R}}} p(x,\xi)R^n\,|\hat{\varphi}(R\xi)|\,đ\xi + c\int_{|\xi|>\frac{1}{\sqrt{R}}} (1+|\xi|)^2 R^n (R\xi)^{-(n+3)}\,đ\xi \\
&\le c \sup_{|\xi|\le\frac{1}{\sqrt{R}}} p(x,\xi)\cdot\|\hat{\varphi}\|_{L^1} + cR^{-3/2} \to 0 \text{ as } R\to\infty
\end{aligned}
$$

uniformly with respect to x, i.e. (i) holds true.
Conversely, if (ii) is not true, then there is a sequence $(\xi_k)_{k\in\mathbb{N}}$, $\xi_k \to 0$ such that

$$\sup_{x\in\mathbb{R}^n} p(x,\xi_k) = \eta > 0$$

and we can choose $x_k \in \mathbb{R}^n$, $k\in\mathbb{N}$, such that $p(x_k,\xi_k) > \eta/2$. Then for a compact neighbourhood $K \subset \mathbb{R}^n$ of the origin and any $x,\xi\in\mathbb{R}^n$

$$
\begin{aligned}
\mu(x,K^{\complement}) &\ge \frac{1}{2}\int_{\complement K} (1-\cos(y,\xi))\,\mu(x,dy) \\
(6.61) \qquad &= \frac{1}{2}p(x,\xi) - \frac{1}{2}\int_K (1-\cos(y,\xi))\,\mu(x,dy).
\end{aligned}
$$

Let ν be the measure defined in Lemma 2.15 and

$$A = \sup_{x\in\mathbb{R}^n} \int_{\mathbb{R}^n} p(x,\xi)\,\nu(d\xi) \le c\int_{\mathbb{R}^n} (1+|\xi|^2)\,\nu(d\xi) < \infty.$$

Choose $a > 0$ such that $a < \frac{\eta}{4A}$. There is a $k_0 = k_0(K)\in\mathbb{N}$ such that for $k \ge k_0$

$$1-\cos(y,\xi_k) \le a\Big(1-\frac{1}{1+|y|^2}\Big) \qquad \text{for all } y\in K$$

and therefore for all $k\ge k_0$

$$
\begin{aligned}
&\int_K (1-\cos(y,\xi_k))\,\mu(x,dy) \le a\int_{\mathbb{R}^n\setminus\{0\}} \Big(1-\frac{1}{1+|y|^2}\Big)\,\mu(x,dy) \\
&= a\int_{\mathbb{R}^n\setminus\{0\}}\int_{\mathbb{R}^n} (1-\cos(y,\xi))\,\nu(d\xi)\,\mu(x,dy) = a\int_{\mathbb{R}^n} p(x,\xi)\,\nu(d\xi) \le a\cdot A < \frac{\eta}{4}.
\end{aligned}
$$

Hence by (6.61) for all $k\ge k_0$

$$\mu(x_k,K^{\complement}) \ge \frac{1}{2}p(x_k,\xi_k) - \frac{1}{2}\cdot\frac{\eta}{4} \ge \frac{\eta}{8} > 0,$$

which contradicts (i).
Finally, if (i) or (ii) is satisfied and $\varphi \in C_0^\infty(\mathbb{R}^n)$, then $p(x,D)\varphi$ is continuous and we have by the Lévy–Khinchin representation (2.19) of $-p(x,D)$ and (6.60) for $x \notin \operatorname{supp}\varphi$

$$|p(x,D)\varphi(x)| = \left| \int_{\mathbb{R}^n\setminus\{0\}} \varphi(x+y)\,\mu(x,dy) \right| \leq \|\varphi\|_\infty \cdot \mu(x, \operatorname{supp}\varphi(x+\cdot)) \to 0$$

as $|x| \to \infty$, i.e. $p(x,D)\varphi \in C_\infty(\mathbb{R}^n)$. □

As $C_0^\infty(\mathbb{R}^n)$ is dense in $C_\infty(\mathbb{R}^n)$ the result implies in particular in the situation considered above

Corollary 6.32. *Let p_2 be as above and assume*

$$\sup_{x\in\mathbb{R}^n} p_2(x,\xi) \to 0 \text{ as } \xi \to 0. \tag{6.62}$$

Then $p_2(x,D)$ maps $C_\infty(\mathbb{R}^n)$ continuously into itself.

Observe that (6.62) is determined directly by the original symbol $p(x,\xi)$, because p_1 in (6.57) always satisfies $\sup_{x\in\mathbb{R}^n}(p_1(x,\xi) - p_1(x,0)) \to 0$ as $\xi \to 0$. This can be seen for example using Theorem 6.31, since the Lévy-measures of p_1 are supported in a bounded set. Thus (6.62) is equivalent to the condition

$$\sup_{x\in\mathbb{R}^n} (p(x,\xi) - p(x,0)) \to 0 \text{ as } \xi \to 0. \tag{6.63}$$

We combine the results with the perturbation argument of Theorem 6.30 and the subsequent remark. Note that (6.58) shows that $-p_2(x,D)$ obviously satisfies the positive maximum principle on $C_\infty(\mathbb{R}^n)$. Thus we have proven

Theorem 6.33. *Let $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite symbol that satisfies (6.56) and (6.63) with decomposition (6.57). Assume that $-p_1(x,D)$ extends to the generator of a Feller semigroup. Then $-p(x,D)$ has the same property.*

Denote by $\lambda(\xi) = (1+\psi(\xi))^{1/2}$ the reference function for the symbolic calculus as in the previous sections, in particular assume (6.4). Then by Theorem 6.29 the assumption that $-p_1(x,D)$ generates a Feller semigroup is satisfied if $p_1(x,\xi)$ is an elliptic symbol in $\mathrm{S}_\varrho^{2,\lambda}$. Moreover we can prove

Corollary 6.34. *Let $\lambda(\xi)$be as above and $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite symbol that satisfies $p(x,\xi) \leq c(1+|\xi|^2)$,*

$$\sup_{x\in\mathbb{R}^n} (p(x,\xi) - p(x,0)) \to 0 \text{ as } \xi \to 0$$

and for some $\delta > 0$

$$p(x,\xi) \geq \delta\,\lambda^2(\xi) \qquad \text{for all } |\xi| \geq 1.$$

Let $\theta \in C_0^\infty(\mathbb{R}^n)$ be an even function such that $0 \leq \theta \leq 1$ and $\theta(0) = 1$. If the mollified symbol

$$(x,\xi) \mapsto \int_{\mathbb{R}^n} p(x,\eta)\hat{\theta}(\xi - \eta)\, đ\eta$$

is in $S_\varrho^{2,\lambda}$, then $-p(x,D)$ has an extension that generates a Feller semigroup.

Proof: We decompose p as in (6.57), see Proposition 3.11,

$$p(x,\xi) = \int_{\mathbb{R}^n} p(x,\eta)\hat{\theta}(\xi-\eta)\, đ\eta - \int_{\mathbb{R}^n} p(x,\eta)\hat{\theta}(\eta)\, đ\eta + p_2(x,\xi).$$

The assumptions yield that $x \mapsto \int_{\mathbb{R}^n} p(x,\eta)\hat{\theta}(\eta)\, đ\eta$ is in $C_b^\infty(\mathbb{R}^n)$. Therefore the symbol $\int_{\mathbb{R}^n} p(x,\eta)(\hat{\theta}(\xi-\eta) - \hat{\theta}(\eta))\, đ\eta$ is in $\mathrm{S}_\varrho^{2,\lambda}$ and, since $p_2(x,\xi)$ is bounded, it is bounded from below by $c\,\lambda^2(\xi)$ for $|\xi|$ large and suitable $c > 0$. Thus it satisfies the assumptions of Theorem 6.29 and the corresponding pseudo differential operator has an extension that generates a Feller semigroup. Moreover the symbol $p_2(x,\xi)$ defines a bounded operator on $C_\infty(\mathbb{R}^n)$ and we conclude as above. □

Finally we turn to L^2-estimates and we study the question whether $p_2(x,D)$ is also a perturbation in the L^2-framework.
For this purpose consider the situation of Example 1 in section 6.2, i.e. let $\tilde{\mu}$ be a symmetric Lévy-measure on $\mathbb{R}^n \setminus \{0\}$ and define the continuous negative definite reference function

$$\psi(\xi) = \int_{0<|y|\le 1} (1 - \cos(y,\xi))\, \tilde{\mu}(dy).$$

Let $\lambda(\xi) = (1 + \psi(\xi))^{1/2}$ and consider for $f : \mathbb{R}^n \times \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^+$ the symbol

$$p(x,\xi) = \int_{\mathbb{R}^n\setminus\{0\}} (1 - \cos(y,\xi))\, f(x,y)\, \tilde{\mu}(dy), \tag{6.64}$$

where we may assume that f is even with respect to the second variable. The symbol is decomposed in

$$p_1(x,\xi) = \int_{0<|y|\le 1} (1 - \cos(y,\xi))\, f(x,y)\, \tilde{\mu}(dy) \tag{6.65}$$

and

$$p_2(x,\xi) = \int_{|y|>1} (1 - \cos(y,\xi))\, f(x,y)\, \tilde{\mu}(dy). \tag{6.66}$$

We have seen that $p_1 \in \mathrm{S}_\varrho^{2,\lambda}$ provided f has bounded derivatives of all order with respect to x. Moreover, if $f(x,y) \ge \delta > 0$ we also have the ellipticity bound

$$p_1(x,\xi) \ge \delta \int_{0<|y|\le 1} (1 - \cos(y,\xi))\, \tilde{\mu}(dy) = \delta\psi(\xi),$$

i.e.

$$p_1(x,\xi) + \tau \ge \delta\lambda^2(\xi)$$

for τ sufficiently large. The Theorems 6.14, 6.25, 6.26 and 6.28 yield the following L^2-results for such symbols.

Theorem 6.35. *Let $p_1 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite symbol in $S_\varrho^{2,\lambda}$ and assume that for some $\delta \ge 0$ and τ sufficiently large*

$$p_1(x,\xi) + \tau \ge \delta\,\lambda^2(\xi).$$

Then for κ sufficiently large the bilinear form $((p_1(x,D)+\kappa)u,v)_0$, $u,v \in C_0^\infty(\mathbb{R}^n)$ has a continuous extension to a coercive form on $H^{1,\lambda}(\mathbb{R}^n)$.
Moreover $p_1(x,D)$ is closable in $L^2(\mathbb{R}^n)$, the closure is given by the continuous extension

$$p_1(x,D) : H^{2,\lambda}(\mathbb{R}^n) \to L^2(\mathbb{R}^n),$$

the estimates

$$c_1 \|u\|_{2,\lambda} \le \|p(x,D)u\|_{L^2} + \|u\|_{L^2} \le c_2 \|u\|_{2,\lambda}, \qquad c_1, c_2 > 0,$$

hold and for κ sufficiently large the equation

$$(p_1(x,D)+\kappa)u = f$$

has a unique solution $u \in H^{2,\lambda}(\mathbb{R}^n)$ for every $f \in L^2(\mathbb{R}^n)$. In particular $-(p_1(x,D)+\kappa)$with domain $H^{2,\lambda}(\mathbb{R}^n)$ is the generator of a strongly continuous contraction semigroup on $L^2(\mathbb{R}^n)$.

The result about the semigroup follows from the Hille-Yosida theorem 4.6, recall that the coercivity of the bilinear form implies dissipativity for the operator $-(p_1(x,D)+\kappa)$.

We now have

Proposition 6.36. *Define $p_2(x,\xi)$ as in (6.66) and assume that $f(x,y) \le M$ for some $M \ge 0$. Then $p_2(x,D)$ is bounded in $L^2(\mathbb{R}^n)$:*

$$\|p_2(x,D)u\|_0 \le c\,\|u\|_0 . \tag{6.67}$$

Proof: Let $u \in C_0^\infty(\mathbb{R}^n)$. Since the Lévy-kernel of p_2 consists of finite measures we have

$$\begin{aligned} -p_2(x,D)u(x) &= \int_{|y|>1} (u(x+y)-u(x))f(x,y)\,\tilde{\mu}(dy) \\ &= \int_{|y|>1} u(x+y)f(x,y)\,\tilde{\mu}(dy) - \left(\int_{|y|>1} f(x,y)\,\tilde{\mu}(dy)\right)\cdot u(x) \end{aligned}$$

and it is enough to prove continuity in $L^2(\mathbb{R}^n)$ for the first term. Let $v \in L^2(\mathbb{R}^n)$.

$$\begin{aligned} \left|\int_{\mathbb{R}^n}\int_{|y|>1} u(x+y) \quad f(x,y)\,\tilde{\mu}(dy)\cdot v(x)\,dx\right| &\le M\int_{|y|>1}\int_{\mathbb{R}^n} |u(x+y)|\cdot|v(x)|\,dx\,\tilde{\mu}(dy) \\ &\le M\int_{|y|>1}\left(\int_{\mathbb{R}^n}|u(x+y)|^2\,dx\right)^{1/2}\cdot\left(\int_{\mathbb{R}^n}|v(x)|^2\,dx\right)^{1/2}\tilde{\mu}(dy) \\ &= M\cdot\tilde{\mu}(\{|y|>1\})\cdot\|u\|_0\cdot\|v\|_0 . \end{aligned}$$

Dividing by $\|v\|_0$ and taking the supremum over all v with $\|v\|_0 = 1$ gives the result □

Therefore $p_2(x,D)$ is a small perturbation of operators as considered in Theorem 6.35 and we obtain

Theorem 6.37. *Assume that p_1 is as in Theorem 6.35 and p_2 is as in Proposition 6.36 and let $p = p_1 + p_2$.*
In particular this holds true if p is given by (6.64) with $\sup_{x,y} \left|\partial_x^\beta f(x,y)\right| < \infty$ for all $\beta \in \mathbb{N}_0^n$ and $f(x,y) \geq \delta > 0$.
Then the following estimates hold

$$\begin{aligned} \|p(x,D)u\|_0 &\leq c\,\|u\|_{2,\lambda}\,, \\ \|u\|_{2,\lambda} &\leq c(\|p(x,D)u\|_0 + \|u\|_0), \\ |(p(x,D)u,v)_0| &\leq c\,\|u\|_{1,\lambda}\cdot\|v\|_{1,\lambda}\,, \\ (p(x,D)u,u)_0 &\geq c_1\,\|u\|_{1,\lambda}^2 - c_2\,\|u\|_0^2 \end{aligned}$$

for $u, v \in C_0^\infty(\mathbb{R}^n)$. Moreover the continuous extension of $-p(x,D)$ to $H^{2,\lambda}(\mathbb{R}^n)$ is the generator of a strongly continuous semigroup in $L^2(\mathbb{R}^n)$.

Proof: This follows immediately from the corresponding estimates for $p_1(x,D)$, the boundedness of $p_2(x,D)$ in $L^2(\mathbb{R}^n)$, the continuous embeddings $H^{2,\lambda}(\mathbb{R}^n) \hookrightarrow H^{1,\lambda}(\mathbb{R}^n) \hookrightarrow L^2(\mathbb{R}^n)$, and (4.11). $\square$

Note that for symbols p defined as in (6.64) with continuous bounded f the operator $p_2(x,D)$ is also a bounded operator in $C_\infty(\mathbb{R}^n)$. Thus if $-p_1(x,D)$ generates a Feller semigroup, by the perturbation result Theorem 6.30 the operator $-(p_1(x,D) + p_2(x,D))$ also generates a Feller semigroup.

Chapter 7

Operators of variable order

7.1 Statement of results

In the previous chapter a symbolic calculus was developed which is suitable for continuous negative definite symbols. As an application we have shown that in an elliptic situation the calculus again yields results concerning the generation of Feller semigroups. The purpose of the chapter is to show that the symbolic calculus also applies in an explicitly non-elliptic situation, that is in the case of operators of variable order. Recently several investigations were made in the case of the best known example of this type, the generator of the so-called stable-like process which is given by the symbol

$$p(x,\xi) = |\xi|^{\alpha(x)} \tag{7.1}$$

or to avoid problems with differentiability

$$p(x,\xi) = (1+|\xi|^2)^{\frac{1}{2}\alpha(x)}, \tag{7.2}$$

where $0 < \alpha(x) \leq 2$. For fixed x the operator $-p(x,D)$ with symbol (7.1) coincides with the generator of a symmetric $\alpha(x)$-stable process, but the order varies with x. Note that in particular for fixed x the symbol is a negative definite function.
In the one-dimensional situation Bass [1] proved well-posedness of the corresponding martingale problem under weak assumptions on $\alpha(x)$. In the higher dimensional case the process corresponding to (7.1) was constructed by Tsuchiya [86] as the solution of a stochastic differential equation. Symbols as in (7.2) are contained in the Hörmander classes $S^m_{\varrho,\delta}$ and were studied by the symbolic calculus of pseudo differential operators by Unterberger, Bokobza [88],[89], Unterberger [87], Višik, Eskin [94], [95] and Beauzamy [3].
In [46] Jacob and Leopold constructed a Feller semigroup generated by operators with symbols (7.2). Their approach is mainly based on the method described in chapter 4 and the results of Leopold for pseudo differential operators of variable order and corresponding Sobolev spaces, [58],[59],[60]. See also Negoro [69] and Kikuchi, Negoro [64] for further results concerning existence of transition densities and path behaviour of stable-like processes.

We will consider a more general situation. The functions $|\xi|^2$ and $1+|\xi|^2$ in (7.1) and (7.2) are associated to a diffusion process, i.e. Brownian motion. Thus the stable-like process can be regarded as a diffusion subordinated by a subordinator given by the exponent $1/2\alpha(x)$, but the subordinator depends on x. Our starting point will be the generator $-\psi(D)$ of a Lévy process or even a generator with variable coefficients which satisfies upper and lower estimates with respect to the reference function ψ. We denote this symbol by $s(x,\xi)$ and consider the symbol

$$p(x,\xi) = s(x,\xi)^{m(x)}, \tag{7.3}$$

where $0 < m(x) \leq 1$. Note that if $s(x,\xi)$ is a negative definite symbol, then $p(x,\xi)$ also is negative definite.

As in Chapter 6 let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite reference function with the property that the Lévy measure of ψ has bounded support. Moreover we will assume that ψ has the minimal growth behaviour at infinity, i.e. there are constants $r > 0$ and $c > 0$ such that

$$\psi(\xi) \geq c\,|\xi|^r, \qquad |\xi| \geq 1. \tag{7.4}$$

As usual let

$$\lambda(\xi) := (1+\psi(\xi))^{1/2}. \tag{7.5}$$

Define the symbol classes $\mathrm{S}_\varrho^{m,\lambda}$ and $\mathrm{S}_0^{m,\lambda}$ as in Chapter 6. Our main result is the following

Theorem 7.1. *Let $s \in S_\varrho^{2,\lambda}$ be a real-valued negative definite symbol which is elliptic in the sense that there is a $\delta > 0$ such that*

$$s(x,\xi) \geq \delta\,\lambda^2(\xi). \tag{7.6}$$

Consider a C^∞-function $m : \mathbb{R}^n \to (0;1]$ with bounded derivatives and let $M := \sup_{x\in\mathbb{R}^n} m(x)$, $\mu := \inf_{x\in\mathbb{R}^n} m(x)$.
If

$$M - \mu < \frac{1}{2} \qquad \textit{and} \qquad \mu > 0 \tag{7.7}$$

then

$$p(x,\xi) = s(x,\xi)^{m(x)} \tag{7.8}$$

defines an operator $-p(x,D) : C_0^\infty(\mathbb{R}^n) \to C_\infty(\mathbb{R}^n)$, which is closable in $C_\infty(\mathbb{R}^n)$ and the closure is the generator of a Feller semigroup (T_t).

The proof of Theorem 7.1 again relies on the theorem of Hille-Yosida. In particular for some $\tau \geq 0$ we have to find solutions of the equation

$$(p(x,D)+\tau)u = f \tag{7.9}$$

for sufficiently many right hand sides. We split the proof into a part concerning the existence of (weak) solutions and a part dealing with regularity. We treat both parts using typical techniques for pseudo differential operators, but in a suitably modified way.
To prove existence of solutions we first apply the modified version of Friedrichs symmetrization that yields again a sharp Garding inequality, that is a lower bound for the corresponding bilinear form in terms of the lower order norm $\|\cdot\|_{\mu,\lambda}$. We then show that the bilinear form is continuous on the space obtained by closing the symmetric part of the form and finally obtain

Theorem 7.2. *Let $p(x,\xi)$ be as in Theorem 7.1 and $\tau > 0$ sufficiently large. Then for any $f \in H^{-\mu,\lambda}(\mathbb{R}^n)$ there is a unique $u \in H^{\mu,\lambda}(\mathbb{R}^n)$ such that*

$$(p(x,D)+\tau)u = f.$$

Next we show that the operator $p(x,D)$ admits a (left-)parametrix, i.e. there is a symbol q such that

$$q(x,D) \circ p(x,D) = \mathrm{id} + r(x,D),$$

where $r(x,D)$ is an operator of negative order, hence has smoothing properties. We then easily obtain the following regularity result.

Theorem 7.3. *Let $p(x,\xi)$ as in Theorem 7.1 and u be a solution of (7.9) for some $f \in H^{k,\lambda}(\mathbb{R}^n)$, $k \geq 0$. Then for all $\varepsilon > 0$ we have*

$$u \in H^{k+2\mu-\varepsilon,\lambda}(\mathbb{R}^n).$$

7.2 Existence of solutions

In this section let $s \in \mathrm{S}_\varrho^{2,\lambda}$ be a negative definite elliptic symbol, i.e. there is a $\delta > 0$ such that

$$s(x,\xi) \geq \delta\, \lambda^2(\xi), \qquad x, \xi \in \mathbb{R}^n, \tag{7.10}$$

and let $m : \mathbb{R}^n \to (0;1]$, $m \in C_b^\infty(\mathbb{R}^n)$ be as in Theorem 7.1, that is for

$$M = \sup_{x\in\mathbb{R}^n} m(x), \qquad \mu = \inf_{x\in\mathbb{R}^n} m(x) \tag{7.11}$$

we assume

$$M - \mu < \frac{1}{2}, \qquad \mu > 0 \tag{7.12}$$

and consider

$$p(x,\xi) = s(x,\xi)^{m(x)}.$$

The first property we have to check is whether $p(x,\xi)$ is a symbol in the symbol classes $\mathrm{S}_\varrho^{m,\lambda}$. Since the exponent $m(x)$ depends on x, differentiation of p with respect to x yields certain logarithmic terms of $s(x,\xi)$. In the case of symbols (7.2) this can be treated by considering the Hörmander classes $S^m_{\varrho,\delta}$ with $\delta > 0$. A similar procedure for symbols in $\mathrm{S}_\varrho^{m,\lambda}$ causes problems for the symbolic calculus, since the order of the derivatives $\partial_\xi^\alpha p(x,\xi)$ does not decrease arbitrarily as $|\alpha| \to \infty$. Therefore it is more convenient to capture the effect of the x-derivatives by slightly increasing the order of the symbol, i.e. to prove $p \in \mathrm{S}_\varrho^{2M+\varepsilon,\lambda}$ for $\varepsilon > 0$. First we need

Lemma 7.4. *Let $G, K, L : \mathbb{R}^N \to \mathbb{R}$ be C^∞-functions, $G > 0$, $L \neq 0$. Then we have for $\gamma \in \mathbb{N}_0^N$, $l = |\gamma|$*

(i)

$$\partial^\gamma \exp K = \exp K \cdot \sum_{\substack{\gamma_1+\ldots+\gamma_{l'}=\gamma \\ l'=0,1,\ldots l}} c_{\{\gamma_i\}} \prod_{i=1}^{l'} \partial^{\gamma_i} K,$$

(ii)

$$\partial^\gamma \log G = \sum_{\gamma_1+\ldots+\gamma_l=\gamma} c'_{\{\gamma_i\}} \prod_{i=1}^{l} \frac{\partial^{\gamma_i} G}{G} \quad \textit{if } \gamma \neq 0,$$

(iii)

$$\partial^\gamma \frac{1}{L} = \frac{1}{L} \cdot \sum_{\gamma_1+\ldots+\gamma_l=\gamma} c''_{\{\gamma_i\}} \prod_{i=1}^{l} \frac{\partial^{\gamma_i} L}{L}.$$

The summation is taken over all choices of multiindices $\gamma_1, \ldots, \gamma_{l'} \in \mathbb{N}_0^N$ *and* $\gamma_1, \ldots, \gamma_l \in \mathbb{N}_0^N$*, respectively, that have sum* γ*. The constants* $c_{\{\gamma_i\}}$*,* $c'_{\{\gamma_i\}}$ *and* $c''_{\{\gamma_i\}}$ *depend on the choice of the multiindices.*

The proof by induction is an elementary application of the chain rule. See also Fraenkel [19] for general higher order chain rules in several dimensions.
We now are able to prove

Proposition 7.5. *Let* $p(x, \xi)$ *be as above. Then for all* $\varepsilon > 0$

$$\left|\partial_\xi^\alpha \partial_x^\beta p(x,\xi)\right| \leq c_{\alpha\beta\varepsilon}\, p(x,\xi) \cdot \lambda^{-\varrho(|\alpha|)+\varepsilon}(\xi). \tag{7.13}$$

In particular $p \in S_\varrho^{2M+\varepsilon,\lambda}$.

Proof: We have to estimate the derivatives

$$\partial_\xi^\alpha \partial_x^\beta p(x,\xi) = \partial_\xi^\alpha \partial_x^\beta s(x,\xi)^{m(x)} = \partial_\xi^\alpha \partial_x^\beta \exp\left(m(x) \cdot \log s(x,\xi)\right).$$

We apply Lemma 7.4 (i) with $N = 2n$, $\gamma = (\alpha, \beta)$, $l = |\alpha| + |\beta|$. Thus

$$\left|\partial_\xi^\alpha \partial_x^\beta p(x,\xi)\right| \leq \exp\left(m(x) \cdot \log s(x,\xi)\right) \cdot \left| \sum_{\substack{\alpha_1+\ldots+\alpha_{l'}=\alpha \\ \beta_1+\ldots+\beta_{l'}=\beta \\ l'=0,1,\ldots l}} c_{\{\alpha_i,\beta_i\}} \cdot \prod_{i=1}^{l'} t_{\alpha_i\beta_i}(x,\xi) \right|, \tag{7.14}$$

where

$$t_{\alpha_i\beta_i}(x,\xi) = \partial_\xi^{\alpha_i} \partial_x^{\beta_i} \left(m(x) \cdot \log s(x,\xi)\right) = \sum_{\beta'_i \leq \beta_i} \binom{\beta_i}{\beta'_i} \partial_x^{\beta_i - \beta'_i} m(x) \cdot \partial_\xi^{\alpha_i} \partial_x^{\beta'_i} \log s(x,\xi). \tag{7.15}$$

Again by Lemma 7.4 (ii), if $k = |\alpha_i| + |\beta'_i| \neq 0$

$$\partial_\xi^{\alpha_i} \partial_x^{\beta'_i} \log s(x,\xi) = \sum_{\substack{\tilde{\alpha}_1+\ldots+\tilde{\alpha}_k=\alpha_i \\ \tilde{\beta}_1+\ldots+\tilde{\beta}_k=\beta'_i}} c_{\{\tilde{\alpha}_j,\tilde{\beta}_j\}} \cdot \prod_{j=1}^{k} \frac{\partial_\xi^{\tilde{\alpha}_j} \partial_x^{\tilde{\beta}_j} s(x,\xi)}{s(x,\xi)}.$$

Since $s(x,\xi)$ is an elliptic symbol im $\mathrm{S}_\varrho^{2,\lambda}$ we find

$$\begin{aligned}\left|\partial_\xi^{\alpha_i}\partial_x^{\beta_i'}\log s(x,\xi)\right| &\leq c_{\alpha_i\beta_i'}\sum_{\substack{\tilde\alpha_1+\ldots+\tilde\alpha_k=\alpha_i\\ \tilde\beta_1+\ldots+\tilde\beta_k=\beta_i'}}\prod_{j=1}^{k}\lambda^{-\varrho(|\tilde\alpha_j|)}(\xi)\\ &\leq c_{\alpha_i\beta_i'}\lambda^{-\varrho(|\alpha_i|)}(\xi),\end{aligned}$$

where we used the subadditivity of ϱ in the last step. Moreover we always have $|\log s(x,\xi)| \leq c\log\lambda(\xi) \leq c_\varepsilon\lambda^{\varepsilon/l}(\xi)$. Since $m\in C_b^\infty(\mathbb{R}^n)$ we therefore get from (7.15)

$$|t_{\alpha_i\beta_i}(x,\xi)| \leq c_{\alpha_i\beta_i\varepsilon}\cdot\begin{cases}\lambda^{-\varrho(|\alpha_i|)}(\xi), & \alpha_i\neq 0\\ \lambda^{\varepsilon/l}(\xi), & \alpha_i=0\end{cases}$$

and finally by (7.14)

$$\begin{aligned}\left|\partial_\xi^\alpha\partial_x^\beta p(x,\xi)\right| &\leq p(x,\xi)\cdot c_{\alpha\beta\varepsilon}\sum_{\substack{\alpha_1+\ldots+\alpha_{l'}=\alpha\\ \beta_1+\ldots+\beta_{l'}=\beta\\ l'=0,1,\ldots l}}\left(\prod_{\substack{i=1,\ldots,l'\\ \alpha_i\neq 0}}\lambda^{-\varrho(|\alpha_i|)}(\xi)\cdot\prod_{\substack{i=1,\ldots,l'\\ \alpha_i=0}}\lambda^{\varepsilon/l}(\xi)\right)\\ &\leq p(x,\xi)\cdot c_{\alpha\beta\varepsilon}\,\lambda^{-\varrho(|\alpha|)+\varepsilon}(\xi).\end{aligned}$$

The second statement follows immediately from $p(x,\xi)\leq c\,\lambda^{2M}(\xi)$. □

We want to consider the equation

$$(p(x,D)+\tau)u = f \tag{7.16}$$

for $\tau\geq 0$. Let $p_\tau(x,\xi)=p(x,\xi)+\tau$ and

$$B_\tau(u,v)=(p_\tau(x,D)u,v)_0,\qquad u,v\in C_0^\infty(\mathbb{R}^n),$$

be the associated bilinear form. We note

Lemma 7.6. *Let $q\in S_0^{2m,\lambda}$. Then the bilinear form*

$$(u,v)\mapsto(q(x,D)u,v)_0,\quad u,v\in C_0^\infty(\mathbb{R}^n),$$

has a continuous extension to $H^{m,\lambda}(\mathbb{R}^n)$.

This follows immediately from

$$|(q(x,D)u,v_0)| = \left|(\lambda^{-m}(D)\circ q(x,D)u,\lambda^m(D)v)_0\right| \leq c\,\|u\|_{m,\lambda}\,\|v\|_{m,\lambda}$$

by Cauchy-Schwarz inequality, since by Corollary 6.12 both $\lambda^{-m}(D)\circ q(x,D)$ and $\lambda^m(D)$ are operators of order m.

Observe that by the ellipticity of s there is a $\delta'>0$ such that

$$p(x,\xi)\geq\delta'\,\lambda^{2\mu}(\xi). \tag{7.17}$$

Theorem 7.7. *Let $\varepsilon > 0$ such that $M - \mu + \varepsilon < 1/2$ and let $M' = M + \varepsilon/2$. Then B_τ extends continuously to $H^{M',\lambda}(\mathbb{R}^n)$ and if τ is sufficiently large, the sharp Garding inequality*

$$B_\tau(u,u) \geq \frac{\delta'}{2} \|u\|^2_{\mu,\lambda} \tag{7.18}$$

holds.

Proof: The first statement is immediate from Lemma 7.6, since $p_\tau \in \mathrm{S}^{2M',\lambda}_\varrho$. Let $Q(x,\xi) = p(x,\xi) - \delta' \lambda^{2\mu}(\xi) \in \mathrm{S}^{2M',\lambda}_\varrho$. By (7.17) we have $Q(x,\xi) \geq 0$ and hence by Theorem 6.20 we know that the Friedrichs symmetrization $Q_F(x,D)$ is a symmetric nonnegative operator with symbol $Q_F \in \mathrm{S}^{2M',\lambda}_0$ such that $r = Q - Q_F \in \mathrm{S}^{2M'-1,\lambda}_0$.
Then by Lemma 7.6

$$\begin{aligned} (p(x,D)u,u)_0 - \delta' \|u\|^2_{\mu,\lambda} &= (Q(x,D)u,u)_0 \\ &= (Q_F(x,D)u,u)_0 + (r(x,D)u,u)_0 \\ &\geq -c\|u\|^2_{M'-1/2,\lambda} \geq -\frac{\delta'}{2}\|u\|^2_{\mu,\lambda} - c(\delta')\|u\|^2_0, \end{aligned}$$

Here in the last step we used (4.11) Choosing $\tau \geq c(\delta')$ proves (7.18). □

B_τ is a continuous bilinear form on $H^{M',\lambda}(\mathbb{R}^n)$ but satisfies a lower bound only with respect to a lower order norm, which of course reflects the character of varying order. To get a weak solution of (7.16) in terms of this form B_τ we will use a method which is known from the case of degenerate elliptic differential operators, see Louhivaara, Simader [62] [63]. For that purpose let

$$\widetilde{B}_\tau(u,v) = \frac{1}{2}(B_\tau(u,v) + B_\tau(v,u)), \qquad u,v \in H^{M',\lambda}(\mathbb{R}^n),$$

be the symmetric part of B_τ. Then obviously

$$\left|\widetilde{B}_\tau(u,v)\right| \leq c\|u\|_{M',\lambda} \cdot \|v\|_{M',\lambda} \tag{7.19}$$

and

$$\widetilde{B}_\tau(u,u) \geq \frac{\delta'}{2}\|u\|^2_{\mu,\lambda}\,. \tag{7.20}$$

Therefore $\widetilde{B}_\tau$ is a symmetric bilinear form on $H^{M',\lambda}(\mathbb{R}^n)$ which by (7.20) is positive and not degenerate, i.e. $\widetilde{B}_\tau(u,u) = 0$ if and only if $u = 0$, that is $\widetilde{B}_\tau$ is an inner product. Of course in general $H^{M',\lambda}(\mathbb{R}^n)$ is not complete with respect to this inner product. By H^{p_τ} we denote the completion of $H^{M',\lambda}(\mathbb{R}^n)$ with respect to the norm $\|\cdot\|_{p_\tau} = \widetilde{B}_\tau^{1/2}$. Then $(H^{p_\tau}, \|\cdot\|_{p_\tau})$ is a Hilbert space. By (7.19) and (7.20) we can construct the completion in such a way that the continuous and dense embeddings

$$H^{M',\lambda}(\mathbb{R}^n) \hookrightarrow H^{p_\tau} \hookrightarrow H^{\mu,\lambda}(\mathbb{R}^n)$$

hold.

Lemma 7.8. *B_τ is a continuous bilinear form on $(H^{p_\tau}, \|\cdot\|_{p_\tau})$.*

Proof: Since $p_\tau(x,\xi)$ is real-valued, Corollary 6.18 yields

$$\frac{1}{2}(p_\tau(x,D)+p_\tau^*(x,D)) = \frac{1}{2}\Big(p_\tau(x,D)+\bar{p}_\tau(x,D)\Big) + r_1(x,D) = p_\tau(x,D)+r_1(x,D),$$

where $r_1 \in \mathrm{S}_0^{2M'-1,\lambda}$ and therefore for $u,v \in C_0^\infty(\mathbb{R}^n)$

$$\begin{aligned} |B_\tau(u,v)| &= |(p_\tau(x,D)u,v)_0| \le \left|\frac{1}{2}((p_\tau(x,D)+p_\tau^*(x,D))u,v)_0\right| + |(r_1(x,D)u,v)_0| \\ &= \left|\widetilde{B}_\tau(u,v)\right| + |(r_1(x,D)u,v)_0|\,. \end{aligned}$$

$\widetilde{B}_\tau$ is continuous on H^{p_τ} by definition and by Lemma 7.6 $(r_1(x,D)u,v)$ is continuous on $H^{M'-1/2,\lambda}(\mathbb{R}^n)$ and therefore also on H^{p_τ}, because by $M'-\frac{1}{2} < \mu$ we have the continuous embedding

$$H^{p_\tau} \hookrightarrow H^{\mu,\lambda}(\mathbb{R}^n) \hookrightarrow H^{M'-1/2,\lambda}(\mathbb{R}^n).$$

□

Remark: Thus B_τ with domain H^{p_τ} is continuous with respect to the norm given by its symmetric part. In other words B_τ is a sectorial form in the sense of Kato [50], VI.2 or a coercive closed form in the sense of Ma, Röckner [66], I.2.3.

It is now easy to give a

Proof of Theorem 7.2: By Lemma 7.8 we know that B_τ is a continuous and by definition coercive bilinear form on H^{p_τ}. Thus by the theorem of Lax-Milgram for any f in the dual space $(H^{p_\tau})'$ there is a unique $u \in H^{p_\tau}$ such that

$$B_\tau(u,v) = \langle f, v\rangle \qquad \text{for all } v \in H^{p_\tau}.$$

We choose a sequence (u_k) in $C_0^\infty(\mathbb{R}^n)$ which converges to u in H^{p_τ} and consequently also in $H^{\mu,\lambda}(\mathbb{R}^n)$. Note that for any $v \in C_0^\infty(R^n)$ the map $u \mapsto (u,v)$ has a continuous extension to $H^{\mu-2M',\lambda}(\mathbb{R}^n)$ and $p_\tau(x,D) : H^{\mu,\lambda}(\mathbb{R}^n) \to H^{\mu-2M',\lambda}(\mathbb{R}^n)$ is continuous. Thus the equation

$$(p_\tau(x,D)u_k, v) = B_\tau(u_k,v), \qquad v \in C_0^\infty(\mathbb{R}^n),$$

yields for $k \to \infty$

$$\langle p_\tau(x,D)u, v\rangle = B_\tau(u,v) = \langle f,v\rangle \qquad \text{for all } v \in C_0^\infty(\mathbb{R}^n)$$

and therefore

$$p_\tau(x,D)u = f.$$

In particular because of the embeddings $H^{p_\tau} \hookrightarrow H^{\mu,\lambda}(\mathbb{R}^n)$ and $H^{-\mu,\lambda}(\mathbb{R}^n) \hookrightarrow (H^{p_\tau})'$ we have a unique weak solution $u \in H^{\mu,\lambda}(\mathbb{R}^n)$ of equation (7.16) for any $f \in H^{-\mu,\lambda}(\mathbb{R}^n)$. □

7.3 Regularity of solutions

Let $p(x,\xi)$ as in Section 7.2 and $\varepsilon > 0$ such that $M-\mu+\varepsilon < 1/2$. Our aim is to construct a (left-)parametix to the operator $p_\tau(x,D)$, that is an inverse modulo a smoothing operator. From the existence of such parametrix we then easily obtain regularity for the solution of equation (7.16).
The symbolic calculus for $\mathrm{S}_\varrho^{m,\lambda}$ does not yield expansion series with remainder terms of arbitrarily small order. But it turns out to be sufficient to use a first order expansion to get a smoothing remainder term, i.e. an operator which is order improving with respect to the scale of Sobolev spaces $H^{s,\lambda}(\mathbb{R}^n)$.
Define

$$q_\tau(x,\xi) = \frac{1}{p_\tau(x,\xi)}$$

Lemma 7.9. *We have*

$$q_\tau \in S_\varrho^{-2\mu+\varepsilon,\lambda}.$$

Proof: We apply Lemma 7.4 (iii) to estimate the derivatives of $q_\tau(x,\xi)$ and it follows with $l=|\alpha|+|\beta|$:

$$\left|\partial_\xi^\alpha \partial_x^\beta q_\tau(x,\xi)\right| \leq \frac{1}{p_\tau(x,\xi)} \sum_{\substack{\alpha_1+\ldots+\alpha_l=\alpha \\ \beta_1+\ldots+\beta_l=\beta}} c_{\{\alpha_i,\beta_i\}} \cdot \prod_{i=1}^{l} \left|\frac{\partial_\xi^{\alpha_i}\partial_x^{\beta_i} p_\tau(x,\xi)}{p_\tau(x,\xi)}\right|.$$

By (7.13) we have $\left|\frac{\partial_\xi^{\alpha_i}\partial_x^{\beta_i} p_\tau(x,\xi)}{p_\tau(x,\xi)}\right| \leq c_{\alpha_i\beta_i\varepsilon}\, \lambda^{-\varrho(|\alpha_i|)+\varepsilon}(\xi)$ for any $\varepsilon > 0$ and therefore by (7.10)

$$\left|\partial_\xi^\alpha \partial_x^\beta q_\tau(x,\xi)\right| \leq c_{\alpha\beta\varepsilon}\, \lambda^{-2\mu}(\xi)\, \lambda^{-\varrho(|\alpha|)+\varepsilon}(\xi)$$

for all $\varepsilon > 0$ by the subadditivity of ϱ. □

Now the proof of Theorem 7.3 is almost immediate.
Proof of Theorem 7.3: Let $f \in H^{k,\lambda}(\mathbb{R}^n)$ and u be the solution of (7.16), which is in $H^{\mu,\lambda}(\mathbb{R}^n)$ by Theorem (7.2). Then Corollary 6.18 gives

$$q_\tau(x,D) \circ p_\tau(x,D) = \mathrm{id} + r(x,D), \tag{7.21}$$

where $r \in \mathrm{S}_0^{-t,\lambda}$ for $-t = (-2\mu+\varepsilon)+(2M+\varepsilon)-1 = 2(M-\mu+\varepsilon-\frac{1}{2}) < 0$. We apply (7.21) to u and obtain

$$u = q_\tau(x,D)\circ p_\tau(x,D)u - r(x,D)u = q_\tau(x,D)f - r(x,D)u.$$

We have $q_\tau(x,D)f \in H^{k+2\mu-\varepsilon,\lambda}(\mathbb{R}^n)$ and $r(x,D)$ is order improving, that is $u \in H^{\mu,\lambda}(\mathbb{R}^n)$ implies $r(x,D)u \in H^{\mu+t,\lambda}(\mathbb{R}^n)$ and hence $u \in H^{(\mu+t)\wedge(k+2\mu-\varepsilon),\lambda}(\mathbb{R}^n)$. Applying this argument recursively finally gives $u \in H^{k+2\mu-\varepsilon,\lambda}(\mathbb{R}^n)$. □

In order to find solutions of (7.16) also in $C_\infty(\mathbb{R}^n)$ we need the Sobolev embedding

$$H^{s,\lambda}(\mathbb{R}^n) \hookrightarrow C_\infty(\mathbb{R}^n) \quad \text{if } s > \tfrac{n}{r},$$

which hold true by assumption (7.4), see Proposition 4.1. Let us give the

Proof of Theorem 7.1: Let again $\varepsilon > 0$ satisfy $M - \mu + \varepsilon < 1/2$. We know that $p \in \mathrm{S}_\varrho^{2M+\varepsilon,\lambda}$. Choose $k > 0$ such that $k > \frac{n}{r}$. Then $H^{k+2M+\varepsilon,\lambda}(\mathbb{R}^n)$ and $H^{k,\lambda}(\mathbb{R}^n)$ can be considered as dense subspaces of $C_\infty(\mathbb{R}^n)$ and it follows that

$$-p(x,D) : H^{k+2M+\varepsilon,\lambda}(\mathbb{R}^n) \to H^{k,\lambda}(\mathbb{R}^n)$$

is a densely defined operator in $C_\infty(\mathbb{R}^n)$.
Moreover $C_0^\infty(\mathbb{R}^n)$ is a core for this operator, because $C_0^\infty(\mathbb{R}^n)$ is dense in $H^{k+2M+\varepsilon,\lambda}(\mathbb{R}^n)$, $-p(x,D)$ maps $H^{k+2M+\varepsilon,\lambda}(\mathbb{R}^n)$ continuously into $H^{k,\lambda}(\mathbb{R}^n)$ and both $H^{k+2M+\varepsilon,\lambda}(\mathbb{R}^n)$ and $H^{k,\lambda}(\mathbb{R}^n)$ are continuously embedded in $C_\infty(\mathbb{R}^n)$.
Furthermore, p is a continuous negative definite symbol. Hence by Theorem 2.16 and Proposition 2.20 we know that $-p(x,D)$ satisfies the positive maximum principle on $H^{k+2M+\varepsilon,\lambda}(\mathbb{R}^n)$.
Finally let $\tau > 0$ be sufficiently large and $f \in H^{k+2(M-\mu+\varepsilon),\lambda}(\mathbb{R}^n)$. Then by Theorem 7.2 and Theorem 7.3 we know that there is a $u \in H^{k+2(M-\mu+\varepsilon)+2\mu-\varepsilon,\lambda}(\mathbb{R}^n) = H^{k+2M+\varepsilon,\lambda}(\mathbb{R}^n)$ such that

$$p_\tau(x,D)u = f.$$

In other words the range of the operator $p_\tau(x,D) = \tau - (-p(x,D))$ with domain $H^{k+2M+\varepsilon,\lambda}(\mathbb{R}^n)$ contains $H^{k+2(M-\mu+\varepsilon),\lambda}(\mathbb{R}^n)$ and is therefore dense in $C_\infty(\mathbb{R}^n)$.
The theorem of Hille-Yosida 4.7 now implies that the closure of $(-p(x,D), H^{k+2M+\varepsilon,\lambda}(\mathbb{R}^n))$ generates a Feller semigroup (T_t). □

7.4 Localization by the martingale problem

The restriction (7.8) for the oscillation of the exponent function $m(x)$ implies in particular that the bilinear form B_τ is continuous with respect to its symmetric part, i.e. sectorial and therefore is necessary in the above approach. We can avoid this condition as well as the boundedness of the derivatives of $m(x)$ if we use an approach via the martingale problem. This is mainly due to the localization technique for the martingale problem, see Theorem 5.3 and moreover to the fact that well-posedness of the martingale problem is closely related to the property that the operator generates a Feller semigroup, see Section 5.4, Proposition 5.18.
In this way we obtain

Theorem 7.10. *Let $s(x,\xi)$ be as in Theorem 7.1, $m : \mathbb{R}^n \to (0;1]$ be a C^∞-function and $p(x,\xi)$ as in (7.8).*
Then $-p(x,D) : C_0^\infty(\mathbb{R}^n) \to C_\infty(\mathbb{R}^n)$ has an extension that generates a Feller semigroup (T_t).

Remark: Let us first note that we may restrict to the conservative case, that is we may consider the symbol $\tilde{p}(x,\xi) := p(x,\xi) - p(x,0)$. Both $p(x,\xi)$ and $\tilde{p}(x,\xi)$ are negative definite symbols and $x \mapsto p(x,0)$ is a bounded continuous function. Therefore both $-p(x,D)$ and $-\tilde{p}(x,D)$ satisfy the positive maximum principle and their difference is a bounded operator in $C_\infty(\mathbb{R}^n)$. By the standard perturbation result for generators of (Feller-)semigroups, see Theorem 6.30, hence $-\tilde{p}(x,D)$ generates a Feller semigroup if and only if $-p(x,D)$ does.

Proof of Theorem 7.10: Let $p(x,\xi)$ be as in Theorem 7.10, i.e $p(x,\xi) = s(x,\xi)^{m(x)}$, where $s \in \mathrm{S}_{\varrho}^{2,\lambda}$ is an elliptic negative definite symbol and m is a C^∞-function on $\mathbb{R}^n$ with values in $(0;1]$. Then

$$\tilde{p}(x,\xi) := p(x,\xi) - p(x,0)$$

is a negative definite symbol such that $\tilde{p}(x,0) = 0$ and for a suitable $c \geq 0$

$$\tilde{p}(x,\xi) \leq c(1 + |\xi|^2).$$

Thus by Theorem 3.15 there is a solution to the martingale problem for $-\tilde{p}(x,D)$ for any initial distribution.

Next fix $x_0 \in \mathbb{R}^n$ and choose open relatively compact neighbourhoods U_{x_0}, V_{x_0} of x_0 such that $x_0 \in U_{x_0} \subset \overline{U_{x_0}} \subset V_{x_0}$ and

$$|m(x) - m(x_0)| < \frac{1}{5} \qquad \text{for all } x \in V_{x_0}.$$

Let $\varphi_{x_0} \in C_0^\infty(\mathbb{R}^n)$ such that $0 \leq \varphi_{x_0} \leq 1$, $\varphi_{x_0} = 1$ in U_{x_0} and $\operatorname{supp} \varphi_{x_0} \subset V_{x_0}$ and define

$$m_{x_0}(x) = \varphi_{x_0}(x) \cdot m(x) + (1 - \varphi_{x_0}(x)) \cdot m(x_0).$$

Then $m_{x_0} \in C_b^\infty(\mathbb{R}^n)$, $\inf\limits_{x\in\mathbb{R}^n} m_{x_0}(x) > 0$ and $\sup\limits_{x\in\mathbb{R}^n} m_{x_0}(x) - \inf\limits_{x\in\mathbb{R}^n} m_{x_0}(x) < \frac{1}{2}$ and therefore the symbol

$$p_{x_0}(x,\xi) = s(x,\xi)^{m_{x_0}(x)}$$

satisfies the conditions of Theorem 7.1 and $-p_{x_0}(x,D)$ has an extension that generates a Feller semigroup. By the above remark the same holds true for $-\tilde{p}_{x_0}(x,D)$, where

$$\tilde{p}_{x_0}(x,\xi) = \tilde{p}_{x_0}(x,\xi) - \tilde{p}_{x_0}(x,0).$$

It is well-known that for a given initial distribution generators of Feller semigroups have at most one solution to the martingale problem (see for example [17], Cor. 4.4.4). Thus again by the above existence result the martingale problem for $-\tilde{p}_{x_0}(x,D)$ is well-posed.

To proceed with the proof of Theorem 7.10 we choose a sequence $\tilde{p}_k(x,\xi) = \tilde{p}_{x_k}(x,\xi)$, $k \in \mathbb{N}$, out of the family $(\tilde{p}_x(x,\xi))_{x\in\mathbb{R}^n}$ such that $\bigcup\limits_{k\in\mathbb{N}} U_{x_k} = \mathbb{R}^n$. Since $\tilde{p}_k(x,\xi)$ coincides with $\tilde{p}(x,\xi)$ for $x \in U_{x_k}$, Theorem 5.3 implies that the martingale problem for $-\tilde{p}(x,D)$ is well-posed.

The statement of Theorem 7.10 for $-p(x,D)$ or equivalently $-\tilde{p}(x,D)$ is therefore implied by Proposition 5.18, once we know that $-\tilde{p}(x,D)$ is an operator in $C_\infty(\mathbb{R}^n)$, that is

$$-\tilde{p}(x,D) : C_0^\infty(\mathbb{R}^n) \to C_\infty(\mathbb{R}^n).$$

But this follows immediately by Theorem 6.31 since by our assumptions $\sup\limits_{x\in\mathbb{R}^n} \tilde{p}(x,\xi) \to 0$ as $\xi \to 0$. $\square$

Chapter 8

Associated Dirichlet forms, hyper-contractivity estimates, and the strong Feller property

In the previous chapters we constructed in a number of different ways semigroups generated by a pseudo differential operator $p(x,D)$ with a continuous negative definite symbol. The approach to Feller semigroups via Hilbert space techniques always gave additional information for the operator and the associated bilinear form in terms of estimates in L^2- Sobolev spaces $H^{s,\lambda}(\mathbb{R}^n)$. In particular these estimates yield that the operator also generates a strongly continuous semigroup in $L^2(\mathbb{R}^n)$. Let us recall the results we have obtained so far in different cases. In the situation considered in Chapter 4, Theorem 4.13, we have by Theorems 4.8, 4.9, 4.11 and 4.12

$$\|-p(x,D)u\|_0 \leq c\,\|u\|_{2,\lambda}, \qquad u \in H^{2,\lambda}(\mathbb{R}^n), \tag{8.1}$$

$$\|u\|_{2,\lambda} \leq c(\|-p(x,D)u\|_0 + \|u\|_0), \qquad u \in H^{2,\lambda}(\mathbb{R}^n) \tag{8.2}$$

and for the bilinear form B_τ obtained by continuous extension of $(u,v) \mapsto ((p(x,D)+\tau)u,v)_0$, $u,v \in C_0^\infty(\mathbb{R}^n)$, we have

$$|B_\tau(u,v)| \leq c\,\|u\|_{1,\lambda} \cdot \|v\|_{1,\lambda}, \qquad u \in H^{1,\lambda}(\mathbb{R}^n), \tag{8.3}$$

$$B_\tau(u,u) \geq \delta\,\|u\|_{1,\lambda}^2, \qquad u \in H^{1,\lambda}(\mathbb{R}^n) \tag{8.4}$$

for τ sufficiently large and a constant $\delta > 0$.
Moreover for all $f \in C_0^\infty(\mathbb{R}^n)$ there is a unique solution $u \in H^{2,\lambda}(\mathbb{R}^n)$ of the equation

$$(p(x,D)+\tau)u = f. \tag{8.5}$$

In particular these results yield that the operator $-p_\tau(x,D) = -(p(x,D)+\tau)$ with domain $H^{2,\lambda}(\mathbb{R}^n)$ is a closed operator in $L^2(\mathbb{R}^n)$ with core $C_0^\infty(\mathbb{R}^n)$. Furthermore by the Hille-Yosida theorem, Theorem 4.6, $(-p_\tau(x,D), H^{2,\lambda}(\mathbb{R}^n))$ generates a strongly continuous contraction semigroup in $L^2(\mathbb{R}^n)$. Moreover B_τ with domain $H^{1,\lambda}(\mathbb{R}^n)$ is a closed coercive form.

Analogously, the symbolic calculus yields the same estimates (8.1) – (8.4) and (8.5) in the situation of Theorem 6.29, see Theorems 6.14, 6.25, 6.26 and 6.28. Moreover note that this remains true in the case of perturbations considered in Theorem 6.37.

Finally also the symbols of variable order in Chapter 7 fit into this framework: In the situation of Theorem 7.1 the associated bilinear form B_τ is a continuous coercive form on the intermediate Hilbert space $(H^{p_\tau}, \|\cdot\|_{p_\tau})$, where for all $\varepsilon > 0$ and μ, M as in Theorem 7.1

$$H^{M+\varepsilon,\lambda}(\mathbb{R}^n) \hookrightarrow H^{p_\tau} \hookrightarrow H^{\mu,\lambda}(\mathbb{R}^n)$$

are continuous and dense embeddings, see Lemma 7.8 and the subsequent remark. Moreover we see as in the prove of Therorem 7.1 (with $k = 0$) that the continuous extension

$$p_\tau(x, D) : H^{2M+\varepsilon,\lambda}(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$$

of the operator satisfies the conditions of the Hille-Yosida theorem, Theorem 4.6, in $L^2(\mathbb{R}^n)$, since it is densely defined in $L^2(\mathbb{R}^n)$, dissipative as $B_\tau(u, u) \geq 0$ and satisfies the dense range condition, because for $\tau' > 0$ the equation

$$p_{\tau+\tau'}(x, D)u = f$$

has a solution $u \in H^{2M+\varepsilon,\lambda}(\mathbb{R}^n)$ for all $f \in H^{2(M-\mu+\varepsilon),\lambda}(\mathbb{R}^n)$.
Therefore the closure of $(-p_\tau(x, D), H^{2M+\varepsilon,\lambda}(\mathbb{R}^n))$ is again the generator of a strongly continuous L^2-contraction semigroup.
Moreover $C_0^\infty(\mathbb{R}^n)$ is a core for the generator of the semigroup, since $C_0^\infty(\mathbb{R}^n)$ is dense in $H^{2M+\varepsilon,\lambda}(\mathbb{R}^n)$ and $p_\tau(x, D) : H^{2M+\varepsilon,\lambda}(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is continuous.

Furthermore in all the cases above the domain of the closed coercive form B_τ is either the Sobolev space $H^{1,\lambda}(\mathbb{R}^n)$ or in the case of operators of variable order is embedded into the Sobolev space $H^{\mu,\lambda}(\mathbb{R}^n)$. Since by the standard assumption (4.4) we have that $H^{s,\lambda}(\mathbb{R}^n) \hookrightarrow H^{\frac{r}{2}s}$, the Sobolev inequality (see [85], p.20) implies

$$H^{s,\lambda}(\mathbb{R}^n) \hookrightarrow L^q(\mathbb{R}^n) \tag{8.6}$$

for $q = \frac{2n}{n-rs} > 2$ if $r \cdot s < n$, otherwise (8.6) holds for all $q > 2$. Therefore the coercivity of B_τ gives

$$\|u\|_{L^q(\mathbb{R}^n)} \leq cB_\tau(u, u), \tag{8.7}$$

for all u in the domain of B_τ and q is determined as above.

To simpify the notation we will cover all these example under the following frame:

Assume

(B.1) A is a closable linear operator in $C_\infty(\mathbb{R}^n)$ as well as in $L^2(\mathbb{R}^n)$ with domain $D(A) = C_0^\infty(\mathbb{R}^n)$. Moreover

the closure $A^{(\infty)}$ of A in $C_\infty(\mathbb{R}^n)$ generates a Feller semigroup $(T_t)_{t\geq 0}$ in $C_\infty(\mathbb{R}^n)$

and

the closure $A^{(2)}$ of A in $L^2(\mathbb{R}^n)$ generates a strongly continuous contraction semigroup $(T_t^{(2)})_{t\geq 0}$ in $L^2(\mathbb{R}^n)$.

(B.2) There is a Hilbert space $(H, \|\cdot\|_H)$ such that the dense and continuous embeddings

$$C_0^\infty(\mathbb{R}^n) \hookrightarrow H \hookrightarrow L^2(\mathbb{R}^n)$$

hold and the bilinear form $B(u,v) = (-Au, v)_0$, $u, v \in C_0^\infty(\mathbb{R}^n)$ extends continuously to a bilinear form on $H \times H$, i.e

$$|B(u,v)| \leq c\, \|u\|_H \cdot \|v\|_H\,, \qquad u, v \in H \tag{8.8}$$

such that

$$B(u,u) \geq \delta\, \|u\|_H^2\,, \qquad u \in H,$$

for some $\delta > 0$.
Moreover assume that for some $2 < q < \infty$ the inequality

$$\|u\|_{L^q(\mathbb{R}^n)} \leq c\, \|u\|_H\,, \qquad u \in H \tag{8.9}$$

holds.

Let us first remark that the L^2-semigroup and the Feller semigroups are compatible in the following sense.

Proposition 8.1 *For the semigroups* $(T_t^{(2)})_{t\geq 0}$ *and* $(T_t)_{t\geq 0}$ *in* (B.1) *we have*

$$T_t^{(2)} u = T_t u \qquad \textit{for all } u \in L^2(\mathbb{R}^n) \cap C_\infty(\mathbb{R}^n),$$

where as usual $L^2(\mathbb{R}^n)$*-functions that admit a continuous version are identified with this uniquely determined function.*

Proof: Let $(R_\alpha^{(2)})_{\alpha>0}$ and $(R_\alpha)_{\alpha>0}$ be the resolvents associated to $(T_t^{(2)})_{t\geq 0}$ and $(T_t)_{t\geq 0}$, respectively. Since the generators $A^{(2)}$ and $A^{(\infty)}$ coincide on $C_0^\infty(\mathbb{R}^n)$ and $C_0^\infty(\mathbb{R}^n)$ is a core for these generators, we know by the theorem of Hille-Yosida, Theorem 4.6, that for all $\alpha > 0$

$$V = (\alpha - A)(C_0^\infty(\mathbb{R}^n))$$

is dense in $L^2(\mathbb{R}^n)$ and in $C_\infty(\mathbb{R}^n)$ and

$$R_\alpha^{(2)} = (\alpha - A)^{-1} = R_\alpha \qquad \text{on } V,$$

hence by continuity

$$R_\alpha^{(2)} = R_\alpha \qquad \text{on } L^2(\mathbb{R}^n) \cap C_\infty(\mathbb{R}^n).$$

Therefore, for the Bochner integrals in $L^2(\mathbb{R}^n)$ and in $C_\infty(\mathbb{R}^n)$, respectively,

$$\int_0^\infty e^{-\alpha t} T_t^{(2)} u\, dt = \int_0^\infty e^{-\alpha t} T_t u\, dt$$

for all $u \in L^2(\mathbb{R}^n) \cap C_\infty(\mathbb{R}^n)$. For all $\varphi \in C_0^\infty(\mathbb{R}^n)$ the map $u \mapsto \int_{\mathbb{R}^n} u \cdot \varphi\, dx$ is a continuous linear functional on $L^2(\mathbb{R}^n)$ and $C_\infty(\mathbb{R}^n)$. Thus by the properties of the Bochner integral

$$\int_0^\infty e^{-\alpha t} (\int_{\mathbb{R}^n} T_t^{(2)} u \cdot \varphi\, dx)\, dt = \int_0^\infty e^{-\alpha t} (\int_{\mathbb{R}^n} T_t u \cdot \varphi\, dx)\, dt.$$

By the strong continuity of the semigroups and the uniqueness of the Laplace transform we consequently obtain

$$\int_{\mathbb{R}^n} T_t^{(2)} u \cdot \varphi \, dx = \int_{\mathbb{R}^n} T_t u \cdot \varphi \, dx \qquad \text{for all } \varphi \in C_0^\infty(\mathbb{R}^n),\ t \geq 0,$$

which completes the proof. □

Next we show that under the assumption above we obtain examples of Dirichlet forms. Recall that a densely defined, positive definite, real-valued bilinear form $(\mathcal{E}, D(\mathcal{E}))$ in $L^2(\mathbb{R}^n)$ is called a closed coercive form in the sense of [66] if the domain $D(\mathcal{E})$ is complete with respect to the norm $\mathcal{E}_1^{1/2}(u,u) = (\mathcal{E}(u,u) + \|u\|_0^2)^{1/2}$ and satisfies the sector condition

$$|\mathcal{E}(u,v)| \leq K\, \mathcal{E}_1^{1/2}(u,u) \cdot \mathcal{E}_1^{1/2}(v,v) \qquad \text{for all } u, v \in D(\mathcal{E}).$$

A closed coercive form $(\mathcal{E}, D(\mathcal{E}))$ is called a **semi-Dirichlet form** if for all $u \in D(\mathcal{E})$ one has the contraction property

$$u^+ \wedge 1 \in D(\mathcal{E}) \quad \text{and}$$

$$\mathcal{E}(u + u^+ \wedge 1, u - u^+ \wedge 1) \geq 0. \tag{8.10}$$

If also the dual form satisfies (8.10), i.e.

$$\mathcal{E}(u - u^+ \wedge 1, u + u^+ \wedge 1) \geq 0,$$

then $(\mathcal{E}, D(\mathcal{E}))$ is called a (non-symmetric) **Dirichlet form**. If $(\mathcal{E}, D(\mathcal{E}))$ is symmetric, then (8.10) is equivalent to $\mathcal{E}(u^+ \wedge 1, u^+ \wedge 1) \leq \mathcal{E}(u,u)$.
A (semi-) Dirichlet form in $L^2(\mathbb{R}^n)$ is called regular if $C_0(\mathbb{R}^n) \cap D(\mathcal{E}))$ is dense in both $C_0(\mathbb{R}^n)$ with respect to $\|\cdot\|_\infty$ and in $D(\mathcal{E})$ with respect to $\mathcal{E}_1^{1/2}$.
We refer to the monograph by Fukushima, Oshima, Takeda [22] as a standard source for all question related to Dirichlet forms as well as to the book [66] by Ma and Röckner, where the non-symmetric case is covered. In this context see also the monograph [72] by Oshima. Semi-Dirichlet forms and the construction of an associated process are investigated by Ma, Overbeck and Röckner in [65].

Now assume (B.1) and (B.2). Note that by (8.8) the bilinear form (B, H) is a closed coercive form on $L^2(\mathbb{R}^n)$ in the sense of [66], which is obtained as the closure of the bilinear form $(-Au, v)$, $u, v \in C_0^\infty(\mathbb{R}^n)$, since $C_0^\infty(\mathbb{R}^n)$ is dense in H.
Moreover, the closure $A^{(2)}$ of A is the Friedrichs extension of A, because $A^{(2)}$ generates the L^2-semigroup $(T_t^{(2)})$. Therefore, see [50], (B, H) is the unique closed coercive form associated to the operator $A^{(2)}$ in the sense that

$$B(u,v) = (-A^{(2)}u, v)_0 \qquad \text{for all } u \in D(A^{(2)}),\ v \in H.$$

Now $(T_t^{(2)})$ coincides on $L^2(\mathbb{R}^n) \cap C_\infty(\mathbb{R}^n)$ with the Feller semigroup (T_t). Thus $(T_t^{(2)})$ is also submarkovian, i.e:

$$0 \leq T_t^{(2)} u \leq 1 \qquad \text{for all } u \in L^2(\mathbb{R}^n) \text{ such that } 0 \leq u \leq 1.$$

But for the strongly continuous contraction semigroup $(T_t^{(2)})$ in $L^2(\mathbb{R}^n)$ the submarkovian property is equivalent to the contraction property (8.10) of the bilinear form associated to the generator $A^{(2)}$, see [66], I.4.3, I.4.4. Therefore (B,H) is semi-Dirichlet form. Moreover, since $C_0^\infty(\mathbb{R}^n) \hookrightarrow H$ densely and continuously, the semi-Dirichlet form is regular. This implies the following results.

Corollary 8.2. *Let $p(x,\xi)$ be a continuous negative definite symbol either as in Theorem 4.13 or as in Theorem 6.29. Then for τ sufficiently large the bilinear form*

$$B_\tau(u,v) = ((p(x,D)+\tau)u,v)_0, \qquad u,v \in C_0^\infty(\mathbb{R}^n),$$

extends continuously to $H^{1,\lambda}(\mathbb{R}^n)$ and $(B_\tau, H^{1,\lambda}(\mathbb{R}^n))$ is a regular semi-Dirichlet form in $L^2(\mathbb{R}^n)$. The domain of the associated L^2- generator is given by $H^{2,\lambda}(\mathbb{R}^n)$.

In the case of symbols of variable order we obtain

Corollary 8.3. *Let p be as in Theorem 7.1. Then for τ sufficiently large the bilinear form $B_\tau(u,v) = ((p(x,D)+\tau)u,v)_0$, $u,v \in C_0^\infty(\mathbb{R}^n)$, is closable in $L^2(\mathbb{R}^n)$ and the closure (B_τ, H^{p_τ}) is a regular semi-Dirichlet form with a domain H^{p_τ} such that for all $\varepsilon > 0$ the embeddings*

$$H^{M+\varepsilon,\lambda}(\mathbb{R}^n) \hookrightarrow H^{p_\tau} \hookrightarrow H^{\mu,\lambda}(\mathbb{R}^n)$$

are dense and continuous.

Note that in general the operator $p(x,D)$ is nor symmetric in $L^2(\mathbb{R}^n)$ neither the adjoint operator generates a submarkovian semigroup. Therefore in general we will not obtain a Dirichlet form. In the special case however, that $p(x,D)$ is a symmetric operator, we obtain a symmetric submarkovian semigroup in $L^2(\mathbb{R}^n)$ associated to $-(p(x,D)+\tau)$. In this case it is well-known, see [14], that the semigroup extends to a submarkovian contraction semigroup on all $L^p(\mathbb{R}^n)$, $1 \le p \le \infty$. Moreover, $-p(x,D)$ generates a Feller semigroup. Thus, if we choose the parameter $\tau = 0$, the submarkovian semigroup is contractive on $L^\infty(\mathbb{R}^n)$ and by duality on $L^1(\mathbb{R}^n)$ and finally by interpolation on $L^2(\mathbb{R}^n)$. Hence the associated form is a Dirichlet form and we obtain

Corollary 8.4. *Asumme that $p(x,\xi)$ is as in Theorem 4.13 or in Theorem 6.29 and $p(x,D)$ is a symmetric operator in $L^2(\mathbb{R}^n)$. Then the associated bilinear form $(B_0, H^{1,\lambda}(\mathbb{R}^n))$ is a regular symmetric Dirichlet form.*

Under assumption (B.2) we have that

$$\|u\|^2_{L^q(\mathbb{R}^n)} \le c\,\|u\|^2_H \le c'\,B(u,u) \qquad \text{for all } u \in H. \tag{8.11}$$

This Sobolev inequality has a number of interesting and important consequences for the semigroup and the associated process. In particular in the examples considered above we know by (8.7)

$$\|u\|^2_{L^q(\mathbb{R}^n)} \le cB_\tau(u,u) = c(B_0(u,u) + \|u\|_0^2)$$

for all u in the domain of the Dirichlet form.

If $p(x,D)$ is symmetric, therefore the following theorem of Fukushima [21], Theo. 2, is applicable; see [22] for the definition of the process and the capacity associated to a Dirichlet form:

Theorem 8.5. *Suppose that $(\mathcal{E}, D(\mathcal{E}))$ is a symmetric regular Dirichlet form in $L^2(\mathbb{R}^n)$ and satisfies*

$$\|u\|^2_{L^q(\mathbb{R}^n)} \le c(\mathcal{E}(u,u) + \|u\|_0^2)$$

for some $q > 2$. Then for the associated standard Markov process there is a Borel set $N \subset \mathbb{R}^n$ of capacity 0 such that the transition functions $p_t(x, \cdot)$, $t > 0$, are absolutely continuous with respect to the Lebesgue measure for all $x \in \mathbb{R}^n \setminus N$.

It is well-known that the Sobolev inequality (8.11) also implies bounds for the norms $\|\cdot\|_{L^p \to L^{p'}}$ for the associated semigroups, see [92] and [10] for the symmetric case. In the non-symmetric case similar estimates are obtained if the semigroup is well-defined in L^p for $1 \le p \le 2$, see [93]. Unfortunately in our case the semigroup is only defined in L^p-spaces for $2 \le p \le \infty$ and the results are not applicable.
Nevertheless we will show that (8.11) implies $\|\cdot\|_{L^2 \to L^\infty}$-estimates for the semigroup also in the non-symmetric case. For this purpose we have to consider the dual semigroup.

Recall that the dual space of $C_\infty(\mathbb{R}^n)$ is the space $\mathcal{M}_b(\mathbb{R}^n)$ of all signed measues on $\mathbb{R}^n$ of bounded variation equipped with the total variation norm $\|\cdot\|_V$. We identify functions $f \in L_1(\mathbb{R}^n)$ with the measure $f \cdot dx$ with density f with respect to the Lebesgue measure. Then $L^1(\mathbb{R}^n)$ becomes a closed subspace of $\mathcal{M}_b(\mathbb{R}^n)$ and we have $\|f\|_{L^1(\mathbb{R}^n)} = \|f \cdot dx\|_V$.

Assume that (B.1) holds. Then it is obvious that the adjoint operators $T_t^{(2)*}$ and T_t^* define semigroups of positivity preserving contractions on $L^2(\mathbb{R}^n)' = L^2(\mathbb{R}^n)$ and $C_\infty(\mathbb{R}^n)' = \mathcal{M}_b(\mathbb{R}^n)$, respectively. Note that in general these dual semigroups are not submarkovian. Moreover the strong continuity of $(T_t^{(2)})$ and (T_t) yields that the dual semigroups $(T_t^{(2)*})$ and (T_t^*) are at least weakly-$*$-continuous. Since weak-$*$-continuity in the Hilbert space $L^2(\mathbb{R}^n)$ is equivalent to weak continuity and weakly continuous semigroups are strongly continuous, see [96], p.233, we see that $(T_t^{(2)*})$ is even a strongly continuous semigroup on $L^2(\mathbb{R}^n)$. It is well-known that the generator of $(T_t^{(2)*})$ is given by the adjoint operator $A^{(2)*}$, see [8]. Moreover we have

Proposition 8.6. *The space $L^1(\mathbb{R}^n)$ is invariant under (T_t^*) and the restriction of $(T_t^*)_{t\ge 0}$ to $L^1(\mathbb{R}^n)$ is a strongly continuous semigroup on $L^1(\mathbb{R}^n)$.*

Proof: Denote the dual pairing of $\mathcal{M}_b(\mathbb{R}^n)$ and $C_\infty(\mathbb{R}^n)$ by $\langle \cdot, \cdot \rangle$ and let $u \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n) = L^2(\mathbb{R}^n) \cap \mathcal{M}_b(\mathbb{R}^n)$. Then by Proposition 8.1

$$(T_t^{(2)*}u, v)_0 = (u, T_t^{(2)}v)_0 = \langle u, T_t v\rangle = \langle T_t^* u, v\rangle \qquad \text{for all } v \in L^2(\mathbb{R}^n) \cap C_\infty(\mathbb{R}^n),$$

that is

$$\int_{\mathbb{R}^n} v \, d(T_t^* u) = \int_{\mathbb{R}^n} v \cdot (T_t^{(2)*}u)\, dx,$$

which implies that the measure $T_t^* u$ has the L^2-density $T_t^{(2)*}u$ with respect to the Lebesgue measure and

$$\int_{\mathbb{R}^n} \left| T_t^{(2)*}u \right| \, dx = \|T_t^* u\|_V \le \|u \cdot dx\|_V = \|u\|_{L^1(\mathbb{R}^n)}$$

by the contraction property of T_t^*. Hence $T_t^* u \in L^1(\mathbb{R}^n)$ and we have shown that

$$T_t^* : L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n) \to L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n).$$

Since $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$, $L^1(\mathbb{R}^n)$ is a closed subspace of $\mathcal{M}_b(\mathbb{R}^n)$ and T_t^* is continuous on $\mathcal{M}_b(\mathbb{R}^n)$, it follows that $L^1(\mathbb{R}^n)$ is invariant under T_t^* and (T_t^*) is a semigroup on $L^1(\mathbb{R}^n)$.
Note that we obtained in particular

$$T_t^{(2)*}u = T_t^* u \qquad \text{for all } u \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n).$$

To prove the strong continuity note that the set

$$M = \{u \in L_1 : T_t^* u \underset{t\to 0}{\longrightarrow} u \text{ in } L^1(\mathbb{R}^n)\}$$

is a closed subspace of $L^1(\mathbb{R}^n)$, cf. [8], Prop. I.1.4.6. Therefore it is enough to prove $L^1_+(\mathbb{R}^n) \cap L^2(\mathbb{R}^n) \subset M$, where $L^1_+(\mathbb{R}^n)$ denotes the non-negative functions in $L^1(\mathbb{R}^n)$.
Let $u \in L^1_+(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Since $(T_t^{(2)*})$ is strongly continuous we have

$$T_t^* u = T_t^{(2)*} u \underset{t\to 0}{\longrightarrow} u \text{ in } L^2(\mathbb{R}^n),$$

in particular $T_t^* u$ converges to $u \in L^1(\mathbb{R}^n)$ in measure with respect to the Lebesgue measure. But then by [2], 21.7,

$$T_t^* u \underset{t\to 0}{\longrightarrow} u \text{ in } L^1(\mathbb{R}^n)$$

provided $\|T_t^* u\|_{L^1(\mathbb{R}^n)} \underset{t\to 0}{\longrightarrow} \|u\|_{L^1(\mathbb{R}^n)}$. To that end let $\varepsilon > 0$. Then there is a compact set $K_\varepsilon \subset \mathbb{R}^n$ such that $\int_{K_\varepsilon^\complement} u\,dx < \frac{\varepsilon}{2}$ and there is a $\varphi_\varepsilon \in C_\infty(\mathbb{R}^n)$ and a $t_\varepsilon > 0$ such that $\varphi_\varepsilon = 1$ on K_ε, $0 \le \varphi_\varepsilon \le 1$ and

$$|\langle T_t u, \varphi_\varepsilon\rangle - \langle u, \varphi_\varepsilon\rangle| < \frac{\varepsilon}{2} \qquad \text{for all } t < t_\varepsilon$$

by the weak-$*$-continuity of (T_t^*). It follows that

$$\begin{aligned} \|T_t^* u\|_{L^1(\mathbb{R}^n)} &\le \|u\|_{L^1(\mathbb{R}^n)} = \int_{\mathbb{R}^n} u\,dx \le \int_{K_\varepsilon} u\,dx + \frac{\varepsilon}{2} \le \int_{\mathbb{R}^n} u \cdot \varphi_\varepsilon\,dx + \frac{\varepsilon}{2} \\ &\le \int_{\mathbb{R}^n} T_t^* u \cdot \varphi_\varepsilon\,dx + \varepsilon \le \int_{\mathbb{R}^n} T_t^* u\,dx + \varepsilon = \|T_t^* u\|_{L^1(\mathbb{R}^n)} + \varepsilon \end{aligned}$$

for all $t < t_\varepsilon$, hence $\|T_t^* u\|_{L^1(\mathbb{R}^n)} \underset{t\to 0}{\longrightarrow} \|u\|_{L^1(\mathbb{R}^n)}$. □

Using the ideas of Nash [68] we now can prove L^1-L^2-estimates for the dual semigroup which imply L^2-L^∞-estimates for the original one.

Theorem 8.7. *Assume* (B.1) *and* (B.2) *and let $N > 0$ be defined by $\frac{1}{N} = \frac{1}{2} - \frac{1}{q}$. If (T_t^*) denotes the dual semigroup on $L^1(\mathbb{R}^n)$ of (T_t), then we have with a suitable constant $c \ge 0$*

(8.12) $$\|T_t^* u\|_{L^2(\mathbb{R}^n)} \le c\,t^{-N/4} \cdot \|u\|_{L^1(\mathbb{R}^n)} \qquad \textit{for all } u \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$$

and

(8.13) $$\|T_t u\|_{L^\infty(\mathbb{R}^n)} \le c\,t^{-N/4} \cdot \|u\|_{L^2(\mathbb{R}^n)} \qquad \textit{for all } u \in L^2(\mathbb{R}^n) \cap C_\infty(\mathbb{R}^n).$$

Proof: We have already mentioned that the bilinear form (B, H) is the unique closed coercive form associated to the operator $A^{(2)}$ in the sense that

$$B(u, v) = (-A^{(2)}u, v)_0 \qquad \text{for all } u \in D(A^{(2)}),\ v \in H.$$

But then, see [50], VI.2.5, the bilinear form (B^*, H) defined by $B^*(u, v) = B(v, u)$ is associated to the adjoint operator $A^{(2)*}$ and therefore by (8.9)

$$D(A^{(2)*}) \subset D(B^*) = H \subset L^q(\mathbb{R}^n)$$

and we have for all $u \in D(A^{(2)*})$ by (8.8) and (8.9)

$$(8.14) \qquad (-A^{(2)*}u, u)_0 = B^*(u, u) = B(u, u) \geq \delta \|u\|_H^2 \geq c \|u\|_{L^q(\mathbb{R}^n)}^2$$

for some $c > 0$.

Now let $u \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, $u \neq 0$ and $t > 0$. Since $A^{(2)*}$ is associated to a closed coercive form, its complexification is sectorial in the terminology of [50] and therefore the semigroup $(T_t^{(2)*})$ is the restriction of a holomorphic semigroup. In particular

$$T_t^{(2)*} : L^2(\mathbb{R}^n) \to D(A^{(2)*})$$

and we have shown

$$T_t^* u = T_t^{(2)*} u \in L^1(\mathbb{R}^n) \cap D(A^{(2)*}) \subset L^1(\mathbb{R}^n) \cap L^q(\mathbb{R}^n).$$

Now $(T_t^{(2)*})$ is strongly continuous, hence by (8.14) following Nash [68]

$$\begin{aligned} \frac{d}{dt} \|T_t^* u\|_{L^2}^2 &= \frac{d}{dt}(T_t^* u, T_t^* u)_0 = 2\,(\frac{d}{dt} T_t^* u, T_t^* u)_0 = 2\,(A^{(2)*} T_t^* u, T_t^* u)_0 \\ &\leq -c \|T_t^* u\|_{L^q}^2 . \end{aligned}$$

Let $\frac{1}{q} + \frac{1}{q'} = 1$. Then by Hölder inequality: $\|f\|_{L^2}^{4/q'} \leq \|f\|_{L^q}^2 \cdot \|f\|_{L^1}^{(4/q')-2}$ for all $f \in L^1(\mathbb{R}^n) \cap L^q(\mathbb{R}^n)$. By the contraction property of T_t^* we have $(\|T_t^* u\|_{L^1} / \|u\|_{L^1})^{(4/q')-2} \leq 1$, hence

$$\frac{d}{dt} \|T_t^* u\|_{L^2}^2 \leq -c \|T_t^* u\|_{L^2}^{4/q'} \cdot \|u\|_{L^1}^{2-4/q'} = -c \|T_t^* u\|_{L^2}^{2+4/N} \cdot \|u\|_{L^1}^{-4/N} .$$

Thus

$$\begin{aligned} \frac{d}{dt} \left(\|T_t^* u\|_{L^2}^{-4/N} \right) &= \frac{d}{dt} \left((\|T_t^* u\|_{L^2}^2)^{-2/N} \right) = -\frac{2}{N} \|T_t^* u\|_{L^2}^{(-4/N)-2} \cdot \frac{d}{dt} \|T_t^* u\|_{L^2}^2 \\ &\geq \frac{c}{N} \|u\|_{L^1}^{-4/N} \end{aligned}$$

This differntial inequality yields

$$\begin{aligned} \|T_t^* u\|_{L^2}^{-4/N} - \frac{c}{N} \|u\|_{L^1}^{-4/N} \cdot t &\geq \lim_{t \to 0} \left(\|T_t^* u\|_{L^2}^{-4/N} - \frac{c}{N} \|u\|_{L^1}^{-4/N} \cdot t \right) \\ &= \|u\|_{L^2}^{-4/N} \geq 0, \end{aligned}$$

which gives

$$\|T_t^* u\|_{L^2} \leq \left(\frac{c}{N} t \right)^{-N/4} \cdot \|u\|_{L^1} ,$$

i.e. (8.12) with a constant c independent of u and t. Now (8.13) follows by duality. □

Applied to our examples Theorem 8.7 immediately yields

Corollary 8.8. *Let $p(x,\xi)$ as in Theorem 4.13 or Theorem 6.29 or Theorem 7.1. Then there are constants $\tau \geq 0$ and $c \geq 0$ such that the Feller semigroup generated by $-p(x,D)$ satisfies*

$$\|T_t u\|_\infty \leq c \frac{e^{\tau t}}{t^{N/4}} \cdot \|u\|_{L^2} \qquad \text{for all } u \in L^2(\mathbb{R}^n) \cap C_\infty(\mathbb{R}^n). \tag{8.15}$$

Here $N = \frac{2q}{q-2}$ and q is determined as in (8.7), In particular, $N = \frac{2n}{r}$ if $n > r$ in the case of Theorem 4.13 or 6.29 and $N = \frac{2n}{r\mu}$ if $n > r\mu$ in the case of Theorem 7.1.

The hypercontractivity estimate (8.15) has another useful application to semigroups generated by pseudo differential operators. Under a weak additional assumption on the symbol, (8.15) implies that the Feller semigroup $(T_t)_{t\geq 0}$ possesses the strong Feller property:
Recall that by the Riesz representation theorem, there exist submarkovian kernels $\mu_t(x,dy)$, such that

$$T_t u(x) = \int_{\mathbb{R}^n} u(y)\, \mu_t(x,dy) \qquad \text{for all } u \in C_\infty(\mathbb{R}^n),\ t \geq 0. \tag{8.16}$$

Therefore (8.16) defines an extension of the semigroup to $B(\mathbb{R}^n)$, the bounded Borel measurable functions, i.e. T_t maps $B(\mathbb{R}^n)$ into itself.
We say that the Feller semigroup (T_t) is a **strong Feller** semigroup if T_t maps $B(\mathbb{R}^n)$ into $C_b(\mathbb{R}^n)$ for all $t > 0$. We now can generalize Theorem 2.1 in [40]

Theorem 8.9. *Let $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite symbol such that*

$$p(x,\xi) \leq c(1 + |\xi|^2)$$

and assume that

$$\sup_{x\in\mathbb{R}^n} (p(x,\xi) - p(x,0)) \underset{\xi\to 0}{\longrightarrow} 0.$$

If $-p(x,D)$ has an extension that generates a Feller semigroup $(T_t)_{t\geq 0}$ and for all $t > 0$

$$\|T_t u\|_\infty \leq c_t \, \|u\|_{L^2(\mathbb{R}^n)} \qquad \text{for all } u \in L^2(\mathbb{R}^n) \cap C_\infty(\mathbb{R}^n) \tag{8.17}$$

holds, then $(T_t)_{t\geq 0}$ is strongly Feller

Proof: Let $x_0 \in \mathbb{R}^n$ and $\eta > 0$ be fixed. We first claim that for the kernels $\mu_t(x,dy)$ defined in (8.16)

$$\sup_{x\in B_\eta(x_0)} \mu_t(x, B_R(0)^{\complement}) \underset{R\to\infty}{\longrightarrow} 0. \tag{8.18}$$

Define $\tilde{p}(x,\xi) = p(x,\xi) - p(x,0)$ and $k(x) = p(x,0)$. Then $k \in C_b(\mathbb{R}^n)$ is nonnegative and $\tilde{p}(x,\xi)$ is a continuous negative definite symbol. Therefore the operator

$$Bu = k \cdot u$$

is a bounded operator in $C_\infty(\mathbb{R}^n)$ and $-p(x,D)$ as well as $-\tilde{p}(x,D) = -p(x,D) + B$ (with domain $C_0^\infty(\mathbb{R}^n)$) are operators in $C_\infty(\mathbb{R}^n)$ that satisfy the positive maximum principle by Theorem 2.18. Therefore by Theorem 6.30 and the subsequent remark also $-\tilde{p}(x,D)$ extends to the generator of a Feller semigroup $(\tilde{T}_t)_{t\geq 0}$, which is given by submarkovian kernels $\tilde{\mu}_t(x, dy)$ as in (8.16). Now for $u \in C_0^\infty(\mathbb{R}^n)$, $u \geq 0$

$$\begin{aligned}
\int_{\mathbb{R}^n} u(y)\,\mu_t(x, dy) &= T_t u(x) = u(x) + \int_0^t T_s(-p(x, D)u)(x)ds \\
&= u(x) + \int_0^t T_s(-\tilde{p}(x, D)u)(x)ds - \int_0^t T_s(Bu)(x)ds \\
&= \tilde{T}_t u(x) - \int_0^t T_s(k \cdot u)(x)ds \leq \tilde{T}_t u(x) \\
&= \int_{\mathbb{R}^n} u(y)\,\tilde{\mu}_t(x, dy).
\end{aligned}$$

Thus (8.18) will follow from

$$\sup_{x\in B_\eta(x_0)} \tilde{\mu}_t(x, B_R(0)^\complement) \underset{R\to\infty}{\longrightarrow} 0. \tag{8.19}$$

To that end let $\varphi \in C_0^\infty(\mathbb{R}^n)$ be such that $0 \leq \varphi \leq 1$, $\varphi|_{B_{1/2}(0)} = 1$ and $\operatorname{supp}\varphi \subset B_1(0)$. Set $\varphi_R(x) = \varphi(\frac{x}{R})$ for $R > 0$. Then by Lemma 5.19

$$\lim_{R\to\infty} \sup_{x\in\mathbb{R}^n} |\tilde{p}(x, D)\varphi_R(x)| = 0. \tag{8.20}$$

Now we have $\varphi_R(x) = 1$ for all $x \in B_\eta(x_0)$ provided $R > 2(|x_0| + \eta)$. For these R's we find

$$\begin{aligned}
\sup_{x\in B_\eta(x_0)} \tilde{\mu}_t(x, B_R(0)^\complement) &\leq \sup_{x\in B_\eta(x_0)} \int_{\mathbb{R}^n} (1 - \varphi_R(y))\,\tilde{\mu}_t(x, dy) \\
&\leq \sup_{x\in B_\eta(x_0)} (\varphi_R(x) - \tilde{T}_t\varphi_R(x)) \\
&\leq \sup_{x\in B_\eta(x_0)} \int_0^t \tilde{T}_s(\tilde{p}(x, D)\varphi_R)(x)\,ds \\
&\leq t \cdot \sup_{0\leq s\leq t} \left\|\tilde{T}_s(\tilde{p}(x, D)\varphi_R)\right\|_\infty \leq t \cdot \|\tilde{p}(x, D)\varphi_R\|_\infty \underset{R\to\infty}{\longrightarrow} 0
\end{aligned}$$

by (8.20) and (8.19), (8.18) are verified.

Now let $u : \mathbb{R}^n \to \mathbb{R}$ be a bounded measurable function and

$$g(x) := T_t u(x) = \int_{\mathbb{R}^n} u(y)\mu_t(x, dy). \tag{8.21}$$

We have to prove that g is continuous. Let $\varepsilon > 0$ and choose $R > 0$ such that

$$\sup_{x\in B_\eta(x_0)} \mu_t(x, B_R(0)^\complement) < \frac{\varepsilon}{4\,\|u\|_\infty}.$$

We take $\widetilde{R} > 0$ such that $B_{\tilde{R}}(x_0)^\complement \subset B_R(0)^\complement$ and define $u_{\tilde{R}}(y) = 1_{B_{\tilde{R}}}(x_0) \cdot u(y)$. Then it follows that

$$\begin{aligned}
|g(x) - g(x_0)| &= \left|\int_{\mathbb{R}^n} u(y)\mu_t(x, dy) - \int_{\mathbb{R}^n} u(y)\mu_t(x_0, dy)\right| \\
&\leq \left|\int_{\mathbb{R}^n} u_{\tilde{R}}(y)\mu_t(x, dy) - \int_{\mathbb{R}^n} u_{\tilde{R}}(y)\mu_t(x_0, dy)\right| \\
&\quad + \|u\|_\infty \cdot (\mu_t(x, B_{\tilde{R}}(x_0)^\complement) + \mu_t(x_0, B_{\tilde{R}}(x_0)^\complement)).
\end{aligned}$$

But for $x \in B_\eta(x_0)$ we have

$$\begin{aligned} & \|u\|_\infty \cdot (\mu_t(x, B_{\tilde{R}}(x_0)^{\complement}) + \mu_t(x_0, B_{\tilde{R}}(x_0)^{\complement})) \\ \leq \; & \|u\|_\infty \cdot (\mu_t(x, B_R(0)^{\complement} + \mu_t(x_0, B_R(0)^{\complement}) \\ < \; & \|u\|_\infty \cdot \left(\frac{\varepsilon}{4\|u\|_\infty} + \frac{\varepsilon}{4\|u\|_\infty} \right) = \frac{\varepsilon}{2}. \end{aligned} \tag{8.22}$$

On the other hand, we claim that $g_{\tilde{R}}(x)$ defined by

$$g_{\tilde{R}}(x) = \int u_{\tilde{R}}(y)\mu_t(x, dy)$$

is continuous. Indeed, since $u_{\tilde{R}} \in L^2(\mathbb{R}^n)$ and $C_0^\infty(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$, we can approximate $u_{\tilde{R}}$ by a sequence of uniformly bounded testfunctions pointwise almost everywhere and in $L^2(\mathbb{R}^n)$. Moreover, by (8.17) it is clear that the measures $\mu_t(x, \cdot)$ are absolutely continuous with respect to the Lebesgue measure. Therefore, we find using (8.17), (8.21), the dominating convergence theorem and the fact that T_t maps $C_0^\infty(\mathbb{R}^n)$ into $C_\infty(\mathbb{R}^n)$ that $g_{\tilde{R}}$ can be uniformly approximated by continuous functions. Thus there exists $\delta \in (0, \eta)$ such that for all $x \in B_\delta(x_0)$ we have $|g_{\tilde{R}}(x) - g_{\tilde{R}}(x_0)| < \frac{\varepsilon}{2}$. Using (8.22) we finally get for $x \in B_\delta(x_0)$

$$|g(x) - g(x_0)| \leq |g_{\tilde{R}}(x) - g_{\tilde{R}}(x_0)| + \frac{\varepsilon}{2} < \varepsilon,$$

proving the theorem. □

Therefore finally Corollary 8.8 yields

Corollary 8.10. *Assume that $p(x, \xi)$ is as in Theorem 4.13, 6.29 or 7.1 and*

$$\sup_{x \in \mathbb{R}^n} |p(x, \xi) - p(x, 0)| \underset{\xi \to 0}{\longrightarrow} 0.$$

Then the Feller semigroup generated by $-p(x, D)$ is strongly Feller.

Note that the strong Feller property of the semigroup offers the possibility to investigate the operator $-p(x, D)$ in the potential theoretical framework of balayage spaces as introduced by Bliedtner and Hansen [6]. In [40] among others this approach was used to study the Dirichlet problem for the operator $-p(x, D)$. The solution obtained in this way then can be identified with the solution obtained by Dirichlet space methods and probabilistic solutions. From the combination of the different approaches one gets more information for the solution. Note in particular that balayage space theory is a pointwise theory, which yields solutions of the Dirichlet problem in the class of continuous functions.

Chapter 9

A non-explosion result

So far the continuous negative definite symbol $p(x, \xi)$ satisfied at least an upper bound of type

$$|p(x,\xi)| \le c(1 + |\xi|^2). \tag{9.1}$$

Since the function $(1 + |\xi|^2)$ describes the maximal growth behaviour of a continuous negative definite function, see Theorem 2.7, the natural interpretation of (9.1) is that $p(x, \xi)$ is a continuous negative definite symbol with bounded "coefficients". In the case of a local operator, i.e. a diffusion operator

$$Lu(x) = \sum_{j,k=1}^{n} a_{jk}(x) \frac{\partial^2}{\partial x_j \partial x_k} u(x) + \sum_{j=1}^{n} b_j(x) \frac{\partial}{\partial x_j} u(x)$$

it is well-known that it is not necessary that the coefficient functions $a_{jk}(x)$ and $b_j(x)$ are bounded in order to have an associated process with infinite life time. In this respect the standard assumption on the coefficients is a quadratic growth condition for the diffusion coefficients

$$|a_{jk}(x)| \le c(1 + |x|^2)$$

and a linear growth condition for the drift coefficients

$$|b_j(x)| \le c(1 + |x|).$$

Under this conditions the associated processes constructed by solving a stochastic differential equation will not explode, i.e. is conservative, see [49], Theorem 5.2.5. Under the same condition also a solution to the martingale problem will not run to infinity in finite time, see [83]. The question we want to study in this chapter is, what growth of a general continuous negative definite symbol is allowed such that an associated process does not explode. To get some idea we first consider a simple heuristic example. Let

$$p(x, \xi) = a(x) \cdot |\xi|^{\alpha}, \qquad 0 < \alpha \le 2,$$

where $a(x)$ is a continuous positive function on $\mathbb{R}^n$, i.e. $p(x, \xi)$ is the symbol of the symmetric α-stable process furnished with a coefficient function $a(x)$. It is well-know that if $n > \alpha$ the potential- or Green-kernel for the generator of the symmetric α-stable process is given by the

Riesz potentials $N_\alpha(x,y) = \frac{c_{n,\alpha}}{|x-y|^{n-\alpha}}$, $x, y \in \mathbb{R}^n$, see [57]. Therefore the potential operator for $-p(x,D)$ is given by the kernel

$$G(x,y) = N_\alpha(x,y) \cdot \frac{1}{a(y)}$$

and we can calculate the expectation of the occupation time in a ball for an associated process started in $x \in \mathbb{R}^n$:

$$\int_0^\infty P^x(X_t \in B_R(0))\,dt = \int_{B_R(0)} G(x,y)\,dy = \int_{B_R(0)} \frac{c_{n,\alpha}}{|x-y|^{n-\alpha}} \cdot \frac{1}{a(y)}\,dy.$$

Thus if $a(x) \geq c(1+|x|^\beta)$, $c>0$, we obtain for $\beta > \alpha$ and $R \to \infty$

$$\int_0^\infty P^x(X_t \in \mathbb{R}^n)\,dt < \infty$$

and thus the process has to leave $\mathbb{R}^n$ in finite time with positive probability. It is therefore reasonable to conjecture that the critical behaviour is given in the case $\alpha = \beta$. In fact the result of this chapter will show that for $\beta \leq \alpha$ no explosion occurs.

To consider a general case fix again a continuous negative definite reference function

$$\psi : \mathbb{R}^n \to \mathbb{R}$$

such that $\psi(0) = 0$. To avoid trivial cases assume $\psi \not\equiv 0$. Now define a function

$$A_\psi(\varrho) = \sup_{|\xi| \leq \frac{1}{\varrho}} \psi(\xi), \qquad \varrho > 0. \tag{9.2}$$

For $\varrho \to \infty$ the function A_ψ measures how rapidly the reference function vanishes in 0. Note that $A_\psi(\varrho) > 0$ for all $\varrho > 0$, since if ψ vanishes in a ball $B_{\frac{1}{\varrho}}(0)$ then by Proposition 2.5 $\psi \equiv 0$, which was excluded. Moreover $A_\psi : (0,\infty) \to (0,\infty)$ is a decreasing function, such that $\lim_{\varrho\to\infty} A_\psi(\varrho) = 0$.

It will turn out that for a symbol $p(x,\xi)$ with a behaviour in ξ defined in terms of $\psi(\xi)$ a sufficient condition for non-explosion is that the growth of the symbol with respect to x as $|x| \to \infty$ is compensated by the decay of $A_\psi(\varrho)$ as $\varrho \to \infty$.

In order to describe non-explosion or conservativeness we use a formulation via the martingale problem. Recall that in Chapter 3 we defined a solution of the martingale problem as a process with paths in $D_{\mathbb{R}^n}$, that is with paths of infinite life time. Thus the existence of such solution is a way to state that the solution is conservative. In fact we used the following procedure in the proof of existence of solutions of the martingale problem in Chapter 3:
First the existence of a solution with paths in the one-point compactification $\overline{\mathbb{R}^n}$ was verified in Proposition 3.14 and next in the proof of Theorem 3.15 we showed that this solution is conservative, i.e. has paths in $D_{\mathbb{R}^n}$.

We will apply a similar procedure to symbols which are not bounded in the sense of (9.1). In the first and essential step we specify a condition which implies that a solution of the martingale problem with paths in $D_{\overline{\mathbb{R}^n}}$ does not explode. In this step we restrict to the case of operators with range in $C_\infty(\mathbb{R}^n)$. Let the reference function ψ be as above and define A_ψ as in (9.2).

Proposition 9.1. *Let $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite symbol, $p(x, 0) = 0$, such that the operator $p(x, D)$ maps $C_0^\infty(\mathbb{R}^n)$ into $C_\infty(\mathbb{R}^n)$. Consider $p(x, D)$ as an operator in $C(\overline{\mathbb{R}^n})$ by identifying $C_\infty(\mathbb{R}^n)$ with the subspace of functions $u \in C(\overline{\mathbb{R}^n})$ such that $u(\Delta) = 0$. If there is a solution of the $D_{\overline{\mathbb{R}^n}}$-martingale problem for $-p(x, D)$ with an initial distribution in $\mathcal{M}_1(\mathbb{R}^n)$ and if*

$$p(x,\xi) \le c\,\frac{1}{A_\psi(|x|)} \cdot \psi(\xi) \qquad \text{for all } |x| \ge 1, \tag{9.3}$$

then the solution has almost surely paths in $D_{\mathbb{R}^n}$.

Proof: For the proof we simply have to repeat literally the first part of the proof of Theorem 3.15. Note that the extension of the operator to functions in $C(\overline{\mathbb{R}^n})$ is denoted by A_θ in this proof. The only thing we have to check is the existence of a sequence of functions $\varphi_k \in C_0^\infty(\mathbb{R}^n)$, $k \in \mathbb{N}$, such that (φ_k) and $(p(x, D)\varphi_k)$ are sequences of functions on $\mathbb{R}^n$ which are uniformly bounded and converge pointwise to 1 and 0, respectively. Thus the proposition follows when we replace Lemma 3.16 by the following lemma. □

This is the key result of this chapter.

Lemma 9.2. *Let $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite symbol, $p(x, 0) = 0$, such that (9.3) holds. Let $\varphi \in C_0^\infty(\mathbb{R}^n)$ such that $0 \le \varphi \le 1$, $\varphi(x) = 1$ for $|x| \le 1$, $\operatorname{supp}\varphi \subset B_2(0)$ and define $\varphi_R(x) = \varphi(\frac{x}{R})$, $R \ge 1$. Then $(\varphi_R)_{R\ge1}$ and $(p(x, D)\varphi_R)_{R\ge1}$ are uniformly bounded and*

$$\begin{array}{rcl} \lim\limits_{R\to\infty} \varphi_R(x) & = & 1 \\ \lim\limits_{R\to\infty} p(x,D)\varphi_R(x) & = & 0 \end{array} \qquad \text{pointwise on } \mathbb{R}^n.$$

To prove the lemma we need an auxiliary result. For that purpose let $\theta \in \mathcal{S}(\mathbb{R}^n)$ such that $0 \le \theta \le 1$, $\theta(-x) = \theta(x)$, $\theta(0) = 1$ and $\hat\theta \ge 0$. Define for $\varrho > 0$

$$\theta_\varrho(x) = \theta\left(\frac{x}{\varrho}\right), \tag{9.4}$$

$$\chi_\varrho(x) = 1 - \theta_\varrho(x). \tag{9.5}$$

Note that by assumption $\hat\theta_\varrho(\xi) = \varrho^n \cdot \hat\theta(\varrho\xi)$ is the density of a probability measure on $\mathbb{R}^n$.

Lemma 9.3. *Let θ be as above and define for the given reference function ψ*

$$A_\psi^\theta(\varrho) = \int_{\mathbb{R}^n} \psi(\xi) \cdot \hat\theta_\varrho(\xi)\, đ\xi, \qquad \varrho > 0. \tag{9.6}$$

Then $A_\psi^\theta : (0,\infty) \to (0,\infty)$ and there is a constant $c_{n,\theta} \ge 0$ such that

$$A_\psi^\theta(\varrho) \le c_{n,\theta}\, A_\psi(\varrho), \qquad \varrho > 0. \tag{9.7}$$

Moreover

$$\lim_{\varrho\to\infty} A_\psi^\theta(\varrho) = 0 \tag{9.8}$$

and

$$A_\psi^\theta(\varrho) \le 4\, A_\psi^\theta(2\varrho). \tag{9.9}$$

If in particular $\theta(x) = e^{-|x|^2}$ then A_ψ^θ is strictly decreasing.

Proof: Note first that $\hat{\theta}_\varrho$ is strictly positive in a neighbourhood of the origin and therefore $A_\psi(\varrho) > 0$ unless $\psi \equiv 0$.
Because $\xi \mapsto \psi(\frac{\xi}{\varrho})$ is a continuous negative definite function for all $\varrho > 0$ we have by Theorem 2.7

$$\begin{aligned}\psi\left(\frac{\xi}{\varrho}\right) &\leq 2 \sup_{|\eta|\leq 1} \psi\left(\frac{\eta}{\varrho}\right) \cdot (1+|\xi|^2) \\ &\leq 2 \sup_{|\eta|\leq \frac{1}{\varrho}} \psi(\eta) \cdot (1+|\xi|^2) = 2\, A_\psi(\varrho) \cdot (1+|\xi|^2).\end{aligned}$$

But $\hat{\theta} \in \mathcal{S}(\mathbb{R}^n)$, hence $\left|\hat{\theta}(\xi)\right| \leq c_{n,\theta}\,(1+|\xi|^2)^{-(\frac{n}{2}+2)}$ and (9.7) follows from

$$\begin{aligned}A_\psi^\theta(\varrho) &= \int_{\mathbb{R}^n} \psi(\xi)\,\hat{\theta}_\varrho(\xi)\,đ\xi = \int_{\mathbb{R}^n} \psi(\xi)\cdot \varrho^n\,\hat{\theta}(\varrho\xi)\,đ\xi = \int_{\mathbb{R}^n} \psi\left(\frac{\xi}{\varrho}\right)\cdot\hat{\theta}(\xi)\,đ\xi \\ &\leq 2\,A_\psi(\varrho)\cdot \int_{\mathbb{R}^n} c_{n,\theta}\,(1+|\xi|^2)^{-(\frac{n}{2}+1)} \leq c'_{n,\theta}\,A_\psi(\varrho).\end{aligned}$$

This implies in particular $\lim_{\varrho\to\infty} A_\psi^\theta(\varrho) = 0$.
Next note that by Proposition 2.5

$$\psi(2\xi) = \sqrt{\psi(2\xi)}^2 \leq \left(\sqrt{\psi(\xi)} + \sqrt{\psi(\xi)}\right)^2 = 4\psi(\xi)$$

and

$$\hat{\theta}_{2\varrho}(\xi) = (2\varrho)^n\,\hat{\theta}(2\varrho\xi) = 2^n\,\hat{\theta}_\varrho(2\xi),$$

which gives (9.9), since

$$\begin{aligned}A_\psi^\theta(\varrho) &= \int_{\mathbb{R}^n} \psi(\xi)\,\hat{\theta}_\varrho(\xi)\,đ\xi = 2^n \int_{\mathbb{R}^n} \psi(2\xi)\,\hat{\theta}_\varrho(2\xi)\,đ\xi \\ &\leq 4 \int_{\mathbb{R}^n} \psi(\xi)\,\hat{\theta}_{2\varrho}(\xi)\,đ\xi = 4\,A_\psi^\theta(2\varrho).\end{aligned}$$

By Corollary 2.14 the continuous negative definite function ψ has a Lévy-Khinchin representation

$$\psi(\xi) = q(\xi) + \int_{\mathbb{R}^n\setminus\{0\}} (1-\cos(y,\xi))\,\mu_\psi(dy),$$

where $q \geq 0$ is a quadratic form. Thus by (9.4) and (9.5)

$$\begin{aligned}A_\psi^\theta(\varrho) &= \int_{\mathbb{R}^n} \psi(\xi)\,\hat{\theta}_\varrho(\xi)\,đ\xi = \int_{\mathbb{R}^n} q(\xi)\,\hat{\theta}_\varrho(\xi)\,đ\xi + \int_{\mathbb{R}^n}\int_{\mathbb{R}^n\setminus\{0\}} (1-\cos(y,\xi))\,\hat{\theta}_\varrho(\xi)\,\mu_\psi(dy)đ\xi \\ &= \int_{\mathbb{R}^n} q\left(\frac{\xi}{\varrho}\right)\,\hat{\theta}(\xi)\,đ\xi + \int_{\mathbb{R}^n\setminus\{0\}} (1-\theta_\varrho(y))\,\mu_\psi(dy) \qquad (9.10) \\ &= \frac{1}{\varrho^2}\int_{\mathbb{R}^n} q(\xi)\,\hat{\theta}(\xi)\,đ\xi + \int_{\mathbb{R}^n\setminus\{0\}} \chi_\varrho(y)\,\mu_\psi(dy).\end{aligned}$$

For the particular choice $\theta(x) = e^{-|x|^2}$ we see that $\chi_\varrho(x) = 1 - e^{-\left|\frac{x}{\varrho}\right|^2}$ is strictly decreasing in ϱ. Therefore, since either q or μ_ψ does not vanish, (9.10) shows thatA_ψ^θ is strictly decreasing with respect to ϱ. □

Proof of Lemma 9.2: The statement on (φ_R) are obvious and that $\lim_{R\to\infty} p(x, D)\varphi_R(x) = 0$ for fixed $x \in \mathbb{R}^n$ follows as in the proof of Lemma 3.16, since $p(x, 0) = 0$ by assumption. It remains to prove that $p(x, D)\varphi_R$ is uniformly bounded with respect to R.

Recall the Lévy-Khinchin representation (2.19) of the pseudo differential operator $p(x, D)$ for $\varphi \in C_0^\infty(\mathbb{R}^n)$:

$$\begin{aligned} p(x, D)\varphi(x) &= -\sum_{j,k=1}^n a_{jk}(x)\frac{\partial^2}{\partial x_j \partial x_k}\varphi(x) + \int_{\mathbb{R}^n\setminus\{0\}} \left(\varphi(x) - \varphi(x+y) + \frac{(y, \nabla\varphi(x))}{1+|y|^2}\right) \mu(x, dy) \\ &= -\sum_{j,k=1}^n a_{jk}(x)\frac{\partial^2}{\partial x_j \partial x_k}\varphi(x) + \int_{\mathbb{R}^n\setminus\{0\}} \left(\varphi(x) - \varphi(x+y) + 1_{\{|y|<\eta\}} \cdot (y, \nabla\varphi(x))\right) \mu(x, dy) \end{aligned}$$

for all $\eta > 0$, where the last equality follows from the fact that the difference of both integrands is even and integrable with respect to the Lévy-kernel. Since for real-valued continuous negative definite functions the Lévy-measures $\mu(x, dy)$ are symmetric, the integrals are equal.

We first consider the second order term. Let $q(x, \xi) = \sum_{j,k=1}^n a_{jk}(x)\, \xi_j \xi_k$, where $a_{jk}(x) = a_{kj}(x)$. Then for $|x| < 1$

$$\sup_{|x|<1} |a_{jk}(x)| \le \sup_{|x|\le 1} \sup_{|\xi|\le 1} q(x, \xi) \le \sup_{|x|\le 1} \sup_{|\xi|\le 1} p(x, \xi) < \infty, \tag{9.11}$$

since p is continuous. For $|x| \ge 1$ we know by (9.3)

$$q(x, \xi) \le c\frac{1}{A_\psi(|x|)} \cdot \psi(\xi).$$

Thus if there is at least one x_0 in $\{|x| \ge 1\}$ such that $q(x_0, \cdot) \not\equiv 0$, then there is a constant $c > 0$ such that

$$A_\psi(\varrho) = \sup_{|\xi|\le \frac{1}{\varrho}} \psi(\xi) \ge c'\, A_\psi(|x_0|) \cdot \sup_{|\xi|\le \frac{1}{\varrho}} q(x_0, \xi) \ge c\, \varrho^{-2}$$

and further

$$|a_{jk}(x)| \le \sup_{|\xi|\le 1} q(x, \xi) \le c\frac{1}{A_\psi(|x|)} \cdot \sup_{|\xi|\le 1} \psi(\xi) \le c \cdot |x|^2.$$

Thus together with (9.11) we have $|a_{jk}(x)| \le c(1 + |x|^2)$ for all $x \in \mathbb{R}^n$ and therefore

$$\left|\sum_{j,k=1}^n a_{jk}(x)\frac{\partial^2 \varphi_R}{\partial x_j \partial x_k}(x)\right| \le c\sum_{j,k=1}^n \frac{1+|x|^2}{R^2} \cdot \left|\frac{\partial^2 \varphi}{\partial x_j \partial x_k}\left(\frac{x}{R}\right)\right|,$$

which is bounded uniformly for all $R \ge 1$, since $\operatorname{supp} \frac{\partial^2\varphi}{\partial x_j \partial x_k}\left(\frac{\cdot}{R}\right) \subset B_{2R}(0)$.

Next consider the second part

$$\tilde{p}(x,D)\varphi(x) = \int_{\mathbb{R}^n\setminus\{0\}} \big(\varphi(x) - \varphi(x+y) + 1_{\{|y|<\eta\}}\cdot(y,\nabla\varphi(x))\big)\ \mu(x,dy). \tag{9.12}$$

By Taylor's formula we have

$$\varphi_R(x+y) = \varphi_R(x) + (y,\nabla\varphi_R(x)) + \int_0^1 (1-t)\sum_{j,k=1}^n \frac{\partial^2\varphi_R}{\partial x_j\partial x_k}((1-t)x+ty)\cdot y_j y_k\,dt$$

and we estimate the remainder term by

$$\left|\int_0^1 (1-t)\sum_{j,k=1}^n \frac{\partial^2\varphi_R}{\partial x_j\partial x_k}((1-t)x+ty)\cdot y_j y_k\,dt\right| \le \frac{1}{2}n^2\cdot \sup_{\substack{x\in\mathbb{R}^n\\ j,k=1,\ldots,n}} \left|\frac{\partial^2\varphi_R}{\partial x_j\partial x_k}(x)\right|\cdot|y|^2 \le \kappa\cdot R^{-2}\cdot|y|^2,$$

where $\kappa = \frac{1}{2}n^2\cdot\sup_{\substack{x\in\mathbb{R}^n\\ j,k=1,\ldots,n}}\left|\frac{\partial^2\varphi}{\partial x_j\partial x_k}(x)\right|$ is independent of R. Therefore

$$|\varphi_R(x) - \varphi_R(x+y) + (y,\nabla\varphi_R(x))| \le \kappa\cdot R^{-2}\cdot|y|^2. \tag{9.13}$$

Set $\theta(x) = e^{-|x|^2}$, $\theta_\varrho(x) = \theta(\frac{x}{\varrho})$, $\chi_\varrho(x) = 1-\theta_\varrho(x)$ and define A_ψ^θ as in Lemma 9.3. Then there is a constant $K\ge 0$ such that $K\,(1-\theta(x))\ge 1$ for $|x|\le 1$ and $K\,(1-\theta(x))\ge\kappa\cdot|x|^2$ for $|x|\le 1$. It follows

$$\begin{array}{lllll} & K\cdot\chi_\varrho(x) & \ge & 1 & \text{for } |x|\ge\varrho \\ (9.14) & & & & \\ & K\cdot\chi_\varrho(x) & \ge & \kappa\cdot\frac{|x|^2}{\varrho^2} & \text{for } |x|\le\varrho \end{array}$$

We consider four different cases:

Case 1: $|x|\le R$.
Then $\varphi_R(x)=1$ and $\nabla\varphi_R(x)=0$ and therefore

$$\tilde{p}(x,D)\varphi_R(x) = \int_{\mathbb{R}^n\setminus\{0\}} (1-\varphi_R(x+y))\ \mu(x,dy).$$

If $|x|<1$ we know by Theorem 2.7 and the continuity of p

$$\tilde{p}(x,\xi)\le p(x,\xi)\le \sup_{\substack{|\eta|\le 1\\ |x|\le 1}} p(x,\eta)\cdot(1+|\xi|^2)\le c\cdot(1+|\xi|^2)$$

and therefore since $R\ge 1$

$$\begin{aligned} |\tilde{p}(x,D)\varphi_R(x)| &\le \int_{\mathbb{R}^n}\tilde{p}(x,\xi)\,\hat{\varphi}_R(\xi)\,đ\xi \le c\int_{\mathbb{R}^n}(1+|\xi|^2)\cdot R^n\,\hat{\varphi}(R\xi)\,đ\xi \\ &\le c\int_{\mathbb{R}^n}(1+|\xi|^2)\,\hat{\varphi}(\xi)\,đ\xi < \infty \end{aligned}$$

independently of R.
Now assume that $|x|\ge 1$. Note that for $|y|\ge R$ by (9.14)

$$1-\varphi_R(x+y)\le 1\le K\cdot\chi_R(y)$$

and for $|y| \le R$ by (9.13) and (9.14)

$$1 - \varphi_R(x+y) \le \kappa \cdot R^{-2}\,|y|^2 \le K \cdot \chi_R(y).$$

Hence

$$\begin{aligned}
|\tilde{p}(x,D)\varphi_R(x)| &\le K \cdot \int_{\mathbb{R}^n\setminus\{0\}} \chi_R(y)\,\mu(x,dy) = K \cdot \int_{\mathbb{R}^n\setminus\{0\}} \int_{\mathbb{R}^n} (1-\cos(y,\xi))\hat{\theta}_R(\xi)\,đ\xi\,\mu(x,dy) \\
&= K \cdot \int_{\mathbb{R}^n} \tilde{p}(x,\xi)\cdot\hat{\theta}_R(\xi)\,đ\xi \le K \cdot \int_{\mathbb{R}^n} p(x,\xi)\cdot\hat{\theta}_R(\xi)\,đ\xi \\
&\le c \cdot K \frac{1}{A_\psi(|x|)} \int_{\mathbb{R}^n} \psi(\xi)\cdot\hat{\theta}_R(\xi)\,đ\xi = c\cdot K\frac{A_\psi^\theta(R)}{A_\psi(|x|)} \\
&\le c\cdot K \frac{A_\psi^\theta(R)}{A_\psi(R)},
\end{aligned}$$

where in the last step we used the monotony of A_ψ. But this in bounded uniformly with respect to R by (9.7).

Case 2: $2R \le |x| \le 4R$.
Then $\varphi_R(x) = 0$, $\nabla\varphi_R(x) = 0$ and

$$\tilde{p}(x,D)\varphi_R(x) = -\int_{\mathbb{R}^n\setminus\{0\}} \varphi_R(x+y)\,\mu(x,dy).$$

Again for $|y| \ge R$

$$\varphi_R(x+y) \le 1 \le K\cdot\chi_R(y)$$

and for $|y| \le R$ by (9.13)

$$\varphi_R(x+y) \le \kappa\cdot R^{-2}\,|y|^2 \le K\cdot\chi_R(y).$$

Thus as in Case 1 by (9.9)

$$\begin{aligned}
|\tilde{p}(x,D)\varphi_R(x)| &\le K\cdot\int_{\mathbb{R}^n\setminus\{0\}} \chi_R(y)\,\mu(x,dy) \\
&\le c\cdot K\frac{A_\psi^\theta(R)}{A_\psi(|x|)} \le 16\,c\cdot K\frac{A_\psi^\theta(4R)}{A_\psi(|x|)} \\
&\le 16\,c\cdot K\frac{A_\psi^\theta(4R)}{A_\psi(4R)}.
\end{aligned}$$

Case 3: $|x| > 4R$.
Then as in Case 2

$$\tilde{p}(x,D)\varphi_R(x) = -\int_{\mathbb{R}^n\setminus\{0\}} \varphi_R(x+y)\,\mu(x,dy)$$

and

$$\begin{aligned}
\varphi_R(x+y) &= 0 \le K\cdot\chi_{|x|-2R}(y) \qquad \text{for } |y| \le |x|-2R, \\
\varphi_R(x+y) &\le 1 \le K\cdot\chi_{|x|-2R}(y) \qquad \text{for } |y| \ge |x|-2R.
\end{aligned}$$

Hence

$$\begin{aligned}|\tilde{p}(x,D)\varphi_R(x)| &\leq K\cdot\int_{\mathbb{R}^n\setminus\{0\}}\chi_{|x|-2R}(y)\,\mu(x,dy)\\ &\leq c\cdot K\frac{A_\psi^\theta(|x|-2R)}{A_\psi(|x|)}\leq c\cdot K\frac{A_\psi^\theta(\frac{|x|}{2})}{A_\psi(|x|)}\\ &\leq 4\,c\cdot K\frac{A_\psi^\theta(|x|)}{A_\psi(|x|)}\end{aligned}$$

by the monotony of A_ψ^θ. Hence (9.7) again gives the uniform bound.

Case 4: $R<|x|<2R$. We choose the free parameter η in (9.12) to be $\eta=R$. Then for $|y|\geq R$

$$|\varphi_R(x)-\varphi_R(x+y)|\leq 1\leq K\cdot\chi_R(y)$$

and for $|y|<R$ by (9.13)

$$|\varphi_R(x)-\varphi_R(x+y)+(y,\nabla\varphi_R(x))|\leq\kappa\cdot R^{-2}\cdot|y|^2\leq K\cdot\chi_R(y)$$

and therefore again

$$\begin{aligned}|\tilde{p}(x,D)\varphi_R(x)| &\leq \int_{\mathbb{R}^n\setminus\{0\}}\left|\varphi_R(x)-\varphi_R(x+y)+1_{\{|y|<R\}}\cdot(y,\nabla\varphi_R(x))\right|\,\mu(x,dy)\\ &\leq K\cdot\int_{\mathbb{R}^n\setminus\{0\}}\chi_R(y)\,\mu(x,dy)\\ &\leq c\cdot K\frac{A_\psi^\theta(R)}{A_\psi(|x|)}\leq 4\,c\cdot K\frac{A_\psi^\theta(2R)}{A_\psi(2R)},\end{aligned}$$

which is a uniform bound also in the last case. □

We now use Proposition 9.1 in order to construct a non-exploding solution of the martingale problem in the general case.

Theorem 9.4. *Let $\psi:\mathbb{R}^n\to\mathbb{R}$ be a continuous negative definite reference function, $\psi\not\equiv 0$, such that $\psi(0)=0$. Define $A_\psi(\varrho)$ as in (9.2).*
Let $p:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ be a continuous negative definite symbol such that $p(x,0)=0$. If

$$p(x,\xi)\leq c\frac{1}{A_\psi(|x|)}\,\psi(\xi)\qquad\textit{for } |x|\geq 1,$$

then for any initial distribution $\mu\in\mathcal{M}_1(\mathbb{R}^n)$ there is a solution of the $D_{\mathbb{R}^n}$-martingale problem for $-p(x,D)$.

Proof: Let $\theta\in C_0^\infty(\mathbb{R}^n)$ such that $0\leq\theta\leq 1$, $\theta(-x)=\theta(x)$, $\theta(0)=1$, $\operatorname{supp}\theta\subset B_1(0)$ and $\hat{\theta}\geq 0$. One easily checks that such function can be obtained for example as $\theta=\tilde{\theta}*\tilde{\theta}$ for a function $\tilde{\theta}\in C_0^\infty(\mathbb{R}^n)$, $\tilde{\theta}\geq 0$, $\tilde{\theta}(-x)=\tilde{\theta}(x)$, $\operatorname{supp}\tilde{\theta}\subset B_{1/2}(0)$ and $\int_{\mathbb{R}^n}\tilde{\theta}^2(x)\,dx=1$.
Moreover let $\theta_\varrho(x)=\theta(\frac{x}{\varrho})$, $\varrho>0$. Define

$$p_1^\theta(x,\xi)=\int_{\mathbb{R}^n}(p(x,\xi+\eta)-p(x,\eta))\,\hat{\theta}_{\frac{|x|}{2}\vee 1}(\eta)\,đ\eta$$

and decompose

$$p(x,\xi) = p_1^\theta(x,\xi) + p_2^\theta(x,\xi).$$

Then Theorem 3.12 yields, fix x for a moment, that $p_1^\theta(x,\xi)$ and $p_2^\theta(x,\xi)$ are negative definite symbols and

$$-p_2^\theta(x,D)u(x) = \int_{\mathbb{R}^n\setminus\{0\}} (u(x+y)-u(x))\,\tilde{\mu}(x,dy), \qquad u \in C_0^\infty(\mathbb{R}^n),$$

is a Lévy-type operator with a Lévy-kernel consisting of finite measures

$$\tilde{\mu}(x,dy) = (1-\theta_{\frac{|x|}{2}\vee 1}(y))\,\mu(x,dy),$$

where $\mu(x,dy)$ is the Lévy-kernel of $p(x,\xi)$.

Moreover p_1^θ and p_2^θ are continuous since for all (x,ξ) in a relatively compact set of $\mathbb{R}^n\times\mathbb{R}^n$ we know $p(x,\eta)\le c(1+|\eta|^2)$ and therefore for all $N\in\mathbb{N}$

$$\begin{aligned}
&\left|(p(x,\xi+\eta)-p(x,\eta))\,\hat{\theta}_{\frac{|x|}{2}\vee 1}(\eta)\right| \\
&\le \ \left(c(1+|\xi+\eta|^2)+c(1+|\eta|^2)\right)\cdot\left(\frac{|x|}{2}\vee 1\right)^n\cdot\hat{\theta}\left(\left(\frac{|x|}{2}\vee 1\right)\cdot\eta\right) \\
&\le \ c(1+|\eta|^2)\cdot\left(1+\left|\left(\frac{|x|}{2}\vee 1\right)\cdot\eta\right|^2\right)^{-N} \\
&\le \ c(1+|\eta|^2)^{1-N}.
\end{aligned}$$

Thus for N sufficiently large we have a uniform integrable bound for the integral which defines $p_1^\theta(x,\xi)$ and continuity follows from dominated convergence.

Moreover note that

$$p_1^\theta(x,D) : C_0^\infty(\mathbb{R}^n)\to C_0(\mathbb{R}^n). \tag{9.15}$$

In fact for $\varphi\in C_0^\infty(\mathbb{R}^n)$ such that $\operatorname{supp}\varphi\subset B_R(0)$ consider $x\in\mathbb{R}^n$ such that $|x|>2R\vee 2$. Then $x\notin\operatorname{supp}\varphi$ and therefore by (3.20) using the Lévy-Khinchin representation

$$p_1^\theta(x,D)\varphi(x) = -\int_{\mathbb{R}^n\setminus\{0\}} \varphi(x+y)\,\theta_{\frac{|x|}{2}}(y)\,\mu(x,dy).$$

But $\operatorname{supp}\varphi(x+\cdot)\cap\operatorname{supp}\theta_{\frac{|x|}{2}}=\emptyset$ by construction and hence $\operatorname{supp}p_1^\theta(x,D)\varphi\subset B_{2R\vee 2}(0)$.

Since $p_1^\theta(x,D)$ satisfies (9.15) it is possible to extend $-p_1^\theta(x,D)$ to an operator A_θ on functions in $C(\overline{\mathbb{R}^n})$ as defined in (3.24), (3.25). Hence by Proposition 3.14 the corresponding $D_{\overline{\mathbb{R}^n}}$-martingale problem is solvable and consequently Proposition 9.1 shows that for any initial distribution $\mu\in\mathcal{M}_1(\mathbb{R}^n)$ there is a solution of the $D_{\mathbb{R}^n}$-martingale problem for $-p_1^\theta(x,D)$. Thus by the perturbation argument of Proprosition 3.6 the theorem is proven provided the total masses of the Lévy-measures $\tilde{\mu}(x,dy)$ of $p_2^\theta(x,\xi)$ are uniformly bounded.

To that end define A_ψ^θ as in Lemma 9.3 for the given cut-off function θ. Then for $|x|<1$ again $p(x,\xi)\le c(1+|\xi|^2)$ and therefore with the notation of Lemma 2.15

$$\begin{aligned}\|\tilde\mu(x,\cdot)\|_\infty &= \int_{\mathbb{R}^n\setminus\{0\}}(1-\theta_1(y))\,\mu(x,dy)\le c\int_{\mathbb{R}^n\setminus\{0\}}\frac{|y|^2}{1+|y|^2}\,\mu(x,dy)\\ &= c\int_{\mathbb{R}^n\setminus\{0\}}\int_{\mathbb{R}^n}(1-\cos(y,\xi))\,\nu(d\xi)\,\mu(x,dy)\\ &\le c\int_{\mathbb{R}^n}p(x,\xi)\,\nu(d\xi)\le c\int_{\mathbb{R}^n}(1+|\xi|^2)\,\nu(d\xi)<\infty.\end{aligned}$$

On the other hand for $|x|\ge 1$

$$\begin{aligned}&\int_{\mathbb{R}^n\setminus\{0\}}\left(1-\theta_{\frac{|x|}{2}\vee 1}(y)\right)\,\mu(x,dy)\\ &= \int_{\mathbb{R}^n\setminus\{0\}}\int_{\mathbb{R}^n}(1-\cos(y,\xi))\cdot\hat\theta_{\frac{|x|}{2}\vee 1}(\xi)\,đ\xi\,\mu(x,dy)\\ &\le \int_{\mathbb{R}^n}p(x,\xi)\cdot\hat\theta_{\frac{|x|}{2}\vee 1}(\xi)\,đ\xi\le\frac{c}{A_\psi(|x|)}\int_{\mathbb{R}^n}\psi(\xi)\cdot\hat\theta_{\frac{|x|}{2}\vee 1}(\xi)\,đ\xi\\ &\le c\frac{A_\psi^\theta\left(\frac{|x|}{2}\vee 1\right)}{A_\psi(|x|)}\le c\frac{A_\psi^\theta\left(\frac{|x|}{2}\right)}{A_\psi(|x|)}\le 4c\frac{A_\psi^\theta(|x|)}{A_\psi(|x|)},\end{aligned}$$

which is uniformly bounded by Lemma 9.3. □

To conclude the investigations we finally combine this result with the uniqueness result derived in Chapter 5 and use the localization technique. We obtain the following result.

Theorem 9.5. *Assume that $\psi:\mathbb{R}^n\to\mathbb{R}$ is a continuous negative definite reference function with $\psi(0)=0$, which satisfies for $|\xi|\ge 1$*

$$\psi(\xi)\ge c\,|\xi|^r$$

for some $r>0$, $c>0$. Let $A_\psi(\varrho)=\sup_{|\xi|\le\frac{1}{\varrho}}\psi(\xi)$, $\varrho>0$, and let M be the smallest integer such that $M>\left(\frac{n}{r}\vee 2\right)+n$, $k=2M+1-n$.
Let $p:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ be a continuous negative definite symbol, $p(x,0)=0$, such that $p(\cdot,\xi)\in C^{(k)}(\mathbb{R}^n)$. If

$$p(x,\xi)\le c\frac{1}{A_\psi(|x|)}\cdot\psi(\xi)\qquad\text{for } |x|\ge 1,\ \xi\in\mathbb{R}^n,$$

and for a locally bounded function $f:\mathbb{R}^n\to\mathbb{R}^+$ and all $|\beta|\le k$

$$\left|\partial_x^\beta p(x,\xi)\right|\le f(x)\cdot\psi(\xi)\qquad\text{for } x\in\mathbb{R}^n,\ |\xi|\ge 1,$$

and for a strictly positive function $g:\mathbb{R}^n\to(0,\infty)$

$$p(x,\xi)\ge g(x)\cdot\psi(\xi)\qquad\text{for } x\in\mathbb{R}^n,\ |\xi|\ge 1,$$

then the $D_{\mathbb{R}^n}$-martingale problem for $-p(x,D)$ is well-posed.

Proof: By Theorem 9.4 there is a solution for every initial distribution. Moreover note that $p(x,D)$ maps $C_0^\infty(\mathbb{R}^n)$ into $C_b(\mathbb{R}^n)$. In fact for $u \in C_0^\infty(\mathbb{R}^n)$ such that $\operatorname{supp} u \subset B_R(0)$ we have for φ_R choosen as in Lemma 9.2

$$|u(x)| \leq \|u\|_\infty \cdot \varphi_R(x).$$

Therefore for $|x| > 2R$ using the Lévy-type representation of $p(x,D)$

$$\begin{aligned} |p(x,D)u(x)| &\leq \int_{\mathbb{R}^n\setminus\{0\}} |u(x+y)|\ \mu(x,dy) \leq \|u\|_\infty \cdot \int_{\mathbb{R}^n\setminus\{0\}} |\varphi_R(x+y)|\ \mu(x,dy) \\ &= \|u\|_\infty \cdot (-p(x,D)\varphi_R)(x). \end{aligned}$$

But $p(x,D)\varphi_R$ is bounded by Lemma 9.2 and hence $p(x,D)u \in C_b(\mathbb{R}^n)$.

For a ball $B_R(x_0)$ let $\chi \in C_0^\infty(\mathbb{R}^n)$ be such that $0 \leq \chi \leq 1$ and $\chi = 1$ in $B_R(x_0)$. Define

$$p_R(x,\xi) = p(x_0,\xi) + \chi(x)(p(x,\xi) - p(x_0,\xi)).$$

Then p_R is a continuous negative definite symbol which satisfies for all $|\beta| \leq k$

$$\begin{aligned} \left|\partial_x^\beta p_R(x,\xi)\right| &\leq \left|\partial_x^\beta p(x_0,\xi)\right| + \left| \sum_{\substack{\gamma\leq\beta \\ \gamma\in\mathbb{N}_0^n}} \binom{\beta}{\gamma} \partial_x^{\beta-\gamma}\chi(x) \cdot \partial_x^\gamma (p(x,\xi) - p(x_0,\xi)) \right| \\ &\leq c \cdot \psi(\xi) \qquad \text{for all } |\xi| \geq 1 \end{aligned}$$

by the support properties of χ and since f is locally bounded. Moreover for $|\xi| \geq 1$

$$\begin{aligned} p_R(x,\xi) &\geq \chi(x)g(x) \cdot \psi(\xi) + (1-\chi(x))g(x_0) \cdot \psi(\xi) \\ &\geq (g(x_0) \wedge g(x)) \cdot \psi(\xi) \qquad \text{for all } |\xi| \geq 1. \end{aligned}$$

Therefore p_R satisfies the assumptions of Theorem 5.7 and the $D_{\mathbb{R}^n}$-martingale problem for $-p_R(x,D)$ is well-posed.

Note that $p_R(x,\xi) \leq c(1+|\xi|^2)$ by the above calculations and therefore $p_R(x,D) : C_0^\infty(\mathbb{R}^n) \to C_b(\mathbb{R}^n)$. Moreover $p_R(x,\xi)$ coincides with $p(x,\xi)$ for $x \in B_R(x_0)$. Cover $\mathbb{R}^n$ by countably many balls $B_R(x_0)$. Then the localization procedure of Theorem 5.3 yields the desired result. □

Bibliography

[1] R.F. Bass, *Uniqueness in law for pure jump Markov processes.* Probab.Th.Rel.Fields 79, (1988) 271-287.

[2] H. Bauer, *Maß- und Integrationstheorie.* De Gruyter, Berlin - New York (1990).

[3] A. Beauzamy, *Espaces de Sobolev et Besov d'ordre variable définis sur L^p.* C.R.Acad. Sci. Paris (Ser. A) 274 (1972) 1935–1938

[4] C. Berg, G. Forst, *Potential theory on locally compact Abelian groups.* Ergebnisse der Mathematik und ihrer Grenzgebiete, II. Ser. Bd.87, Springer Verlag, Berlin - Heidelberg - New York (1975).

[5] P. Billingsley, *Convergence of probability measures.* John Wiley, New York - London - Sydney - Toronto (1968).

[6] J. Bliedtner, W. Hansen, *Potential theory.* Universitext, Springer Verlag, Berlin - Heidelberg -New York - Tokyo (1986).

[7] S. Bochner, *Harmonic analysis and the theory of probability.* California Monographs in Mathematical Science, University of California Press, Berkeley (1955).

[8] P.L. Butzer, H. Berens, *Semigroups of operators and approximation.* Die Grundlehren der mathematischen Wissenschaften, Bd. 145, Springer Verlag, Berlin - Heidelberg - New York (1967).

[9] A. Calderón, R. Vaillancourt, *On the boundedness of pseudo-differential operators.* J.Math.Soc.Japan 23, No.2 (1971) 374–378.

[10] E. Carlen, S. Kusuoka, D.W. Stroock, *Upper bounds for symmetric Markov transition functions.* Ann. Inst. Henri Poincaré, Probabilités et Statistiques, Sup. au n.23,2 (1987) 245–287.

[11] Z. Ciesielski, G. Kerkyacharian, B. Roynette, *Quelques espaces fonctionnels associés à des processus Gaussiens.* Studia Math. 107 (1993) 171–204.

[12] Ph. Courrège, *Générateur infinitésimal d'un semi-groupe de convolution sur R^n et formule de Lévy-Khinchin.* Bull. Sci. Math. 2[e] sér. 88 (1964) 3–30.

[13] Ph. Courrège, *Sur la forme intégro-différentielle des opérateurs de C_k^∞ dans C satisfaisant au principe du maximum.* Sém. Théorie du Potentiel (1965/66) Exposé 2.

[14] E.B. Davies, *Heat kernels and spectral theory*. Cambridge Tracts in Mathematics, Vol. 92, Cambrigde University Press, Cambridge (1989).

[15] J.L. Doob, *Stochastic processes*. Wiley Classics Library, John Wiley, New York - Chicester - Brisbane - Toronto - Singapore (1990).

[16] E.B. Dynkin, *Markov processes*. Vol. I. Grundlehren der mathematischen Wissenschaften, Vol. 121, Springer Verlag, Berlin - Göttingen - Heidelberg (1965).

[17] S.N. Ethier, Th.G. Kurtz, *Markov processes — characterization and convergence*. Wiley Series in Probability and Mathematical Statistics, John Wiley, New York - Chicester - Brisbane - Toronto - Singapore (1986).

[18] W. Feller, *An introduction to probability theory and its applications*. Vol. II, 2nd edition, Wiley series in Probability and Mathematical Statistics, John Wiley & Sons, New York (1971).

[19] L.E. Fraenkel, *Formulae for high derivatives of composite functions*. Math.Proc.Camb. Phil. Soc. 83 (1978) 159–165.

[20] A. Friedman, *Partial Differential Equations*. Holt, Rinehart and Winston, Inc., New York - Chicago - San Francisco - Atlanta - Dallas - Montreal - Toronto - London - Sydney (1969).

[21] M. Fukushima, *On an L^p-estimate of resolvents of Markov processes*. Publ. R.I.M.S. Kyoto Univ. 13 (1977) 191–202.

[22] M. Fukushima, Y. Oshima, M. Takeda, *Dirichlet forms and symmetric Markov processes*. De Gruyter Studies in Mathematics, Vol. 19, Walter de Gruyter Verlag, Berlin - New York (1994).

[23] K. Harzallah, *Fonctions opérant sur les fonctions définies-négatives*. Ann. Inst. Fourier 17,1 (1967) 443–468.

[24] K. Harzallah, *Sur une démonstration de la formule de Lévy-Khinchine*. Ann. Inst. Fourier. 19,2 (1969) 527–532.

[25] V. Herren, *Lévy-type processes and Besov spaces*. Potential Analysis 7 (1997) 689–704.

[26] C.J. Himmelberg, *Measurable relations*. Fund. Math 87 (1975) 53–72.

[27] F. Hirsch, *Principes de maximum pour les noyaux de convolution*. Sém. Théorie du Potentiel, Lecture Notes in Mathematics, Vol. 713, Springer Verlag, Berlin-Heidelberg-New York (1979) 113–136.

[28] L. Hörmander, *The analysis of linear partial differential operators III*. Die Grundlehren der mathematischen Wissenschaften, Bd. 274, Springer Verlag, Berlin - Heidelberg - New York - Tokyo (1985).

[29] W. Hoh, *Some commutator estimates for pseudo differential operators with negative definite functions as symbols*. Integr.Equ.Oper.Th. 17 (1993) 46–53.

[30] W. Hoh, *Das Martingalproblem für eine Klasse von Pseudodifferentialoperatoren.* Dissertation, Universität Erlangen-Nürnberg, 1992.

[31] W. Hoh, *The martingale problem for a class of pseudo differential operators.* Math.Ann. 300 (1994) 121–147.

[32] W. Hoh, *Pseudo differential operators with negative definite symbols and the martingale problem.* Stoch. and Stoch. Rep. 55 (1995) 225–252.

[33] W. Hoh, *Feller semigroups generated by pseudo differential operators.* Intern. Conf. Dirichlet Forms and Stoch. Processes, Walter de Gruyter Verlag 1995, 199–206.

[34] W. Hoh, *A symbolic calculus for pseudo differential operators generating Feller semigroups.* To appear in Osaka J. Math.

[35] W. Hoh, *Pseudo differential operators with negative definite symbols of variable order.* (submitted)

[36] W. Hoh, *Un calcul symbolique pour des opérateurs pseudo-différentiels engendrant des semigroupes de Feller.* C.R. Acad. Sci. Paris, t.325, Série I (1997) 1121–1124.

[37] W. Hoh, *On perturbations of pseudo differential operators with negative definite symbol.* Preprint.

[38] W. Hoh, N. Jacob, *Some Dirichlet forms generated by pseudo differential operators.* Bull. Sc. math., 2^{e} série, 116 (1992) 383–398.

[39] W. Hoh, N. Jacob, *Pseudo differential operators, Feller semigroups and the martingale problem.* In: Stochastic processes and optimal control, Stochastics Monographs, vol.7, Gordon and Breach Science Publishers, Amsterdam (1993) 95–103.

[40] W. Hoh, N. Jacob, *On the Dirichlet problem for pseudodifferential operators generating Feller semigroups.* J. Funct. Anal. 137, 1 (1996) 19–48.

[41] N. Ikeda, S. Watanabe, *Stochastic Differential equations and Diffusion Processes.* North-Holland Mathematical Library, Vol. 24, North-Holland Publishing Company, Amsterdam - Oxford - New York and Kodansha Ldt., Tokyo, second edition (1985).

[42] N. Jacob, *Further pseudo differential operators generating Feller semigroups and Dirichlet forms.* Rev.Mat.Iberoamericana 9, No.2 (1993) 373-407.

[43] N. Jacob, *A class of Feller semigroups generated by pseudo differential operators.* Math.Z. 215 (1994) 151–166.

[44] N. Jacob, *Pseudo-differential operators and Markov processes.* Mathematical Reseach, Vol.94, Akademie Verlag, Berlin (1996).

[45] N. Jacob, *Characteristic functions and symbols in the theory of Feller processes.* Potential Analysis 8 (1998) 61–68.

[46] N. Jacob, H.-G. Leopold, *Pseudo-differential operators with variable order of differentiation generating Feller semigroups.* Integr.Equat. Oper.Th. 17 (1993) 544–553.

[47] N. Jacob, R.L. Schilling, *Subordination in the sense of S. Bochner — An approach through pseudo differential operators.* Math. Nachr. 178 (1996) 199-231.

[48] N. Jacob, R.L. Schilling, *An analytic proof of the Lévy-Khinchin formula on* $\mathbb{R}^n$. To appear in: Publ. Math. Debrecen (1998).

[49] I. Karatzas, S.E. Shreve, *Brownian motion and stochastic calculus.* Graduate Texts in Mathematics 113, Springer Verlag, New York - Berlin - Heidelberg (1991).

[50] T. Kato, *Perturbation theory for linear operators.* Die Grundlehren der mathematischen Wissenschaften, Bd. 132, Springer Verlag, Berlin New York (1966).

[51] T. Komatsu, *Markov processes associated with certain integro-differential operators.* Osaka J. Math. 10 (1973) 271–303.

[52] T. Komatsu, *On the martingale problem for generators of stable processes with perturbations.* Osaka J.Math. 21 (1984) 113-132.

[53] T. Komatsu, *Pseudo-differential operators and Markov processes.* J.Math.Soc. Japan 36, No.3 (1984) 387-418.

[54] H. Kumano-go, *Pseudodifferential operators.* M.I.T. Press, Cambridge Mass. - London (1981).

[55] K. Kuratowski, C. Ryll-Nardzewski, *A general theorem on selectors.* Bull. Acad. Polon. Sci. Ser. Sci Math. Astronom. Phys. 12 (1965) 397–403.

[56] T.G Kurtz, *Semigroups of conditioned shifts and approximation of Markov processes.* Ann. Probab. 3 (1975) 618–642.

[57] N.S. Landkof, *Foundations of modern potential theory.* Die Grundlehren der mathematischen Wissenschaften, Bd. 180, Springer Verlag, Berlin - Heidelberg - New York (1972).

[58] H.-G. Leopold, *On a class of function spaces and related pseudo-differential operators.* Math. Nachr. 127 (1986) 65–82.

[59] H.-G. Leopold, *On Besov spaces of variable order of differentiation.* Z. Anal. Anwendungen 8 (1989) 69–82. Integr.Equat. Oper.Th. 17 (1993) 544–553.

[60] H.-G. Leopold, *On function spaces of variable order of differentiation.* Forum Math. 3 (1991) 69–82.

[61] J.P. Lepeltier, B. Marchal, *Problème des martingales et équations différentieles stochastiques associées à un opérateur intégro-différentiel.* Ann. Inst. Poincaré, Vol.XII, no.1 (1976) 43–103.

[62] I.S. Louhivaara, C.G. Simader, *Über koerzitive lineare partielle Differentialoperatoren: Fredholmsche verallgemeinerte Dirichletprobleme und deren Klasseneinteilung*. Complex analysis and its applications. Collect. Artic., Steklov Math Inst. (1978) 342–345.

[63] I.S. Louhivaara, C.G. Simader, *Fredholmsche verallgemeinerte Dirichletprobleme für koerzitive lineare partielle Differentialoperatoren*. Proc. Rolf Nevanlinna Symp. on complex analysis, Silivri 1976, Publ.Math. Res.Inst. 7 (1978) 47–57.

[64] K. Kikuchi, A. Negoro, *On Markov process generated by pseudodifferential operator of variable order*. Osaka J. Math 34 (1997) 319–335.

[65] Zh.-M. Ma, L. Overbeck, M. Röckner, *Markov processes associated with semi-Dirichlet forms*. Osaka J. Math. 32 (1995) 97–119.

[66] Zh.-M. Ma, M. Röckner, *Dirichlet forms*. Universitext, Springer Verlag, Berlin (1992).

[67] M. Nagase, *On the algebra of a class of pseudo-differential operators and the Cauchy problem for parabolic equations*. Math.Jap. 17, No.2 (1972) 147–172.

[68] J. Nash, *Continuity of solutions of parabolic and elliptic equations*. Amer. J. Math. 80 (1958), 931–954.

[69] A. Negoro, *Stable-like processes. Construction of the transition density and behavior of sample paths near* $t = 0$. Osaka J. Math. 31 (1994) 189–214.

[70] A. Negoro, M. Tsuchiya, *Stochastic processes and semigroups associated with degenerate Lévy generating operators*. Stoch. and Stoch.Rep. 26 (1989) 29-61.

[71] O. Okitaloshima, J. van Casteren, *On the uniqueness of the martingale problem* . Intern. J. Math. 7 no.6 (1996) 775–810.

[72] Y. Oshima, *Lectures on Dirichlet spaces*. Universität Erlangen-Nürnberg, Erlangen-Nürnberg (1988).

[73] S. Port, C. Stone, *Infinitely divisible processes and their potential theory. Part I.* Ann. Inst. Fourier 21,2 (1971) 157–275.

[74] Ph. Protter, *Stochastic integration and differential equations*. Application of Mathematics, vol. 21, Springer Verlag, Berlin (1990).

[75] M. Rogalski, *Le théorème de Lévy-Khincin*. Sém. Choquet (Initiation à l'Analyse) 3e année (1963/63) exposé 2.

[76] J.P. Roth, *Opérateurs dissipatifs et semigroupes dans les espaces de fonctions continues*. Ann. Inst. Fourier, Grenoble, Vol.26,4 (1976), 1–97.

[77] R.L. Schilling, *Comparable processes and the Hausdorff dimension of their sample paths*. Stoch. and Stoch. Rep. 57 (1996) 89–110.

[78] R.L. Schilling, *On Feller processes with sample paths in Besov spaces*. Math. Ann. 309 (1997) 663–675.

[79] I.J. Schoenberg, *Metric spaces and positive definite functions*. Trans. Amer. Math. Soc. 44 (1938) 522–536.

[80] A.V. Skorohod, *Studies in the theory of random processes*. Addison-Wesley, Reading Mass. (1965).

[81] D.W. Stroock, *Diffusion processes associated with Lévy generators*. Z. Wahr. verw. Geb. 32 (1975) 209-244.

[82] D.W. Stroock, S.R.S. Varadhan, *Diffusion processes with continuous coefficients, I,II*. Commun. Pure Appl. Math., Vol. XXII (1969) 345-400, 479-530.

[83] D.W. Stroock, S.R.S. Varadhan, *Multidimensional diffusion processes*. Grundlehren der mathematischen Wissenschaften 233, Springer Verlag, Berlin - Heidelberg - New York (1979).

[84] H. Tanabe, *Equations of evolution*. Monographs and Studies in Mathematics, Vol. 6, Pitman, London - San Francisco - Melbourne (1979).

[85] M. Taylor, *Pseudodifferential operators*. Princeton University Press, Princeton (1981).

[86] M. Tsuchiya, *Lévy measure with generalized polar decomposition and the associated SDE with jumps*. Stoch. and Stoch.Rep. 38 (1992) 95-117.

[87] A. Unterberger, *Sobolev spaces of variable order and problems of convexity for partial differential operators with constant coefficients*. Astérisque 2–3 (1973) 325–341.

[88] A. Unterberger, J. Bokobza, *Sur une généralisation des opérateurs de Calderon-Zygmund et des espaces* H^s. C.R.Acad. Sci. Paris (Ser. A) 260 (1965), 3265–3267.

[89] A. Unterberger, J. Bokobza, *Les opérateurs pseudodifférentiels d'ordre variable*. C.R.Acad. Sci. Paris (Ser. A) 261 (1965), 2271–2273.

[90] J. van Casteren, *Generators of strongly continuous semigroups*. Research Notes in Mathematics, Vol. 115, Pitman Publishing Inc., Marshfield MA - London (1985)

[91] J. van Casteren, On martingales and Feller semigroups. Results in Mathematics 21 (1992) 274–288

[92] N.Th. Varopoulos, *Hardy-Littlewood theory for semigroups*. J. Funct. Anal. 63 (1985) 240–260.

[93] N.Th. Varopoulos, L. Saloff-Coste, T. Coulhon, *Analysis and geometry on groups*. Cambridge Tracts in Mathematics, Vol. 100, Cambrigde University Press, Cambridge (1992).

[94] M.I. Višik, G.I. Eskin, *Elliptic convolution equations in a bounded region and their applications*. Russian Math. Surveys 22 (1967) 13–75.

[95] M.I. Višik, G.I. Eskin, *Convolution equations of variable order*. Trans. Moskow Math. Soc. 16 (1968) 27–52.

[96] K. Yosida, *Functional analysis*. Die Grundlehren der mathematischen Wissenschaften 123, Springer Verlag, Berlin - Heildelberg - New York (1980).