The Feynman graph representation of convolution semigroups and its applications to Lévy statistics

Hanno Gottschalk\textsuperscript{\textdagger}, Boubaker Smii\textsuperscript{\textdaggerdbl} and Horst Thaler\textsuperscript{\textdagger}

March 7, 2006

\textsuperscript{\textdagger}: Département des Mathématiques, Université de Tunis El Manar
\textsuperscript{\textdaggerdbl}: Institut für angewandte Mathematik, Rheinische Friedrich-Wilhelms-Universität Bonn

\textbf{Abstract:} We consider the Cauchy problem for a pseudo differential operator which has a translation invariant and analytic symbol. For a certain set of initial conditions, a formal solution is obtained by a perturbative expansion. The so-obtained series can be re-expressed in terms of generalized Feynman graphs and Feynman rules. The logarithm of the solution then can be represented by a series containing the connected Feynman graphs, only. Under some conditions, it is shown that the formal solution uniquely determines the real solution by the means of Borel transforms. The formalism is then applied to probabilistic Lévy distributions. Here, the Gaussian part of such a distribution is re-interpreted as a initial condition, and a large diffusion expansion for Lévy densities is obtained. It is outlined, how this expansion can be used in statistical problems that involve Lévy distributions.

\textbf{Key words:} Convolution semigroups, Feynman graphs and -rules, Borel summability, Lévy distributions, maximum likelihood principle.

\textbf{MSC (2000):} 60J35, 60J75, 62E20, 82B28

\textbf{Contents}

1 Introduction and overview 2

2 Generalized Feynman graphs and the convolution semigroup 4
   2.1 A solution in the sense of formal series .......................... 4
   2.2 Expansion in the operator’s coefficients ................................ 7
   2.3 The Feynman graph representation ...................................... 9
   2.4 The logarithm of the solution .......................................... 13
   2.5 Borel summability .................................................. 15
1 Introduction and overview

It is a well-known problem, that there is no closed formula for the transition densities of general Lévy processes. In fact, though there are many examples of Lévy laws where the density is known, see e.g. [2, 8, 17], a general formula, even in the sense of series expansions, seems to be missing. Small time expansions for pure jump Lévy densities have been considered e.g. in [3, 11]. But such expansions require the knowledge of convolution powers of the Lévy measures which in most cases will be hard to calculate analytically. Expansions are available for stable laws, cf. [17, 19, 20] and references therein.

It seems that there is no expansion formula for Lévy densities in the case where the Lévy process has moments of all orders that is simple enough to be applicable in statistics. In this work we provide such a formula in the sense of Borel summable asymptotic expansions in two cases: Either the Lévy process has to be started in a initial distribution that is given by a density of special form, or it can be started in a point but has to have a (large) diffusive part. The techniques employed stem from the renormalization group of statistical mechanics [16] and the recent proposal of a generalized Feynman graph calculus, [5, 6]. Some of our ideas are similar to recent work of physicists motivated by Feynman path integration, cf. [8, Chapter 20.4.6] for a large diffusion expansion for one-dimensional and symmetric Lévy densities, but non symmetric densities on \( \mathbb{R}^d \), Borel summability and the graphic representation have not been considered before.

Most of our considerations are not necessarily restricted to Lévy processes but hold for fairly general convolution semigroups. We study the Cauchy problem of the partial (pseudo-) differential equation

\[
\begin{aligned}
\frac{\partial}{\partial t} \Phi_t(\phi) & = \Psi(\nabla) \Phi_t(\phi) & (t, \phi) \in ]0, \infty[ \times \mathbb{R}^d \\
\Phi_0(\phi) & = f(\phi) & \phi \in \mathbb{R}^d
\end{aligned}
\tag{1}
\]

for a given initial state \( f : \mathbb{R}^d \rightarrow \mathbb{C} \) and a (pseudo-) differential operator with constant coefficients \( \Psi(\nabla) \).

We consider the case where is the symbol \( \Psi \) of the (pseudo-) differential operator, i.e. has an expansion

\[
\Psi(\xi) = \sum_{n=0}^{\infty} (C^{(n)}(\xi) \circ \cdots \circ \xi) = \sum_{n=0}^{\infty} \sum_{X_1, \ldots, X_n=1}^{d} C^{(n)}_{X_1 \cdots X_n} \xi X_1 \cdots \xi X_n ,
\tag{2}
\]
\( \xi = (\xi_1, \ldots, \xi_d) \in \mathbb{R}^d \), that converges on some neighborhood of 0 and \( f \) is of the form

\[
f = e^{-V}, \quad V(\phi) = \sum_{p=0}^{\tilde{p}} (\lambda^{(p)} , \phi^{\otimes p})_p = \sum_{p=0}^{\tilde{p}} \sum_{X_1, \ldots, X_p=1}^{d} \lambda^{(p)}_{X_1 \cdots X_p} \phi_{X_1} \cdots \phi_{X_p}
\]

for \( \phi = (\phi_1, \ldots, \phi_d) \in \mathbb{R}^d \), where \( \tilde{p} \) is even and \( (\lambda^{(p)} , \phi^{\otimes p})_p = \sum_{X_1, \ldots, X_p=1}^{d} \lambda_{X_1 \cdots X_p} \phi_{X_1} \cdots \phi_{X_p} > 0 \) for \( \phi \neq 0 \). We obtain a formal expansion of the solution of (1) in powers of \( \lambda_p \) that can be represented in terms of generalized Feynman graphs \([5, 6]\). Here \( C^{(p)} D_{X_1 \cdots X_p} \) and \( \lambda^{(p)} \) \( (\lambda^{(p)}_{X_1 \cdots X_p}) \) are symmetric tensors of \( p \)-th degree over \( \mathbb{R}^d \) (their components, respectively). \( \langle \cdot, \cdot \rangle_p \) is the canonical scalar product on \( (\mathbb{R}^d)^{\otimes p} \).

E.g. in the case where \( \lambda^{(p)} = 0 \) unless \( p = 2 \) (this will be the most interesting case) and the symbol \( \Psi(\xi) \) fulfills\(^1\) \( C^{(1)} = 0 \), the lowest orders in this expansion are

\[
\Phi_t(\phi) = 1 - \phantom{\text{1}} - \phantom{\text{2}} + \frac{1}{2} \left[ \text{3rd order \& higher in } \lambda^{(2)} \right]
\]

with each graph stands for a numerical expression that can be calculated in terms of the basic parameters \( C^{(n)}_{X_1 \cdots X_n} \), \( \lambda^{(p)}_{X_1 \cdots X_p} \), \( \phi \) and \( t \) by an easy algorithm ("Feynman rule"), cf. the following section.

This solution is the well-known graphical expansion for the renormalization group equation \([10]\) for the case when \( \Psi(\nabla) \) is a second order partial differential operator (e.g. the Laplacian) and \( \Phi_t \) hence is the heat equation. The recently discovered generalized Feynman graph calculus allows the generalization of this technique to a large class of symbols. This generalization is carried through in the present article, having in mind in particular Lévy processes with generators given by symbols of the form

\[
\Psi(-i\xi) = i \langle a, \xi \rangle - \langle \xi, D \xi \rangle + z \int_{\mathbb{R}^d \setminus \{0\}} (e^{i \langle \phi, \xi \rangle} - 1) \ dr(\phi),
\]

which lead to equations \([8]\) in the following jump-diffusion form (Lévy-Itô-decomposition)

\[
\frac{\partial \Phi_t(\phi)}{\partial t} = - \sum_{X=1}^{d} a_X \frac{\partial}{\partial \phi_X} \Phi_t(\phi) + \sum_{X_1, X_2=1}^{d} D_{X_1 X_2} \frac{\partial^2}{\partial \phi_{X_1} \partial \phi_{X_2}} \Phi_t(\phi) \\
+ z \int_{\mathbb{R}^d \setminus \{0\}} [\Phi_t(\phi + \varphi) - \Phi_t(\phi)] \ dr(\varphi).
\]

Here \( a \in \mathbb{R}^d \), \( D \) is a real and positive semidefinite \( d \times d \) matrix, \( z \geq 0 \) and \( r \) is a probability measure on \( \mathbb{R}^d \setminus \{0\} \) s.t. its Fourier transform is entire analytic. The well-known interpretation of \( a \) is the drift vector, \( D \) determines the diffusion part, whereas \( z \) and \( r \) give the frequency and distribution of jumps, respectively.

The restriction of the set of initial conditions in \([8]\) in many cases is unwanted. It can be circumvented by splitting the symbol \([5]\) into a pure jump and a pure diffusion part. After

\(^1\)This assumption is without loss of generality, since a non zero \( C^{(1)} \) can be compensated by a shift of \( \phi \).
some time, the transition kernel of the heat semigroup generated by the diffusive part alone is of course a Gaussian density, which, up to normalization, is of the form given in (3). This simple observation makes it possible to develop a perturbative calculus for densities of Lévy processes starting in a point in the limit of large diffusion. A part from this assumption, the perturbative formalism developed in this article is pretty universal and we therefore hope that it can contribute to a better handling of Lévy distributions in statistics.

The article is organized as follows: In Section 2 we develop a perturbative formalism for the solution of equation (1). In Subsection 2.1 we prove the differentiability of the solution $\Phi_t$ in the parameters $\lambda^{(p)}$ that define the initial condition and we conclude that the so-obtained series is asymptotic. Section 2.2 gives an expansion of the perturbation series in the coefficients of $\Psi$ which in the case of Lévy processes are just the cumulants. The re-arrangement in terms of a generalized kind of Feynman graphs is presented in Section 2.3. In the following Subsection 2.4 we obtain an expression for the logarithm of the solution $\Phi_t$ in terms of connected Feynman graphs. This does not only reduce the number of Feynman graphs that one has to calculate, but is also of importance in connection with applications to the maximum likelihood method. A short regression on Borel summability of the solution for the case that the logarithm of the initial condition, $V$, is quadratic concludes this Section (Subsect. 2.5).

In Section 3 we apply the general results to the special case of Lévy distributions. Subsection 3.1 presents the large diffusion expansion of Lévy densities. In Section 3.2 we illustrate our method by a second order calculation of the (non-normalized) Lévy distributions using Padé resummation. The result is used for a calculation of densities and quantiles that we compared with the Monte Carlo simulation of a Lévy process with compound or pure Poisson jump part. In Section 3.3 we describe the strategies how to apply the results of this work to the maximum likelihood estimation of parameters and we give an activity estimate to second order in $\beta = \lambda^{(2)}$ as an illustration.

Some technical stuff that presumably is standard to some part of the readers, but possibly not to some other part, is referred to the Appendices A–C.

2 Generalized Feynman graphs and the convolution semigroup

2.1 A solution in the sense of formal series

After the previous rough description, let us now fix the mathematical details. Let the pseudodifferential operator $\Psi$ fulfill the following requirements:

P1) The associated symbol $\Psi(-i\xi)$ is analytic for $\xi$ in a neighborhood of 0;

P2) $\Re \Psi(-i\xi) \leq c \log(1 + \|\xi\|)$; $\xi \in \mathbb{R}^d$, with $\Re$ the real part.

P3) $\forall \alpha \in \mathbb{N}_0^d \exists N_\alpha \in \mathbb{N}, b_\alpha > 0$ such that $|D_{\alpha} \Psi(-i\xi)| \leq b_\alpha (1 + \|\xi\|)^{N_\alpha} \forall \xi \in \mathbb{R}^d$.

Here $D_{\alpha} h(\phi) = \frac{\partial^{\mid\alpha\mid}}{\partial^{\alpha_1}_{\xi_1} \cdots \partial^{\alpha_d}_{\xi_d}} h(\phi)$. We furthermore assume without loss of generality that $\Psi(0) = 0$. If this is not the case and $\Psi(0) = b$ (i.e. if the associated Lévy process undergoes killing or
creation with rate \( b \), then the solution \( \Phi_t \) of equation (11) can be obtained from the solution \( \tilde{\Phi}_t \) of the equation (11) with \( \Psi \) replaced with \( \tilde{\Psi} = \Psi - \Psi(0) \) setting \( \Phi_t(\phi) = e^{t\Psi(0)}\tilde{\Phi}_t(\phi) \).

We take specific initial conditions that fulfill

I1) \( f(\varphi) = e^{-V(\varphi)} \);

I2) \( V \) is a polynomial as given in equation (2).

Certainly, the restriction in the initial conditions by I1–2) is disturbing. Nevertheless, I1–2) are general enough to determine the convolution kernel \( \nu_t \) completely, as one can construct approximating sequences of the Dirac delta distribution in that class, e.g. take \( V(\varphi) = \frac{\lambda}{2}||\varphi||^2 - \frac{d}{2}\log(2\pi\epsilon) \), \( \epsilon > 0 \).

Let \( \mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d) \) be the Fourier transform on the space of Schwartz test functions over \( \mathbb{R}^d \), i.e.

\[
\mathcal{F}(f)(\xi) = \int_{\mathbb{R}^d} f(\varphi) e^{i\langle \xi, \varphi \rangle} \, d\varphi, \quad f \in \mathcal{S}(\mathbb{R}^d).\tag{7}
\]

We denote the space of tempered distributions, i.e. the topological dual of the Schwartz space \( \mathcal{S}(\mathbb{R}^d) \), by \( \mathcal{S}'(\mathbb{R}^d) \). By the duality given by the \( L^2 \) scalar product on \( \mathbb{R}^d \) with respect to the Lebesgue measure, \( \mathcal{S}(\mathbb{R}^d) \) is densely embedded into \( \mathcal{S}'(\mathbb{R}^d) \) and \( \mathcal{F} \) extends uniquely to the latter space by continuity w.r.t. the weak topology \( \mathcal{S}' \). We denote this extension by the same symbol \( \mathcal{F} \).

By condition P2), \( e^{t\Psi(-i\xi)} \) is polynomially bounded for all \( t \), i.e. \( |e^{t\Psi(-i\xi)}| \leq (1 + ||\xi||)^ct \), hence \( e^{t\Psi(-i\xi)} \) is in \( \mathcal{S}'(\mathbb{R}^d) \) for all \( t \geq 0 \) and \( c > 0 \) as in that condition. Furthermore, by conditions I1) and I2) – the positivity condition on \( \lambda(\beta) \) in particular – \( f_\beta = e^{-\beta V} \) is in the test function space \( \mathcal{S}(\mathbb{R}^d) \) for \( \beta > 0 \). Hence for \( f = f_1 \), the convolution \( \Phi_t(\phi) = \nu_t * f(\phi) = \langle \nu_t, f_\phi \rangle \) is well-defined, where \( \nu_t \in \mathcal{S}'(\mathbb{R}^d) \) is the inverse Fourier transform of \( e^{t\Psi(-i\xi)} \) and \( f_\phi(\varphi) = f(\varphi - \phi) \).

**Lemma 2.1.** \( \Phi_t \) is the (unique) solution of the Cauchy problem (4).

**Proof.** The proof is a standard argument based on the Fourier transform of tempered distributions. For the convenience of the reader, the details are given in Appendix A.1. □

Following Poincaré, a formal power series \( \sum_{m=0}^\infty \beta^m a_m \) is called an asymptotic series for the function \( h(\beta), \beta > 0 \), at \( \beta = 0 \), if for \( N \in \mathbb{N} \)

\[
\lim_{\beta \searrow 0} \beta^{-N} \left| h(\beta) - \sum_{m=0}^N \beta^m a_m \right| = 0.\tag{8}
\]

Here we want to expand the solution \( \Phi_t \) in powers of \( V \). In many cases, this expansion is not convergent. We however show that such an expansion is asymptotic. To this aim let, for \( \beta > 0 \), \( \Phi_t^\beta = \nu_t * f_\beta \). Expansion in powers of \( V \) then means to expand in the auxiliary parameter \( \beta \) and then set \( \beta = 1 \). By definition, the series in \( V \) is asymptotic to \( \Phi_t \), if and only if the \( \beta \) expansion exists and is asymptotic to \( \Phi_t^\beta \).

The **moments** of the distribution \( \nu_t \) are defined as follows:

\[
\langle \varphi x_1 \cdots \varphi x_n \rangle_{\nu_t} = \frac{\partial^n}{\partial x_{1} \cdots \partial x_{n}} e^{t\Psi(\xi)}|_{\xi=0}, \quad n \in \mathbb{N}, \ x_1, \ldots, x_n \in \{1, \ldots, d\}. \tag{9}
\]

For a polynomial \( P \) in \( \varphi \), the expression \( \langle P(\varphi) \rangle_{\nu_t} \) is defined by linearity of the moments in the polynomial’s coefficients.
Proposition 2.2. The $V$-expansion of $\Phi_t$ is given by

$$\Phi_t(\phi) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \langle V_{\partial}^m(\phi) \rangle_{\nu_t}, \quad V_\phi(\phi) = V(\phi - \phi).$$

Equation (10) has to be understood in the sense that the right hand side gives the asymptotic series for the left hand side.

Proof. The statement is evident in the sense of formal power series. That the right hand side is asymptotic follows from the fact that $\Phi_t^\beta(\phi)$ is in $C^\infty((0,\infty))$ with respect to $\beta$ and for any $m \in \mathbb{N}$, $\frac{\partial^m}{\partial \beta^m} \Phi_t^\beta(\phi)$ can be continuously extended to $\beta = 0$ such that $\lim_{\beta \to 0} \frac{\partial^m}{\partial \beta^m} \Phi_t^\beta(\phi) = (-1)^m \langle V_{\partial}^m(\phi) \rangle_{\nu_t}$, cf. Appendix A.1 for the details. A posteriori we conclude that the expansion of $\Phi_t^\beta$ is $C^\infty([0,\infty))$ in $\beta$ and apply Taylor’s lemma to conclude the argument.

In general the asymptotic expansion is not convergent. This applies in particular to $\Psi(-i\xi)$ given in (5), where $\Psi$ is the generator of a Lévy process. We will come back to this point in Section 2.5, where Borel summability is proven in an important case. By the fact that we have an entire analytic Fourier transform of $\Phi_t^\beta$, the expansion is asymptotic, it is already clear that the first $N$ terms in (10) give a good approximation to the actual solution for $V$ small and $\phi$ not too large, an error estimate is provided by Taylor’s formula. In order to extract reliable data also for larger $V$, resummation techniques have to be applied. Some first steps in this direction are given in Section 3.2.

Obviously, all partial differential operators of the form $\Psi(V) = \sum_{n=0}^{\hat{n}} \sum_{X_1,\ldots,X_n=1}^{\hat{d}} C_{X_1,\ldots,X_n} \times \frac{\partial^m}{\partial \phi X_1 \cdots \partial \phi X_n}$ of even order $\hat{n}$ with $(-1)^{\hat{n}/2}(C^{(n)},\psi^{(n)})_{\hat{n}} > 0$ for $\xi \neq 0$ or of odd order $\hat{n}$ with $C^{(n)}$ real and $(-1)^{(\hat{n}-1)/2}(C^{(\hat{n}-1)},\psi^{(\hat{n}-1)}_{\hat{n}-1}) > 0$ for $\xi \neq 0$ fulfill the conditions $P1$–$P3$. This of course includes the heat equation ($\hat{n} = 2$). More interestingly, not only local generators can be treated, but also generators of Lévy processes that have a symbol that is analytic at 0:

Proposition 2.3. The generator (5) with $r$ a probability measure with analytic Fourier transform at 0 fulfills the conditions $P1$–$P3$.

Proof. P1), analyticity at 0, is immediate from the conditions. The real part of (5) is obviously bounded from above by $2\xi$, which establishes P2). To see P3), let us note that the partial derivatives of the third part in (5) are of the form $\int_{\mathbb{R}^d \setminus \{0\}} \varphi^\alpha e^{i\xi \cdot \varphi} d\rho(\varphi)$. As $r$ has analytic Fourier transform at zero, all moments of $r$ exist. This implies differentiability everywhere and that $D_\alpha \Psi(-i\xi)$ is even uniformly bounded in $\xi \in \mathbb{R}^d$.

For the sake of mathematical completeness, we also give a sufficient condition for the convergence of the expansion (10). Its practical value is however limited, as it excludes the most interesting examples.

Proposition 2.4. If the symbol $\Psi$ does not necessarily fulfill P3) but fulfills P3'): $\Psi(-i\xi)$ is entire analytic and $\Psi(-i\xi) \leq K(1 + \|\Im \xi\|_{\max})$, $\xi \in \mathbb{C}^d$, for some $K > 0$, then the expansion (10) is convergent.

Proof. The condition implies that $\nu_t$ has an entire analytic Fourier transform $|e^{i\Psi(\xi)}| \leq e^{Kt} e^{tK\|\Im \xi\|_{\max}}$. Application of the Paley–Wiener–Schwartz theorem implies that $\nu_t$ has support in a square of length $2tK$ centered at the origin and hence compact support. Let $\chi$ be a test function with compact support and $\chi|_{\text{supp} \chi} = 1$. Then $\Phi_t^\beta(\phi) = \nu_t * f_\beta(\phi) = \langle \nu_t, \chi \cdot e^{-\beta V_\phi} \rangle$. As the series expansion in $\beta, \beta \in \mathbb{C}$, of $e^{-\beta V_\phi}$ and all its derivatives converge compactly on $\text{supp} \chi$, the mapping $\mathbb{C} \ni \beta \to \chi \cdot e^{-\beta V_\phi} \in \mathcal{S}(\mathbb{R}^d)$ is entire analytic in $\beta$. Since $\nu_t \in \mathcal{S}'(\mathbb{R}^d)$, the same applies to $\Phi_t^\beta(\phi)$. 

H. Gottschalk, B. Smii and H. Thaler
2.2 Expansion in the operator’s coefficients

In this section we define the truncated moments of the convolution kernel \( \nu_t \) as partial derivatives of \( t\Psi \) – as the coefficients of \( \Psi \) times \( t \), in other words. A result known as the linked cluster theorem (sometimes also named the cumulant expansion) establishes a combinatorial connection between the moments \( \langle \varphi_{X_1} \cdots \varphi_{X_n} \rangle_{\nu_t} \) and the truncated moments \( \langle \varphi_{X_1} \cdots \varphi_{X_n} \rangle^T_{\nu_t} \). The asymptotic expansion of the solution of the Cauchy problem \( \Phi_t(\phi) \) can thus be directly expressed in terms of the coefficients of the symbol \( C^{(n)}_{X_1 \cdots X_n} \), the time \( t \) and the coefficients \( \lambda_{X_1 \cdots X_n}^{(p)} \) of the logarithm of the initial condition.

**Definition 2.5.** The truncated moments of the distribution \( \nu_t \) are defined as

\[
\langle \varphi_{X_1} \cdots \varphi_{X_n} \rangle^T_{\nu_t} = t \frac{\partial^n}{\partial \xi_{X_1} \cdots \partial \xi_{X_n}} \Psi(\xi)|_{\xi=0}, \quad n \in \mathbb{N}, \quad X_1, \ldots, X_n \in \{1, \ldots, d\}.
\]  

(11)

By (2) one obtains that the truncated moments are the coefficients of the (pseudo-) differential operator \( \Psi \), i.e.

\[
\langle \varphi_{X_1} \cdots \varphi_{X_n} \rangle^T_{\nu_t} = t C^{(n)}_{X_1 \cdots X_n}.
\]  

(12)

The following classical theorem gives the relation between the ordinary moments \( \langle \varphi_{X_1} \cdots \varphi_{X_n} \rangle_{\nu_t} \) and the truncated ones introduced above.

**Theorem 2.6.** (Linked cluster) For \( n \in \mathbb{N} \) we have

\[
\langle \varphi_{X_1} \cdots \varphi_{X_n} \rangle_{\nu_t} = \sum_{I \in \mathcal{P}(\{1, \ldots, n\})} \prod_{I_k \neq \emptyset} \left( \prod_{j \in I_k} \varphi_{X_j} \right)^T_{\nu_t}
\]  

(13)

where, for a finite set \( A \), \( \mathcal{P}(A) \) stands for the set of all partitions \( I \) of \( A \) into nonempty disjoint subsets \( \{I_1, \ldots, I_k\} \).

For the proof we refer to the literature, e.g. [5, 15].

**Proposition 2.7.** The solution of the Cauchy problem \( \Phi_t(\phi) \), \( \phi \in \mathbb{R}^d \), has the following asymptotic expansion:

\[
\Phi_t(\phi) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \sum_{p_1, \ldots, p_m=0}^{\hat{p}} \sum_{K \subseteq \Omega(p_1, \ldots, p_m)} \prod_{I \in \mathcal{P}(\{p_1, \ldots, p_m\}) \setminus K} t C_{X_1^{p_1} \cdots X_{p_1}^{1}} \cdots t C_{X_1^{p_m} \cdots X_{p_m}^{1}}
\]  

(14)

Here \( \Omega(p_1, \ldots, p_m) = \bigcup_{n=1}^{m} \{(1, n), \ldots, (p_n, n)\} \) and \( C_{I_l} = C^{(q)}_{X_{a_1}^{b_1} \cdots X_{a_q}^{b_q}} \) for \( I_l = \{(a_1, b_1), \ldots, (a_q, b_q)\} \).
Proof. The proof is by a straightforward calculation using (11) and (12)

\[
\Phi_t(\phi) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \left( V^{(m)}(\varphi) \right)_{\nu_2}
\]

\[
= \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \sum_{p_1,\ldots,p_m=0}^{\tilde{p}} \prod_{k=1}^{d} \lambda^{(p_k)}_{X^1_{p_k} \cdots X^d_{p_k}} \nu_2
\]

\[
\times \sum_{K \subset \Omega(p_1,\ldots,p_m)} \prod_{(s,q) \in K} (-\varphi X^s_q) \left( \prod_{(r,t) \in \Omega(p_1,\ldots,p_m) \setminus K} \varphi X^r_t \right)_{\nu_2}
\]

\[
= \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \sum_{p_1,\ldots,p_m=0}^{\tilde{p}} \prod_{k=1}^{d} \lambda^{(p_k)}_{X^1_{p_k} \cdots X^d_{p_k}} \nu_2
\]

\[
\times \sum_{K \subset \Omega(p_1,\ldots,p_m)} \prod_{(s,q) \in K} (-\varphi X^s_q) \sum_{l \in \mathcal{P}(\Omega(p_1,\ldots,p_m) \setminus K)} \prod_{l=1}^{k} tC_{l_i}
\]

\[
(15)
\]

Though the expansion is explicit, it is also rather messy. To clean it up, we introduce a graphical calculus in the next subsection.

If \( d = 1 \), the expression (11) simplifies considerably, one can drop all subscripts \( X_1, \ldots \) and one obtains the asymptotic expansion

\[
\Phi_t(\phi) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \prod_{p_1,\ldots,p_m=0}^{\tilde{p}} \lambda^{p_1+\cdots+p_m}_{k=0} \prod_{l=1}^{p_1+\cdots+p_m-k} \frac{(-\phi)^k}{k!}
\]

\[
\times \sum_{q=1}^{\infty} \frac{t^q}{1+\sum_{l_1+\cdots+l_q=0}^{l_1+\cdots+l_q=0} h_q(l_1,\ldots,l_q) \prod_{s=1}^{q} C^{(l_s)}}
\]

\[
(16)
\]

Here \( h_q(l_1,\ldots,l_q) \) is a combinatorial factor given by the number of partitions of \( l_1 + \cdots + l_q \) objects in subsets with \( l_1, l_2, \ldots, l_q \) elements. Explicitly it can be calculated as follows: Let \( \{l_s : s = 1, \ldots, q\} = \{i_1, \ldots, i_s\} \) be the set of different values of the \( l_1, \ldots, l_q \) and let \( q_u = \sum_{n=1}^{\#}\left\{ n \in \mathbb{N} : l_n = i_u \right\}, u = 1, \ldots, s \). Then

\[
h_q(l_1,\ldots,l_q) = \left( \frac{l_1 + \cdots + l_q}{l_1 \cdots l_q} \right) \frac{1}{q_1! \cdots q_s!}
\]

\[
(17)
\]

The expressions (12) and (16) simplify further if certain coefficients \( C^{(n)} \) and \( \lambda^{(p)} \) vanish, e.g. because of symmetries. As this is easier to understand in terms of graphs, we postpone the discussion of this point to the following subsection.
In the case that \( f \) is a probability density and \( \Psi \) is of the form (5), it is clear that \( \Phi_t \), as a convolution of a probability measure with a probability density, is a probability density for all \( t \). Note that the first \( N \) terms of the expressions (14) or (16) give a polynomial in \( \phi \) and \( t \). Obviously, this plays havoc with normalization conditions. We thus have to keep in mind, that the approximation given by the first \( N \) terms in the perturbation series is only good for \( t \) and \( \phi \) sufficiently small. In Sections 2.4, 3.1 and 3.2 we will discuss strategies to deal with this problem.

On the other hand, the expansion in terms of polynomials in \( \phi \) makes it most simple to integrate the probability density and to get a formula for the distribution function. Taking e.g. the one dimensional formula (16), we obtain in the sense of asymptotic series

\[
P(a < Z_t \leq b) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \sum_{p_1,\ldots,p_m=0}^{\tilde{p}} \prod_{l=1}^{m} \lambda(p_l) \sum_{k=0}^{p_1+\cdots+p_m-k} \left( \sum_{l_1+\cdots+l_q=p_1+\cdots+p_m-k} \prod_{s=1}^{q} C(l_s) \right) \frac{a^{k+1} - b^{k+1}}{k+1}
\]

where \( Z_t \) is a random variable with probability density function \( \Phi_t \), i.e. a Lévy process with generator \( \Psi \) and initial distribution \( f \). Again, the first \( N \) terms of the right hand side give a good approximation for \( a, b, t \) and \( \lambda(p) \) sufficiently small.

### 2.3 The Feynman graph representation

In this subsection we represent the solution of the Cauchy problem (14) as a sum over generalized Feynman graphs that are evaluated by a simple algorithm called ”Feynman rules”. In Physics the result is known as the Feynman graph representation for amputated Green’s functions for the case when the generator is some kind of Laplacian, i.e. \( \Psi(\nabla) = \sum_{x_1,x_2=1}^{d} X_{x_1,x_2} \frac{\partial^2}{\partial \phi_{x_1} \partial \phi_{x_2}} \). Combining this with the generalized Feynman graph calculus developed in [5, 6] we obtain the first main result of this article.

A graph is a geometrical object which consists of vertices and edges, i.e lines that connect exactly two vertices. Here we need graphs with non directed edges that have vertices of different types. Vertices can be distinguishable or non distinguishable, depending on their type, and they can have distinguishable or non distinguishable legs. We use the term ”leg” for the part of the edge meeting the vertex. All these graph theoretic notions are properly defined in Appendix A.2. But for a first reading an intuitive comprehension is sufficient.

Our expansion is represented as a sum over generalized Feynman graphs. They have the following properties:

**Definition 2.8.** A generalized amputated Feynman graph is a graph with three types of vertices, called inner full\(^2\), inner empty and outer empty vertices, respectively, see Table 1. By definition full vertices are distinguishable and have distinguishable legs whereas empty vertices are non distinguishable and have non distinguishable legs\(^3\). Outer empty are met by one edge only.

\(^2\)In physics this type is called interaction vertex.
\(^3\)Empty inner vertices can be distinguished from empty outer vertices, however.
Table 1: Different types of vertices.

<table>
<thead>
<tr>
<th></th>
<th>Full</th>
<th>Empty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inner</td>
<td>⨿</td>
<td>○</td>
</tr>
<tr>
<td>Outer</td>
<td>×</td>
<td>⊗</td>
</tr>
</tbody>
</table>

Edges are non directed and connect full and empty (inner and outer) vertices, but never connect two full or two empty vertices.

Let \( m \in \mathbb{N}_0 \). The set of generalized amputated Feynman graphs with \( m \) inner full vertices with \( p_1, \ldots, p_m \) the number of legs of the inner full vertices such that \( p_j \leq \bar{p} \) and \( \lambda(p_j) \neq 0 \), \( j = 1, \ldots, m \), is denoted by \( \pF(m) \).

The name generalized amputated Feynman graphs has been chosen to distinguish the graphs used here from classical Feynman graphs but also from the generalized Feynman graphs with outer full vertices used in \([5, 6]\) (this latter type gives the right combinatorics to calculate the moments of \( \Phi_t \)). To simplify the notions, we here use the term Feynman graph when we mean a generalized amputated Feynman graph and there is no danger of confusion.

Let \( m, p_1, \ldots, p_m \in \mathbb{N}_0 \). We construct a one to one correspondence between pairs \((K; I)\) where \( K \subseteq \Omega(p_1, \ldots, p_m) \) and \( I \in \mathcal{P}(\Omega(p_1, \ldots, p_m) \setminus K) \) and the Feynman graphs \( \pF(m, p_1, \ldots, p_m) \) with \( m \) inner full vertices with \( p_1, \ldots, p_m \) legs in the following way:

- For \( j = 1, \ldots, m \) label the legs of the \( j \)-th inner full vertex with \((1, j), \ldots, (p_j, j)\). In this way we obviously obtain a one to one correspondence between the legs of the inner full vertices and the elements in \( \Omega(p_1, \ldots, p_m) \).

We first consider the case where the Feynman graph \( G \) is given and we construct \((K, I)\):

- Define \( K \subseteq \Omega(p_1, \ldots, p_m) \) as the set corresponding (under the above bijection) to those legs of inner full vertices in the Feynman graph \( G \) that run into an outer empty vertex.

- Let \( k \) be the number of inner empty vertices in \( G \). For \( q = 1, \ldots, k \) let \( I_q \) be the subset in \( \Omega(p_1, \ldots, p_m) \) corresponding to those legs of inner full vertices that run into the \( q \)-th inner empty vertex. Let finally \( I = \{I_1, \ldots, I_k\} \) then evidently \( I \in \mathcal{P}(\Omega(p_1, \ldots, p_m) \setminus K) \).

Conversely let \( K \subseteq \Omega(p_1, \ldots, p_m) \) and \( I \in \mathcal{P}(\Omega(p_1, \ldots, p_m) \setminus K) \), \( I = \{I_1, \ldots, I_k\} \) be given. We construct a Feynman graph \( G \in \pF(m, p_1, \ldots, p_m) \)

- Draw \( m \) inner full vertices with \( p_1, \ldots, p_m \) legs and \( k \) inner empty vertices with \( \sharp I_1, \ldots, \sharp I_k \) legs and \( p = \sharp K \) outer empty vertices.

- Connect the legs of the \( q \)-th inner empty vertex with the legs of the inner full vertices corresponding to the elements in \( I_q \).

- Connect also the legs outer empty vertices with the legs of the inner full vertices corresponding to the elements of \( K \). The result is a Feynman graph \( G \in \pF(m) \).

For an illustration of the above construction cf. Figure 1. We have established the following result:
Feynman graph representation of semigroups

Lemma 2.9. Let \( m, p_1, \ldots, p_m \in \mathbb{N}_0 \). Then there exists a one to one correspondence between \( \{(K, I) : K \subseteq \Omega(p_1, \ldots, p_m), I \in \mathcal{P}(\Omega(p_1, \ldots, p_m) \setminus K)\} \) and \( \tilde{F}(m, p_1, \ldots, p_m) \), the set of Feynman graphs with \( m \) inner full vertices with \( p_1, \ldots, p_m \) legs.

The following algorithm in physics is named "Feynman rules" and gives a numerical value to each Feynman graph:

Definition 2.10. Let \( G \in \tilde{F}(m) \) such that \( p_1, \ldots, p_m \) are the numbers of legs of the inner full vertices of \( G \). Then, for \( \phi \in \mathbb{R}^d \) and \( t \in (0, \infty) \), \( V[G] = V[G](t, \phi) \) is a real number obtained by the following method\(^4\):

1. To each leg of an inner full vertex assign a value \( X_a = X_j^i \in \{1, \ldots, d\} \) with \( a \) the element in \( \Omega(p_1, \ldots, p_m) \) corresponding to that leg.

2. Multiply with a factor \( (-\phi X_b) \) each outer empty vertex where \( b \in \Omega(p_1, \ldots, p_m) \) is the element corresponding to a leg of the inner full vertex which is directly connected with that outer empty vertex.

3. For each inner empty vertex with \( l \) legs multiply with a factor \( tC_J = tC_{X_{a_1} \ldots X_{a_l}} \) where \( J = \{a_1, \ldots, a_l\} \subseteq \Omega(p_1, \ldots, p_m) \) is the subset corresponding to the \( l \) legs of inner full vertices that are directly connected to the inner empty vertex.

4. For the \( j \)-th inner full vertex, \( j = 1, \ldots, m \), multiply with a factor \( \lambda_{X_{k_j}^{i_1} \cdots X_{k_j}^{i_{l_j}}}^{(p_j)} \) and sum up over \( \sum_{X_{k_j}^{i_1} \cdots X_{k_j}^{i_{l_j}}}^{d} \).

\(^4\)We use the labeling of legs of inner full vertices by \( \Omega(p_1, \ldots, p_m) \) introduced above – but also any other labeling would do.
Combining Proposition 2.7, Lemma 2.9 and Definition 2.10 one thus gets the Feynman graph expansion of the convolution semigroup:

**Theorem 2.11.** The solution of the Cauchy problem (1) has the following asymptotic expansion:

\[
\Phi_t(\phi) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \sum_{G \in \mathcal{F}(m)} V[G](t, \phi).
\]

(19)

Apart from the more efficient notation that comes along with the use of Feynman graphs, Feynman graphs can be used to reduce the number of terms in the expansion. Note that the value \(V[G](t, \phi)\) attributed to the graph \(G\) only depends on the topological Feynman graph. Hence, one can change the sum in (19) into a sum over topological Feynman graphs by multiplying \(V[G](t, \phi)\) with a weight factor or multiplicity, i.e. with the number of Feynman graphs that give the same topological Feynman graph. Such multiplicities in low orders can be calculated by hand, see (4) for an example.

Equation (19) still is fairly general. Often, a number of Feynman graphs do not give contributions. E.g. in the case where \(\Psi(\xi) = \Psi(-\xi)\), one has \(C^{(n)} = 0\) for \(n\) odd and hence all Feynman graphs with at least one empty vertex with an odd number of legs can be omitted in the expansion.

Another simplification occurs, if the \(V(\varphi)\) is quadratic in \(\varphi\). Then, the only kind of inner full vertex that occurs is of the type \(\bullet\) and can thus be considered as a edge of new type. Note that the legs of the inner full vertex are distinguishable, hence this new type of edge is directed, which changes multiplicities. Finally, if \(\lambda_{X_1X_2} = \lambda \delta_{X_1X_2}\) is a multiple of the Kronecker symbol, the evaluation of a graph as e.g. on the right hand side of Fig. 2 can be simplified as follows:

**Corollary 2.12.** If \(V(\varphi)\) is quadratic, the set of Feynman graphs \(\mathcal{F}(m)\) of order \(m\) can be identified with the set of all graphs \(\mathcal{Q}(m)\) with two types of indistinguishable vertices, called outer empty and inner empty, with \(m\) directed edges such that an outer empty vertex is hit by exactly one edge.

If \(\lambda^{(2)}\) in addition is diagonal\(^5\), the value \(V[G](t, \phi)\) attributed to such a graph \(G \in \mathcal{Q}(m)\) can be calculated according to the following rules:

1. To each of the \(m\) edges assign a value \(X_a \in \{1, \ldots, d\}\), \(a = 1, \ldots, m\).
2. For each inner empty vertex with \(l\) legs multiply with a factor \(tC_{X_{a_1} \cdots X_{a_l}}\), where \(a_1, \ldots, a_l\) correspond to the numbers given to the edges that are connected to this vertex\(^6\).
3. For each edge add a factor \(\lambda\), alternatively add an over all factor \(\lambda^m\).
4. Sum up over all edges, i.e. perform the sum\(^7\) \(\sum_{X_1, \ldots, X_m = 1}^{d} X_1, \ldots, X_m = 1\).

As a side remark we notice that the combinatorics of the Feynman graph formalism is dimension independent. Only Feynman rules depend on dimension. The number of steps needed in the evaluation of a given Feynman graph grow like \(\sim d^m\) with the dimension, where \(m\) is the

\(^5\)If \(\lambda^{(2)}\) is symmetric, this can always be archived by choosing a suitable basis on \(\mathbb{R}^d\).

\(^6\)If an edge connects the vertex with itself, the corresponding index has to be repeated twice.

\(^7\)Of course, for \(d = 1\) the same simplifications as discussed at the end of Section 2.2 occur.
order of the graph. If one compares this with an evaluation of $\Phi_t$ by a multi-dimensional Fourier transform, the steps needed for the evaluation grows rather like $\sim m^d$ with $m$ a parameter standing, e.g., for the number of summands in an approximation of the one dimensional Fourier integrals by a sum. For large dimensions, the Feynman graph formalism thus promises better results than numerical evaluation of Fourier transforms. Since the Feynman graph formalism has its roots in infinite dimensions, this should not be a surprise.

2.4 The logarithm of the solution

In this subsection we obtain an asymptotic expansion of the logarithm of the solution in terms of connected Feynman graphs. The argument in many points follows [5].

A graph $G \in \bar{F}(m)$ is said to be connected if and only if each two vertices are connected by a walk passing through finitely many edges. We denote the collection of connected $m$-th order Feynman graphs by $\bar{F}_c(m)$.

On the level of partitions, connectedness can be expressed as follows:

**Definition 2.13.** Let $m, p_1, \ldots, p_m \in \mathbb{N}$ be given and $J_j = \{(j, 1), \ldots, (j, p_j)\}$, $j = 1, \ldots, m$. We say that the pair $(K, I)$ with $K \subseteq \Omega(p_1, \ldots, p_m)$, $I \in \mathcal{P}(\Omega(p_1, \ldots, p_m) \setminus K)$, $I = \{I_1, \ldots, I_k\}$, is connected, if it fulfills the following condition: $\forall 1 \leq i_1, \ldots, i_q \leq k$, $1 \leq q < k$ and $1 \leq j_1, \ldots, j_s \leq m$, $1 \leq s < m$ such that

$$\bigcup_{a=1}^{q} I_{i_a} = \bigcup_{a=1}^{s} (J_{j_a} \setminus K).$$  \hspace{1cm} (20)

Let $\mathcal{P}_c(\Omega(p_1, \ldots, p_m) \setminus K)$ be the collection of partitions $I \in \mathcal{P}(\Omega(p_1, \ldots, p_m) \setminus K)$ such that $(K, I)$ is connected.

**Lemma 2.14.** A graph $G \in \bar{F}(m)$ is connected if and only if the corresponding pair $(K, I)$ is connected.

**Proof.** We only sketch the idea of the proof: In terms of the corresponding Feynman graph $G$, the condition in Definition 5.1 means that one cannot find a true subset $Q_1, Q_2$ of empty and full vertices such that all edges from an empty vertex in $Q_1$ end up at a full vertex in $Q_2$ and vice versa. This is evidently equivalent to the connectedness of $G$. For details see [5].

**Definition 2.15.** Let $J_j$, $j = 1, \ldots, m$ be as in Definition 2.13, $\phi \in \mathbb{R}^d$, $X_a = X^l_a \in \{1, \ldots, d\}$ for $a = (j, l) \in \Omega(p_1, \ldots, p_m)$. We use the abbreviation $J_j^\phi = \prod_{a \in J_j} (\varphi X_a - \phi X_a)$, $j = 1, \ldots, m$.

The block truncated moment function $\left\langle \prod_{l=1}^{m} J_l^\phi \right\rangle^{(T)}_{\nu_t}$ of $\nu_t$ are recursively (in $m \in \mathbb{N}$) defined as follows:

$$\left\langle \prod_{l=1}^{m} J_l^\phi \right\rangle^{(T)}_{\nu_t} = \sum_{I \in \mathcal{P}(1, \ldots, m)} \prod_{l=1}^{k} \left\langle \prod_{q \in I_l} J_q^\phi \right\rangle^{(T)}_{\nu_t}$$  \hspace{1cm} (21)

The symbol $(T)$ in the definition means that each polynomial (in $\varphi$) $J_q^\phi$ in the combinatorics of truncation is treated as one object.
The one to one correspondence. Furthermore, the mappings are inverses of each other, which establishes constructions giving mappings between the index sets of the sums in (23) and (24) such that the \( l \) \( Q \) show that \( \{ J \} \) vertices corresponding to \( K, R \) We have to prove the equality between the right hand side of (23) and (24), i.e., we have to defining equation for the left hand side. We thus have to show that

\[ \langle J_1^{\phi} \cdots J_m^{\phi} \rangle_{\nu_1} \mathcal{T} = \sum_{K \subseteq \Omega(p_1, \ldots, p_m), \Omega(I_1, \ldots, I_k)} \prod_{l=1}^{K} (-\phi_{X_l}) \prod_{j \in K} (I_l)_T \quad (22) \]

Proof. To verify equation (22) we have to insert the right hand side of this equation into the defining equation for the left hand side. Thus, we have to show that

\[ \langle J_1^{\phi} \cdots J_m^{\phi} \rangle_{\nu_1} = \sum_{I \subseteq \Omega(p_1, \ldots, p_m)} \prod_{l=1}^{I} (-\phi_{X_l}) \prod_{j \in I} (I_l)_T \quad \nu_1 \]

On the other hand, \( \langle J_1^{\phi} \cdots J_m^{\phi} \rangle_{\nu_1} \) can be directly expanded in terms of truncated moments:

\[ \langle J_1^{\phi} \cdots J_m^{\phi} \rangle_{\nu_1} = \sum_{K \subseteq \Omega(p_1, \ldots, p_m) \cup \Omega(L, \ldots, L_k)} \prod_{l=1}^{K} (-\phi_{X_l}) \prod_{j \in K} (I_l)_T \quad \nu_1 \]

We have to prove the equality between the right hand side of (23) and (24), i.e., we have to construct a correspondence between the pairs \( (K, R) \) and the objects \( (I, (K_1, Q_1), \ldots, (K_k, Q_k)) \) indexing the sum in (23) such that their contributions to the sum are equal.

Let \( (R, K) \) be given. We say that \( R \) connects \( q \) and \( j \) in notation \( q \sim_R j \) if the full inner vertices corresponding to \( J_q \) and \( J_j \) respectively are connected in the generalized Feynman graph corresponding to \( R \). It is easy to see that \( R \) is an equivalence relation on \( \{1, \ldots, m\} \). Let \( I = \{I_1, \ldots, I_k\} \) be the equivalence class of \( \sim_R \), then \( I \in \mathcal{P}(\{1, \ldots, m\}) \). We set \( K_l = (\bigcup_{q \in I_l} J_q \cap K) \), \( l = 1, \ldots, k \), and \( Q_l = \{S \in R : S \subseteq \bigcup_{q \in I_l} J_q\} \). Using the equivalence relation \( \sim_R \), it is now easy show that \( Q_l \in \mathcal{P}(\bigcup_{q \in I_l} J_q) \), cf. [3] for the details.

The converse construction is trivial: Take \( K = \bigcup_{l=1}^{k} K_l \) and \( R = \bigcup_{l=1}^{k} Q_l \). These two constructions give mappings between the index sets of the sums in (23) and (24) such that the summands are equal. Furthermore, the mappings are inverses of each other, which establishes the one to one correspondence. \( \blacksquare \)

We have now completed the preparations to show the following theorem:

**Theorem 2.17.** In the sense of asymptotic series \( \log \Phi_t(\phi) \) has the following expansion in terms of connected Feynman graphs

\[ \log \Phi_t(\phi) = \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \sum_{G \in F_c(m)} \mathcal{V}[G](t, \phi) . \quad (25) \]

Proof. Note that by the ordinary linked cluster theorem (Theorem 2.6) the following holds in the sense of asymptotic series

\[ \log \Phi_t(\phi) = \log \langle \nu_t, e^{-V} \rangle = \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \langle V^{m}_\phi \rangle_{\nu_t} . \quad (26) \]
The assertion thus follows by application of Proposition 2.16 and Lemma 2.14.

The above result is useful in three respects:

If $\Psi$ is the symbol of a Lévy process and the initial condition $f = e^{-V}$ is a probability density, then the solution $\Phi_t(\phi)$ is a probability density for all $t$. We have seen in Proposition 2.7 and in equation (16) that the first $N$ summands of the asymptotic expansion for $\Phi_t(\phi)$ is a polynomial in $\phi$ and thus non normalizable. We however derived an asymptotic expansion for $\log \Phi_t(\phi)$ such that the first $N$ terms are a polynomial $P_N(\phi)$ in $\phi$ that goes to $-\infty$ for large $\phi$ (often this is the case for $N$ odd), we obtain $\Phi_t(\phi) \approx e^{-P_N(\phi)}$ as the solution in $N$-th order perturbation theory. While the large $\phi$ behavior of this approximate solution obviously is still problematic, one has at least gained normalizability.

Secondly, in statistical estimations by the maximum likelihood method one needs an explicit formula for $\log \Phi_t(\phi)$, which is just the sum over connected Feynman graphs, cf. Section 3.3.

Finally, the connectedness condition also reduces the number of terms in the expansion, e.g. equation (4), only the 2nd – 4th and the 7th term contribute to the logarithm of the solution.

2.5 Borel summability

In certain cases a function may be recovered from its asymptotic series expansion. This class of functions comprises the Borel summable functions. A sufficient condition for a function to equal its Borel sum is the version of Watson’s theorem by A. Sokal [15]:

**Theorem 2.18.** Let $h$ be analytic in the interior of the circle $C_R := \{ \zeta \in \mathbb{C} : \Re \zeta^{-1} > R^{-1} \}$ (cf. Figure 2 a)), such that

$$|h(\zeta) - \sum_{m=0}^{N-1} a_m \zeta^m| \leq A\sigma^N N! |\zeta|^N,$$

uniformly in $N$ and $\zeta \in C_R$, with constants $A, \sigma > 0$. Under this assumption one may show that $B(\tau) := \sum_{m=0}^{\infty} a_m \tau^m / m!$ converges for $|\tau| < 1/\sigma$ and has an analytic continuation to the striplike region $S_\sigma = \{ \tau \in \mathbb{C} : \text{dist}(\tau, \mathbb{R}_+) < 1/\sigma \}$ (cf. Figure 2 b)) obeying the bound

$$|B(\tau)| \leq K \exp(|\tau|/R),$$

uniformly in every $S_{\sigma'}$, with $\sigma' > \sigma$. Furthermore, $h$ can be represented by the absolutely convergent integral

$$h(\zeta) = \frac{1}{\zeta} \int_0^\infty e^{-\tau/\zeta} B(\tau) \, d\tau.$$

Conversely, if $B(\tau)$ is a function analytic in $S_\sigma$ and there satisfying (28), then the function defined by (29) is analytic in $C_R$ and satisfies (27) with $a_m = \frac{d^m}{d\tau^m} B(\tau)|_{\tau=0}$ uniformly in every $C_{R'}$ with $R' < R$.

If $h$ is analytic on $C_\infty = \bigcup_{R>0} C_R = \{ \zeta \in \mathbb{C} : \Re \zeta > 0 \}$ and (27) holds on that domain, the representation (29) of course also holds for $\zeta \in C_\infty$ and for $\zeta$ real and $\beta = \Re \zeta \in \mathbb{R}_+$ in particular. The aim now is to apply this result to the series given in (19). It is easy to show analyticity:

The probability that a random variable $Z_t$ distributed according $\Phi_t$ takes large values $\phi$ is suppressed much too strongly. To adjust the large $\phi$ behavior this first step has to be combined with resummation techniques, cf. Section 3.2.
Lemma 2.19. The function $\mathbb{R}_+ \ni \beta \to \Phi_t^\beta(\phi)$ (cf. Section 2.1) has analytic continuation $C_\infty \ni \zeta \to \Phi_t^\zeta(\phi)$, $\forall t \in \mathbb{R}_+, \phi \in \mathbb{R}^d$.

Proof. Note that for $\zeta \in C_\infty$, $f_\zeta = e^{-\zeta V} \in S(\mathbb{R}^d)$ and hence $\Phi_t^\zeta(\phi) = \nu_t \ast f_\zeta(\phi)$ is well defined. To show the complex differentiability in $\zeta$ we now proceed as in the proof of Lemma A.1.

In order to check (27), we need an estimate on the multiple $\zeta$ derivatives of $\Phi_t^\zeta(\phi)$ on $C_R$ or even $C_\infty$. As $\Phi_t^\zeta(\phi)$ is not known explicitly, such an estimate is difficult to obtain. But in the case that $\Psi$ is the symbol of a Lévy process, the probabilistic structure helps:

Lemma 2.20. If $\Psi$ is the symbol of a Lévy process and $V(\varphi) \geq 0 \ \forall \varphi \in \mathbb{R}^d$, then

$$\left| \frac{d^m}{d\zeta^m} \Phi_t^\zeta(\phi) \right| \leq \langle V^m \rangle_{\nu_t}, \ \zeta \in C_\infty, t \in \mathbb{R}_+, \phi \in \mathbb{R}^d. \quad (30)$$

Proof. We have $\left| \frac{d^m}{d\zeta^m} \Phi_t^\zeta(\phi) \right| = \left| \langle \nu_t, V^m \cdot e^{-\zeta V} \rangle \right| \leq \langle V^m \rangle_{\nu_t}$ as $\nu_t$ is a probability measure and $|e^{-\zeta V}| \leq 1$.

Lemma 2.20 allows us in the given situation to check (27) by Taylor’s estimate for the remainder and the calculation of the $N$-th coefficient in the perturbation series. Being a sum over Feynman graphs, the coefficient can be dominated by the number of such Feynman graphs times the maximal value of $V[G](t, \phi)$ over all these graphs. The latter is easy to control. We therefore start with an estimate on the number of Feynman graphs. Let us first assume that the shift $\phi$ is zero. We furthermore assume that $V$ is homogeneous of degree $\bar{p}$. Then only graphs without outer empty vertices and with inner full vertices with $\bar{p}$ legs contribute. As such graphs are in one to one correspondence with the number of partitions of $\bar{p}N$ objects, we have to control the number of such partitions.

The number of partitions of $m$ objects is known as $m$-th Bell number and we denote it by $b_m$. It has the asymptotic behavior, see e.g. [10],

$$b_m \sim m^{-\frac{1}{2}} \lambda(m)^{m+\frac{1}{2}} e^{\lambda(m)-m-1}, \quad (31)$$

where $\lambda(m)$ is implicitly defined by $\lambda(m) \ln \lambda(m) = m$. Using Stirling’s formula $N! \sim \sqrt{2\pi e(N+\frac{1}{2})} \ln N - N$ we get the following asymptotic formula for $b_{\bar{p}N}/N!$:

$$\frac{1}{N!} b_{\bar{p}N} \sim \frac{1}{\sqrt{2\pi}} \bar{p}^{-\frac{1}{2}} e^{(\bar{p}-1)(N \ln \bar{p} - N - N) - \frac{1}{2} \ln N - \frac{1}{2} \ln \bar{p} + \ln \ln \lambda(\bar{p}N)} e^{\frac{1}{2} \ln \lambda(\bar{p}N) + \lambda(\bar{p}N)}, \quad (32)$$
from which we obtain that
\[ \frac{1}{N!} p_N \leq A' K^N N! \] (33)
is fulfilled for \( p \leq 2 \) and \( K \) sufficiently large. In the following we restrict ourselves to \( p = 2 \).

Let us now consider a Feynman graph \( G \in \bar{F}(N) \) without empty outer vertices. If the modulus of the coefficients \( C^{(n)}_{X_1 \ldots X_n} \) is uniformly bounded by \( \hat{A} \kappa^n \) for \( \hat{A}, \kappa \) sufficiently large, we obtain
\[ |V(G)(t,0)| \leq \max\{\hat{A},1\}^N \kappa^N d^N \kappa 2^N \max\{1,t\}^{2N}, \quad \forall G \in \bar{F}(N), \] (34)
with \( \kappa = \max_{X_1, X_2 \in \{1, \ldots, d\}} |\lambda^{(2)}_{X_1 X_2}|. \) Thus, under the given conditions, the estimate \( \hat{A} \) is fulfilled for \( \phi = 0 \) with \( h(\zeta) = \Phi^\zeta(0), a_m = \frac{(-1)^m}{m!} \sum_{G \in F(m)} V[G](t,0), A = A' \) and \( \sigma = d \kappa \kappa 2^N \max\{1,t\}^2 \max\{\hat{A},1\}. \) It turns out that the case \( \phi \neq 0 \) can be easily traced back to the case \( \phi = 0: \)

**Theorem 2.21.** Let \( \Psi \) be a Lévy symbol as in \( \hat{A} \) such that \( \text{supp}(r) \) is compact and let \( f = e^{-V} \) with \( V(\varphi) = \langle \lambda^{(2)}, \varphi \otimes \varphi \rangle_2. \) Then \( \Phi^\zeta(\phi) \) as a function of \( \zeta \) fulfills the conditions of Theorem \( \hat{A} \).

In particular, the series
\[ B(\tau, t, \phi) = \sum_{m=0}^{\infty} \frac{(-\tau)^m}{(m!)^2} \sum_{G \in F(m)} V[G](t, \phi) \] (35)
converges on a ball of radius \( 1/\sigma \) centered at 0 and has analytic continuation to the strip \( S_\sigma \) (for some \( \sigma = \sigma(t, \phi) < \infty \).) Denoting this continuation by the same symbol, we get the solution \( \Phi^\zeta_t \) of the Cauchy problem \( \hat{A} \) through the formula
\[ \Phi^\zeta_t(\phi) = \int_0^\infty e^{-\tau} B(\tau, t, \phi) d\tau. \] (36)

**Proof.** If \( \text{supp}(r) \subseteq \{ x \in \mathbb{R}^d : |x| \leq \kappa \}, \) then the uniform bound \( |C^{(n)}_{X_1 \ldots X_n}| \leq \hat{A} \kappa^n \) holds for \( \hat{A} \) sufficiently large, since \( C^{(n)}_{X_1 \ldots X_n} \) for \( n > 2 \) up to a factor \( z \) is a \( n \)-th moment of \( r. \) From the considerations above the assertion now follows for \( \phi = 0. \)

If \( \phi \neq 0, \) then modify for a fixed time \( t = t_0 \) the Lévy symbol \( \Psi \) by a deterministic term \( \Gamma_{\phi,t_0}(-i\xi) = \Psi(-i\xi) + i\phi \cdot \xi/t_0 \) generating the shift \( \phi \) at the time \( t_0. \) Let \( \Theta_{t_0,\phi} \) be the solution of the Cauchy problem with \( \Psi \) replaced by \( \Gamma_{t_0,\phi}, \) then obviously \( \Phi^\zeta_{t_0}(\phi) = \Theta^\zeta_{t_0,\phi}(0). \) Since the coefficients of \( \Psi \) fulfill the above uniform bound, the same holds for the coefficients of \( \Gamma_{t_0,\phi} \) (with \( \hat{A} \) changed depending on \( t_0 \) and \( \phi \)). The assertion holds for \( \Theta^\zeta_{t_0,\phi}(0), \) hence it also holds for \( \Phi^\zeta_{t_0}(\phi). \)

In fact, both expansions are equal. The only difference is that in the graphic expansion of \( \Theta_{t_0,\phi}(0) \) there are no empty outer vertices. This is compensated by evaluation of the empty inner vertices with one leg by \( (C^{(1)}_X - \phi X). \) Multiplying out the parenthesis just generates the contributions of a non-modified inner empty vertex with one leg and the outer empty vertex.

The restriction to \( p = 2 \) still covers a set of initial conditions \( f \) that determine the convolution kernel \( \nu_t \) completely, cf. Section \( \hat{A}. \) It is also sufficient for the most important application that we have in mind, see the following Section \( \hat{A}. \)

The second technical restriction on the support of the jump distribution \( r \) is more severe. It can be overcome by considering generalized notions of Borel summability, cf. \( \hat{A}, \) that for the
case of Lévy processes can be adapted to a growth like \((\sim N!)^\gamma\) of the moments of the jump distribution \(r\). In applications this matters if one wants to choose \((36)\) or generalizations thereof as the point of departure of the numerical evaluation, cf. \([9\, \text{Chapters 16.3–16.6}]\) and one has to take care about the speed of convergence of the (generalized) Borel transform. This program, as all high precision numerics based on our proposal, is beyond the scope of the present article.

Here we take the somewhat more theoretical point of view that only a finite number of terms of the perturbation series can be calculated. Hence, only a finite number of moments of \(r\) enters. From the values of finitely many moments in general it is not possible to determine, whether \(r\) is compactly supported or not. In statistical applications with a numerics that is not based on the Borel transform (see again Section 8), the compact support of \(r\) does not pose any problems, as this property cannot be proven/disproven on the basis of finite samples\(^9\).

It would be desirable to have a similar theorem for \(\log \Phi_t(\phi)\). The analyticity of \(\log \Phi^C_t(\phi)\) on some circle \(C_R\) is immediate. The bounds \((21)\) for the perturbation series at zero is a corollary of the above considerations, since the connectedness condition reduces the total number of Feynman graphs. In particular, the Borel sum \(B_c(\tau, t, \phi)\) of \(\log \Phi^C_t(\phi)\) has a non-vanishing radius of convergence. But for the log-solution the analogue of Lemma 2.20 is missing and therefore the bounds on the series coefficients at \(\zeta = 0\) do not produce a uniform estimate \((24)\) on some \(C_R, R > 0\). So the existence of an analytic continuation of \(B_c(\tau, t, \phi)\) is not proven. Of course, one can repeatedly use the chain rule to obtain a bound on \(\log \Phi^C_t(\phi)\), but that generates an additional factor \(\sim b_N\). So more sophisticated estimates or more powerful versions of Theorem 2.18 are needed \([12]\).

Here we do not enter into this subject. For practical use, again, the existence of the analytic continuation of \(B_c(\tau, t, \phi)\) is of little importance. Numerical procedures to evaluate the integral transform \((29)\) are based on the convergence of the Borel sum which permits resummation as e.g. the Borel-Padé or the conformal mapping method \([24\, \text{Chapter 16}]\). Both methods, relying on the first \(N\) terms of the perturbation expansion only, automatically produce analytic continuations of an approximated \(B_c(\tau, t, \phi)\) on some strip region that are sufficiently bounded in order to render \((29)\) absolutely convergent\(^10\).

3 Application to Lévy distributions

3.1 Large diffusion expansion of Lévy densities

It is more interesting to get explicit (approximate) formulae for the distribution of a Lévy process starting at a fixed point, e.g. 0, than to determine the density of the distribution for an initial condition \(f = e^{-V}\) with \(V\) as above. Though the densities of Lévy distributions can be expressed in terms of more or less elementary functions for many interesting examples, see e.g. \([2\, \text{Chapter 1}]\), or can at least be expanded in a series in other cases \([16]\), a generic formula is missing.

In this section we present asymptotic expansions of (non-normalized) Lévy densities and their logarithms, respectively, for the case where the generator of the process has a large diffusion constant \(D > 0\). The formulae are generic in the sense that no detailed assumptions on the distribution of jumps, \(r\), are required. I.e. we consider the generator \((6)\) with \(D_{X_1,X_2} = D\delta_{X_1,X_2}\) with \(\delta_{X_1,X_2}\) the Kronecker symbol and we assume that \(D\) is large and hence \(1/D\) is a small

\(^9\)Note that equation \((36)\) does not depend on \(\kappa\!^!\)

\(^{10}\)Strictly speaking, in the Padé case this is only true if the Padé approximant has no poles on \(\mathbb{R}_+\).
parameter in which we would like to expand. As the initial value problem for the density of the jump-diffusion process starting at zero, we get for $t, \phi \in (0, \infty) \times \mathbb{R}^d$

$$\begin{align*}
\frac{\partial \Phi_t}{\partial t}(\phi) &= D\Delta \Phi_t(\phi) - \sum_{X=1}^d a_X \frac{\partial}{\partial \phi X} \Phi_t(\phi) + \sum_{R \in \mathbb{Q}_1(0)} [\Phi_t(\phi + \varphi) - \Phi_t(\phi)] \, dr(\varphi) \\
\Phi_0(\phi) &= \delta_0(\phi)
\end{align*}$$

(37)

with $\delta_0$ the Dirac measure with weight one in zero and $\Delta$ the Laplacian on $\mathbb{R}^d$. The solution, $\Phi_t^{\text{diff}}$, of the diffusive part alone is of course

$$\Phi_t^{\text{diff}}(\phi) = (4\pi Dt)^{-\frac{d}{2}} e^{-\frac{||\phi||^2}{4Dt}}, \quad \beta = 1/(4Dt), \quad t > 0.$$  

(38)

Let $\nu_t = \nu_t^{\text{jump}}$ be the inverse Fourier transform of $e^{t\Psi^{\text{jump}}(-i\xi)}$ where $\Psi^{\text{jump}}$ is obtained from $\Psi$ letting $D = 0$. Obviously we obtain the solution of (37) by $\Phi_t(\phi) = \nu_t * \Phi_t^{\text{diff}}$. With $V(\phi) = ||\phi||^2$, $f_\beta = e^{-\beta V}$ one thus gets

$$\left(\frac{\beta}{\pi}\right)^{-d/2} \Phi_t(\phi) = \nu_t * f_\beta(\phi).$$

We note that the expansion of the right hand side in powers of $1/D$ is equivalent to the expansion in powers of $\beta$ as the m-th coefficient only differs by a factor $1/(4t)^m$. Furthermore, from now on we omit the multiplication with factors $\beta$ in the evaluation rules for the Feynman graphs in order to obtain a more explicit $\beta$ dependence. Combining this with Corollary 2.12 and Theorem 2.17 we obtain the $1/D$ (large diffusion) expansion for the non-normalized (log-) density function:

**Theorem 3.1.** In the sense of asymptotic series in $\beta$, we obtain the following large diffusion expansion of the non-normalized solution $\left(\frac{\beta}{\pi}\right)^{-d/2} \Phi_t$, where $\Phi_t$ solves (30),

$$\Phi_t(\phi) = \sum_{m=0}^{\infty} \frac{(-\beta)^m}{m!} \sum_{G \in \mathcal{Q}(m)} V[G](t, \phi).$$

(39)

Furthermore, again in the sense of asymptotic series, we have

$$\log \left[ \left(\frac{\beta}{\pi}\right)^{-d/2} \Phi_t(\phi) \right] = \sum_{m=1}^{\infty} \frac{(-\beta)^m}{m!} \sum_{G \in \mathcal{Q}_{c}(m)} V[G](t, \phi)$$

(40)

with $\mathcal{Q}_{c}(m)$ the set of connected graphs in $\mathcal{Q}(m)$.

Combination of (39) with Theorem 2.21 yields the following convergent representation as a Borel transform:

**Corollary 3.2.** Suppose that the support of the jump distribution $r$ is compact. Let $B(\tau, t, \phi)$ be the analytic continuation of $\sum_{m=0}^{\infty} \frac{(-\tau)^m}{(m!)^2} \sum_{G \in \mathcal{Q}(m)} V[G](t, \phi)$ to some strip of the form $S_\sigma$. Then,

$$\Phi_t(\phi) = \pi^{-\frac{d}{2}} \beta^{rac{d}{2}-1} \int_0^{\infty} e^{-\tau/\beta} B(\tau, t, \phi) \, d\tau.$$  

(41)

A comparison between Corollary 2.12 and Corollary 3.2— and other possible expansions known from statistical mechanics seems to be of interest. In particular, readers familiar with the theory of classical gases might find the convergent small $z$ ("low activity") expansion more suited than the small $\beta$
(“high temperature”) expansion. For a small $z$ expansion, a graphical representation by Meyer graphs exists, too. But there is a decisive disadvantage from the point of view of applications. The small $\beta$ expansion produces expressions that directly depend on the quantities of statistical importance – $z$ and the moments $r_n$ of the jump distribution $r$. Hence these quantities can be obtained from an empirical distribution by fits. This can be done to some extent non-parametrically, i.e. without any a priory assumption on the jump distribution, as any sequence of moments of the jump measure depending on arbitrary parameter sets can be inserted into our formula. In the Meyer series however, the jump distribution $r$ enters via nontrivial integrals over products of Meyer functions $f(s, s', \phi, \phi') = e^{-ss'\mid\phi-\phi'\mid^2} - 1$ which, except for a few cases with $r$ particularly simple, can only be done numerically\(^{11}\).

Similar difficulties arise, if one wants to solve the inverse Fourier integral of $e^{t\Psi(-i\xi)}$. Note that this integral is of oscillatory type – so even its numerics is not really trivial.

### 3.2 Second order perturbation theory

In this section we give an illustration, how to actually calculate the $1/D$-expansion of a Lévy density and how to improve the result using resummation techniques. We take \(^{40}\) as starting point. Calculations are carried through up to second order in $\beta$, only, and we use the simplest possible resummation algorithm. Thus the idea of this section is by no means to provide high accuracy numerics, but to illustrate techniques up to second order of perturbation theory, that in higher orders will provide reasonable numerics. On the other hand, this does not mean that the second order calculation is not interesting at all from a numerical point of view. A more detailed discussion can be found below. Calculations to higher orders and more sophisticated resummation, see e.g. \(^{7, 8, ?}\) are left for the future.

In this subsection, we assume without loss of generality that $\Psi^{\text{jump}}(-i\xi)$ has a vanishing first coefficient $C^{(1)}$. If this is not the case, this can be achieved by a shift $\phi \rightarrow \phi - a$, cf. the proof of Theorem \(^{22, 24}\).

Under these conditions, the second order perturbation expansion in $\beta$ for the left hand side of \(^{39}\) is given in \(^{41}\). Extracting the connected graphs, we obtain

$$
\log \left[ \left( \frac{\beta}{\pi} \right)^{-d/2} \Phi_t(\phi) \right] = -\beta \left[ \begin{array}{c} \circ \rightarrow \circ \\
+ \end{array} \right] + \frac{\beta^2}{2} \left[ \begin{array}{c} \circ \rightarrow \circ + 2 \circ \rightarrow \circ + 4 \circ \rightarrow \circ + 4 \circ \rightarrow \circ \\
+ \end{array} \right] + \mathcal{O}(\beta^3). \tag{42}
$$

As already in equation \(^{41}\), we omitted the evaluation symbol $\mathcal{V}$ and we multiplied with the multiplicities of the topological graphs.

Equation \(^{42}\) should be o.k. for $z$ and $1/D$ sufficiently small and\(^{12}\) $t \approx 1$, but the positive sign in front of the $\beta^2$-term implies that the exponential of the second order approximation becomes non-normalizable (not $L^1$ integrable) if $\beta$ (or $z$) passes a certain threshold. To extract reliable data also for higher values of $\beta$, resummation is required\(^{13}\). The idea of Padé approximants is to replace the Taylor polynomial by a rational function that has the same Taylor series as the polynomial. For definitions and details we refer to Appendix \(^{A.3}\).

\(^{11}\)Note that the small time expansion in \(^{3, 11}\) is equivalent to the small $z$ expansion.

\(^{12}\)By a scaling of $D$ and $z$ one can always obtain $t = 1$.

\(^{13}\)In fact, Padé resummation of expansions in physics leads to results with impressive numerical precision, cf. e.g. \(^{14}\) p. 506. Often, validity of the resummed approximation is only stopped by phase transitions – which are of course absent on the discrete, finite space $\{1, \ldots, d\}$, as considered here.
Feynman graph representation of semigroups

rel. frequency and densities

QQ-plot predicted vs simulated

Sample quantiles of pseudo random numbers

Figure 3: a) Relative frequency of pseudo random numbers (sample size \( n = 10^6 \)), density of 2nd order Padé approximant \([44]\) (solid line) and a best Gaussian fit (dashed line); b) the qq-plot of sample quantiles (1% to 99%) of pseudo random numbers \( (n = 10^5) \) vs the predicted quantiles of \([44]\) (circle) and a best Gaussian fit (triangle). The black line is the graph of \( f(x) = x \) (optimal line). Here \( \beta = 0.2, z = 1, r = \delta_0 \).

Following the standard calculation in A.3 we obtain as the second order Padé approximant

\[
\log \left( \frac{\beta}{\pi} \right)^{-d/2} \Phi_t(\phi) = \frac{-\beta \left[ + + + + \right]}{1 + \frac{\beta}{2} \left[ + + + + \right]} + \mathcal{O}(\beta^3), \tag{43}
\]

which, if exponentiated, gives a normalizable approximation to \( \Phi_t \) for \( t, z \) and \( \beta \) arbitrary. Evaluation for \( d = 1 \) with \( C^{(n)} = zr_n, r_n = \int_{\mathbb{R}} s^n \, dr(s) \) yields an expression in terms of the time \( t \), the activity (mean frequency of jumps in unit time) \( z \), and the 2nd, 3rd and 4th moment of the jump distribution \( r \):

\[
\log \left( \frac{\beta}{\pi} \right)^{-d/2} \Phi_t(\phi) = \frac{-\beta(tzr_2 + \phi^2)}{1 + \frac{\beta}{2} \left[ + t z r_4 + 2r_2^2 + 4t z r_2^2 + 4t z r_3 \phi^2 \right] \frac{1}{t z r_2 + \phi^2}} + \mathcal{O}(\beta^3). \tag{44}
\]
We have tested formula (44) by comparison with large samples of pseudo random numbers that simulate a compound Poisson process with additional diffusion, $Z$ i.e.

$$Z = X(a, \sigma^2) + s_1 Y_1(z_1) - s_2 Y_2(z_2),$$  \hspace{1cm} (45)$$

where $X(a, \sigma^2)$ has Gaussian distribution with mean $a$ and standard deviation $\sigma = 1/\sqrt{2D}$. $Y_j(z_j)$ is Poisson distributed with parameter $z_j$, $j = 1, 2$. $s_1, s_2$ give the length of positive/negative jumps, respectively. $X, Y_1, Y_2$ are all independent. This choice is motivated by the fact that pseudo random samples for $Z$ can be obtained using standard software routines. The parameters of the model are $z = z_1 + z_2$ and $r = (z_1 \delta_{s_1} + z_2 \delta_{s_2})/(z_1 + z_2)$, and for the shift $\phi \to \phi - a$ one obtains $a = z_1 s_1 - z_2 s_2$. The normalization constant of the exponentiated right hand side of (44) has been calculated numerically. The range of parameters has been $0.1 \leq \beta \leq 10$ and $0.5 \leq z \leq 4$, $1 \leq s_j \leq 10$, $j = 1, 2$.

If $\Psi_{\text{jump}}(-i\xi)$ is symmetric under $\xi \to -\xi$ (here: $s_1 = s_2$ and $z_1 = z_2$), the results of the 2nd order calculation is very good, as long as jumps are not too large compared with the diffusive part, which gives a nice illustration for the power of resummation.

The parameters of the model are $z$ close to one and to the Gaussian quantiles else, we found reasonable quantile functions for a large range of the Gaussian model that gives a larger weight to the 2nd order quantiles if the probability is small or close to one and to the Gaussian quantiles else, we found reasonable quantile functions for a large range of parameters.

### 3.3 Second order maximum likelihood estimation

As already mentioned at the end of Subsection 3.1, the Feynman graph expansion for Lévy densities can be directly expressed in terms of quantities of statistical interest, as e.g. the activity $z$ and the moments $r_n$ of the jump distribution. It it therefore natural to ask for a method, how to determine these quantities by comparison with statistical data.

The formula for the log of the density in (40) makes it natural to apply the maximum likelihood method, where the log-density plays a crucial rôle, too. In fact, in this respect there seems to be an interesting parallel between the calculation of thermodynamic potentials in statistical mechanics and the maximum likelihood method.

When applying the maximum likelihood method, see e.g. [1] for a historical review, to (40), it is convenient to treat $\beta$ as a known parameter since the term $-\frac{d}{\pi} \log(\beta/\pi)$ can then be ignored in the maximization problem

$$\text{"find } \hat{\theta} \text{ such that } \sum_{i=1}^{n} \log \Phi_i(\phi_i, \hat{\theta}) \geq \sum_{i=1}^{n} \log \Phi_i(\phi_i, \theta) \forall \theta \"}$$  \hspace{1cm} (46)$$

---

14 Actually, in this example $\sigma_{\text{Gauss}} = \sqrt{2D} = 1.58 < 6$ =jump length, hence the jump part for $z = 1$ not small compared with the diffusive part, which gives a nice illustration for the power of resummation.

15 Using an ad hoc interpolation between the quantiles of 2nd order perturbation theory and the quantiles of the Gaussian model that gives a larger weight to the 2nd order quantiles if the probability is small or close to one and to the Gaussian quantiles else, we found reasonable quantile functions for a large range of parameters.
where $\phi_1, \ldots, \phi_n \in \mathbb{R}^d$ are sample values and $\theta$ either stands for the coefficients\(^{16}\) of $\Psi^{\text{jump}}$, $C^{(1)}, C^{(2)}, \ldots$, or for some parametrization of these coefficients. $\hat{\theta}$ stands for the best estimation of the parameters $\theta$ out of the sample $\{\phi_i\}_{i=1}^n$. If the diffusion part is unknown, one can still try a rough guess for $\beta$ and do the fine tuning by including a small diffusive (quadratic) term to $\Psi^{\text{jump}}$, which only changes $C^{(2)}$ by a constant which then can be treated as an additional parameter.

Parametrizations $\theta$ will in general include the above mentioned correction for the diffusion constant $D$, the activity $z$ and some parametrization $\alpha$ of the jump distribution $r$ such that the moments of $r$ can be expressed in terms of the parameters $\alpha$. Of course, some of these parameters could again be known, s.t. the maximization problem simplifies.

The right hand side of (40) is then truncated after $N$ terms and resummed. The order $N$ of the perturbative maximum likelihood parameter estimation has to be chosen as a function on the precision demanded and the number of parameters that should be estimated. At the present stage, this problem requires further study.

As an example, we apply a maximum likelihood activity estimation based on the 2nd order formula (44). Here we assume that $D$ and the law of the jumps are known. Figure 4 shows a typical result. Again, we only want to show that the perturbative maximum likelihood method works in principle and we do not have any ambition in precise numerics. In fact, as $z$ enters into the variance via $\sigma^2 = \sigma^2_{\text{gauss}} + zr^2$, one can solve for $z$ replacing $\sigma^2$ by the sample variance which gives a more precise estimate. But note that the variance method cannot be extended to estimate parameters like the skewness $C^{(3)}$ and kurtosis $C^{(4)}/(2D + C^{(2)})^2$ of a Lévy process – whereas it is possible to adapt the perturbative maximum likelihood method. Such estimates have been done on the basis of assumptions on the distribution of the Lévy process, as e.g. (truncated) stable laws or Meixner distributions. The advantage of the Feynman graph method however is that there is no unmotivated assumption on the class of distributions\(^{17}\).

### A Appendices

#### A.1 Proofs of Section 2.1

**Proof of Lemma 2.1**

\[ e^{\mu \Psi(-i\xi)} \rightarrow 1 \text{ as } t \downarrow 0 \text{ in the } \mathcal{S}'(\mathbb{R}^d) \text{ topology, hence } \nu_t = \mathcal{F}^{-1}(e^{\mu \Psi(-i\xi)}) \rightarrow \mathcal{F}^{-1}(1) = \delta_0 \text{ in } \mathcal{S}'(\mathbb{R}^d) \text{ for } t \downarrow 0. \]

Here $\delta_0$ is the Dirac measure of mass one in zero. Consequently $\Phi_t(\phi) = \nu_t * f(\phi) \rightarrow \delta_0 * f(\phi) = f(\phi)$ as $t \downarrow 0$.

It remains to show that $\Phi_t$ is differentiable in $t$ and fulfills equation (11). Firstly, the differential quotient

\[ \frac{e^{\mu \Psi(-i\xi)} - e^{\mu \Psi(-i\xi')}}{t - t'} \]

converges to $\Psi(-i\xi)e^{\mu \Psi(-i\xi)}$ pointwisely for all $\xi \in \mathbb{R}^d$. By

---

\(^{16}\) We note that the positivity of $r$ leads to some constraints for these coefficients, that have to be taken into account in the maximization problem.

\(^{17}\) In future applications, such assumptions have to be introduced for the jump distribution in order to restrict the number of parameters in higher order calculations. But the parametrization can be changed at any time, as it plays no crucial rôle in the formalism. In particular, different assumptions on the jump distributions can also be compared with each other by the maximum likelihood principle.
Figure 4: a) The 2nd order log. density as function of $z$ for $\phi = 1, 2, 3$, $r = \delta_1$ and $\beta = 0.1$, hence $\sigma_{\text{Gauss}} = \sqrt{2D} = 2.236$. For $\phi = 1$ (solid line), the deviation from zero is smaller than expected from the Gaussian fluctuations alone, the maximum is at $z = 0$, for $\phi = 2$ (dashed line) a maximum is being formed for $z > 0$, which is shifted further to the right for $\phi = 3$ (dotted line). b) Maximum likelihood estimation for the activity based on (44) for $\beta = 0.1$, and $r = \delta_1$. The size of the sample of (pseudo) random numbers is $n = 10^4$. 
Lemma A.1. For continuity of $F$ in $(t, t')$, we get

$$\left| \frac{e^{i\Psi(-i\xi)} - e^{i\Psi(-i\xi)}}{t - t'} \right| \leq \left| \Psi(-i\xi)e^{i\Psi(-i\xi)} \right| (1 + \|\xi\|)|t - t'|c$$

(47)

and by P2) this implies that the differential quotient is uniformly polynomially bounded (for, say, $|t - t'| < 1$). It follows by dominated convergence that the differential quotient converges in $S'(\mathbb{R}^d)$ as $t' \to t$ which implies that $F(\nu_t)$ is differentiable in $t$ in the $S'(\mathbb{R}^d)$ topology. By the continuity of $F$, the same applies to $\nu_t$ and hence $\Phi_t(\phi)$ is differentiable.

It is immediate that $F\left( \frac{d}{dt} \nu_t \right)(\xi) = \Psi(-i\xi)e^{i\Psi(-i\xi)}$, from which we get that $F\left( \frac{d}{dt} \Phi_t \right)(\xi) = \Psi(-i\xi)F(\Phi_t)(\xi)$. From the fact that the time derivative can be taken out of the Fourier transform, uniqueness of the Fourier transformed solution follows. The argument is concluded by application of the inverse Fourier transform.

**Proof of Proposition 2.2**

In the following let $\|h\|_{\alpha, N}$ be a Schwartz norm, $N \in \mathbb{N}$, $\alpha \in \mathbb{N}_0^d$, i.e.

$$\|h\|_{\alpha, N} = \sup_{\phi \in \mathbb{R}^d} (1 + \|\phi\|)^N |D_\alpha h(\phi)|$$

(48)

**Lemma A.1.** For $\beta > 0$, $\Phi_t^\beta(\phi)$ is infinitely differentiable in $\beta$ and $\frac{d^m}{d\beta^m}\Phi_t^\beta(\phi) = (-1)^m \nu_t \ast (V^m f_\beta)(\phi)$.

**Proof.** Since $\Phi_t^\beta(\phi) = \nu_t \ast f_\beta(\phi)$ and $\nu_t \in S'(\mathbb{R}^d)$, it suffices to show that the mapping $(0, \infty) \ni \beta \to f_\beta = e^{-\beta V} \in S(\mathbb{R}^d)$ is $m$ times differentiable on $(0, \infty)$, $m \in \mathbb{N}$ arbitrary, and $\frac{d^m}{d\beta^m} f_\beta = (-1)^m V^m f_\beta$. In fact, this statement is trivial for $m = 0$. Suppose that we have proven the statement up to $m - 1$. We have to show that the differential quotient $(-1)^{m-1}V^{m-1}(\frac{f_\beta - f_{\beta'}}{\beta - \beta'})$ converges to $(-1)^m V^m f_\beta$ as $\beta' \to \beta$. As the common pre-factor $(-1)^{m-1}V^{m-1}$ is a polynomial, it suffices to treat the case $m = 1$. For $\phi \in \mathbb{R}^d$ we have

$$- V(\phi) f_\beta(\phi) - f_\beta(\phi) - f_{\beta'}(\phi) = - V(\phi) \frac{f_{\beta'}(\phi) - f_\beta(\phi)}{\beta - \beta'}$$

(49)

Hence

$$\left\| - V f_\beta(\phi) - \frac{f_\beta - f_{\beta'}}{\beta - \beta'} \right\|_{\alpha, N} \leq K \sum_{\alpha' \in \mathbb{N}_0^d: |\alpha'| < |\alpha|} \sup_{\phi, \tau \in \mathbb{R}^d} (1 + \|\phi\|)^{N_{\alpha'}} \sup_{\tau \in (\beta, \beta')} |D_{\alpha'} f_\beta(\phi) - D_{\alpha'} f_{\tau}(\phi)|.$$  

(50)

for $K$ and $N_{\alpha'}$ sufficiently large. The derivative $D_{\alpha'} f_\beta(\phi)$ is of the form $P_{\alpha'}(\phi, \beta) f_\beta(\phi)$ where $P_{\alpha'}(\phi, \beta)$ is a polynomial in $\beta$ and $\phi$. Hence, for $|\beta - \beta'| < 1$, $\beta, \beta' > 0$, $\epsilon > 0$,

$$\sup_{\tau \in (\beta, \beta')} |D_{\alpha'} f_\beta(\phi) - D_{\alpha'} f_{\tau}(\phi)| = \sup_{\tau \in (\beta, \beta')} |P_{\alpha'}(\phi, \beta) f_\beta(\phi) - P_{\alpha'}(\phi, \tau) f_{\tau}(\phi)|$$

$$\leq \sup_{\tau \in (\beta, \beta')} |P_{\alpha'}(\phi, \beta) - P_{\alpha'}(\phi, \tau)| f_\beta(\phi) + \sup_{\tau \in (\beta, \beta')} |P(\phi, \tau)(f_\beta(\phi) - f_{\tau}(\phi))|$$

$$\leq K_{\alpha'} |\beta - \beta'| (1 + \|\phi\|)^{Q_{\alpha'}} \left( f_\beta(\phi) + |V(\phi)|e^{(\beta+1)L} e^{-\epsilon(\lambda(\beta), \phi, \rho)} \phi \right)$$

(51)
with \( L = \min_{\phi \in \mathbb{R}^d} V(\phi) - \epsilon (\lambda^{(p)}, \phi^{\otimes p})_p, K_{\alpha'}, Q_{\alpha'} \) sufficiently large. Multiplication by \((1 + \|\phi\|)^{N_{\alpha'}}\) and taking the supremum over \( \phi \in \mathbb{R}^d \) implies that the right hand side of \( 50 \) vanishes as \( \beta' \to \beta \). This concludes the proof. ■

It remains to show that the derivatives of \( \Phi^\beta_\epsilon(\phi) \) extend continuously to \( \beta = 0 \). We start with two technical lemmas:

Let \( h_r \in \mathcal{S}(\mathbb{R}^d), r \in (0, \infty), \) be a family of functions such that \( h_r \to h \) for \( r \to 0 \) in \( \mathcal{S}(\mathbb{R}^d) \) where \( \int_{\mathbb{R}^d} h \, d\xi = 1 \). Let furthermore \( h^\epsilon(\xi) = e^{-\epsilon h(\xi/\epsilon)} \) and \( h^\epsilon(\xi) \) be defined likewise. Let \( r(\epsilon) \) be a function such that \( r(\epsilon) \to 0 \) as \( \epsilon \to 0 \).

**Lemma A.2.** Let \( F : \mathbb{R}^d \to \mathbb{C} \) be a function that is polynomially bounded and has polynomially bounded partial derivatives. Let \( h^\epsilon_r \) and \( r(\epsilon) \) as above. Then \( \lim_{\epsilon \to 0} \int_{\mathbb{R}^d} F(\xi) h^\epsilon_{r(\epsilon)}(\xi) \, d\xi = F(0) \).

**Proof.** We consider the following estimate

\[
\left| \int_{\mathbb{R}^d} F(\xi) h^\epsilon_{r(\epsilon)}(\xi) \, d\xi - F(0) \right|
= \left| \int_{\mathbb{R}^d} F(\epsilon \xi) h_{r(\epsilon)}(\xi) \, d\xi - F(0) \right|
\leq \left| \int_{\mathbb{R}^d} F(\epsilon \xi) (h_{r(\epsilon)}(\xi) - h(\xi)) \, d\xi \right| + \left| \int_{\mathbb{R}^d} (F(\epsilon \xi) - F(0)) h(\xi) \, d\xi \right|
\]

(52)

The first term on the right hand side can be estimated as follows

\[
\left| \int_{\mathbb{R}^d} F(\epsilon \xi) (h_{r(\epsilon)}(\xi) - h(\xi)) \, d\xi \right| \leq \int_{\mathbb{R}^d} \frac{|F(\epsilon \xi)|}{(1 + \|\xi\|)^N} \, d\xi \|h_{r(\epsilon)} - h\|_{0,N}.
\]

(53)

Recall that \( F \) is polynomially bounded. Thus, for \( N \) sufficiently large, the integral on the right hand side converges to \( |F(0)| \int_{\mathbb{R}^d} \frac{1}{(1 + \|\xi\|)^N} \, d\xi \) by dominated convergence. Hence the right hand side of (52) tends to zero as \( \epsilon \to 0 \).

The second term on the right hand side of (52), for \( 0 < \epsilon < 1 \), can be estimated by

\[
\left| \int_{\mathbb{R}^d} (F(\epsilon \xi) - F(0)) h(\xi) \, d\xi \right| = \left| \int_0^\epsilon \int_{\mathbb{R}^d} \langle \nabla F(s \xi), \xi \rangle h(\xi) \, d\xi \, ds \right|
\leq \epsilon K \int_{\mathbb{R}^d} (1 + \|\xi\|)^N |h(\xi)| \, d\xi
\]

(54)

for \( K \) and \( N \) sufficiently large by the polynomial boundedness of the partial derivatives of \( F \). Hence also the second term on the right hand side of (52) tends to zero as \( \epsilon \to 0 \). ■

**Lemma A.3.** Let \( V(\varphi) = \langle \lambda^{(p)}, \varphi^{\otimes p} \rangle_p + \sum_{p=0}^{\bar{p}-1} \langle \lambda^{(p)}, \varphi^{\otimes p} \rangle_p \) such that \( \lambda^{(p)} \to 0 \) as \( \epsilon \to 0 \), \( p = 0, \ldots, \bar{p} - 1 \) and \( f^\epsilon = e^{-V(\cdot)} \). Let \( f^0(\varphi) = e^{-\langle \lambda^{(p)}, \varphi^{\otimes p} \rangle_p} \). Then \( f^\epsilon \to f^0 \) in \( \mathcal{S}(\mathbb{R}^d) \) as \( \epsilon \to 0 \).

**Proof.** We first assume \( \alpha = 0 \), then \( |f^0 - f^\epsilon| \leq \int_0^1 |Q e^{-sQ} | \, ds |f^0| \leq |Q| |e^{Q} f^0| \) with \( Q \) a polynomial that has vanishing coefficients as \( \epsilon \to 0 \). Multiplying this by another polynomial and taking the supremum over \( \mathbb{R}^d \) thus gives an expression that vanishes in that limit. Hence \( \|f^\epsilon - f^0\|_{0,N} \to 0 \) for \( \epsilon \to 0 \).

Let next \( \alpha \in \mathbb{N}_0^d, \alpha \neq 0 \). We note that each contribution to \( D_\alpha (f^0 - f^\epsilon) \) is of the form \( f_0 e^{Q} P \) or of the form \( S(f^0 - f^\epsilon) \). \( Q \) is of degree at most \( \bar{p} - 1 \) (potentially \( Q = 0 \)). \( P, Q \) and \( S \) are
polynomials. The coefficients of \( P \) and \( Q \) are vanishing as \( \epsilon \to 0 \). Hence we can conclude as above \( \| f^\epsilon - f^0 \|_{\alpha,N} \to 0 \) as \( \epsilon \to 0 \).

We now conclude the proof of Proposition \ref{prop:22} By Plancherel’s formula, P 2–3) and Lemma \ref{lem:A.1} we have

\[
\frac{d^m}{d\beta^m} \Phi_t^\beta (\phi) = (-1)^m \nu_t \ast (V^m f_\beta)(\phi) = (-1)^m \int_{\mathbb{R}^d} \mathcal{F}^{-1}(f_{\beta,\phi})(\xi)V^{\phi_m}(i\nabla_\xi)e^{t\Phi(-i\xi)} \, d\xi
\]  

(55)

For a polynomial \( W(\phi) = \sum_{p=0}^p \langle w(p), \phi \otimes p \rangle_p \) we set \( W_\beta(\phi) = \sum_{p=0}^p \beta^{1-p} \langle w(p), \phi \otimes p \rangle_p \). Let \( f^0(\phi) = e^{-\langle \lambda(\phi), \phi \otimes p \rangle} \) and let \( f_0^\beta = e^{-\langle V_\phi, \beta \rangle} \). By Lemma \ref{lem:A.2} \( f_0^\beta \rightarrow f^0 \) in \( \mathcal{S}(\mathbb{R}^d) \) as \( \beta \to 0 \). Hence, \( \mathcal{F}^{-1}(f_0^\beta) \rightarrow \mathcal{F}^{-1}(f^0) \) in \( \mathcal{S}(\mathbb{R}^d) \) and \( \int_{\mathbb{R}^d} \mathcal{F}^{-1}(f^0) \, d\xi = f^0(0) = 1 \).

Furthermore, \( f_{\beta,\phi}(\phi) = e^{-\beta V_\phi(\phi)} = e^{-\langle V_\phi, \beta \rangle} = f_0^\beta(\beta^{1/\beta} \phi) \). Inserting this into the right hand side of (55) we obtain

\[
\frac{d^m}{d\beta^m} \Phi_t^\beta (\phi) = (-1)^m \int_{\mathbb{R}^d} \beta^{1-p} \mathcal{F}^{-1}(f_{\beta,\phi})(\xi/\beta^p) V^{\phi_m}(i\nabla_\xi)e^{t\Phi(-i\xi)} \, d\xi
\]  

(56)

Note that by P2–P3) \( V_\phi(i\nabla_\xi)e^{t\Phi(-i\xi)} \) as a function in \( \xi \) is polynomially bounded with polynomially bounded derivatives. We can thus apply Lemma \ref{lem:A.2} and we get

\[
\lim_{\beta \to 0} \frac{d^m}{d\beta^m} \Phi_t^\beta (\phi) = (-1)^m V_\phi^m(i\nabla_\xi)e^{t\Phi(-i\xi)} \bigg|_{\xi=0} = (-1)^m \langle V_\phi^m(\phi) \rangle_{\nu_t}.
\]  

(57)

This finishes the proof.

### A.2 Notions from graph theory

For the readers convenience we develop some graph-theoretic notions, some of them being non-standard, following essentially \[3\].

Given a set \( M \), let \( M_2^2 \) be the set of subsets of \( M \) containing two elements. A graph \( G \) over \( M \) then is a subset of \( M_2^2 \). The points of \( M \) are called vertices of \( G \) and the elements (i.e. non ordered pairs of \( M \)) \( \{ m_1, m_2 \} \in G \) are called the edges (or bonds). \( G(M) \) stands for the collection of all graphs over \( M \). We could also replace \( M_2^2 \) with \( M^2 = M \times M \), the set of ordered pairs. We then obtain a graph with directed edges.

Let \( N \subseteq M \). The following construction explains the notion of a graph with distinguishable legs (endpoints of edges) at the vertices in \( N \). Let \( M \) be a set and \( \hat{\pi} : \hat{M} \rightarrow M \) a surjective map and let \( \hat{\pi}^2 : G(\hat{M}) \rightarrow G(M) \) be the induced mapping on graph level defined by application of \( \hat{\pi} \) on the endpoints of edges. Let \( G \in G(M) \) be fixed and for \( m \in M \) let \( p(m,G) \) be the number of edges connected with \( m \). Let \( (M_G, \hat{\pi}_G) \) be as above such that for \( m \in M \setminus N \) \( \# \hat{\pi}^{-1}_G(\{ m \}) = 1 \) and \( \# \hat{\pi}^{-1}_G(\{ m \}) = p(m,G) \) for \( m \in N \). \( \# A \) denotes the cardinality of the set \( A \). A leg \( \hat{m} \) at a vertex \( m \) is a point in \( \hat{\pi}^{-1}_G(\{ m \}) \). The class of graphs with distinguishable legs at vertices in \( N \) associated with an ordinary graph \( G \in G(M) \) is then given by

\[
[G|N] = \{ \hat{G} \in (\hat{\pi}_G)^{-1}(G) : \forall m \in N, \hat{m} \in \hat{\pi}^{-1}(m), \# \{ e \in E : \hat{m} \in e \} = 1 \}.
\]

The graphs with distinguishable legs at vertices in \( N \) is then \( [G|N](M) = \cup_{G \in G(M)} [G|N] \).

Let \( L \subseteq M \) and \( s \) in the permutation group of \( L, \text{Perm}(L) \). Obviously, \( s \) extends to a permutation on \( M \) by putting \( s(m) = m \) for \( m \in M \setminus L \). Let \( s^2 : G(M) \rightarrow G(M) \) be the induced mapping on graph level. We can then consider the equivalence classes under exchange of vertices in \( L, [G : L](M) = G(M)/\text{Perm}(L)^2 \). Such an equivalence class is called a graph with
indistinguishable vertices in $L$. It is easy to check that both of the above constructions are compatible, i.e. a vertex can belong to a set $L$ of indistinguishable vertices with distinguishable or indistinguishable legs and a distinguishable vertex can also have distinguishable or indistinguishable legs.

A graph with several types of vertices has sets of vertices $M_1, \ldots, M_k$ and is a graph over $M = \bigcup_{l=1}^{k} M_k$.

Finally, a topological graph is a graph with indistinguishable vertices in each class of vertices and with indistinguishable legs. The multiplicity of a topological graph, given some other class of graphs (generalized Feynman graphs in the present case), is the number of such graphs that coincide with the topological graph up to the labeling of indistinguishable vertices and legs.

Let us also point out that graph-theoretic notions like the connectedness for a graph with distinguishable legs $\hat{G}$ are meant in the sense that the associated "projected" graph $G$ or even topological graph $G_{\text{top}}$ are connected.\(^{18}\)

### A.3 Padé resummation

The idea of Padé resummation is to replace the Taylor polynomial with a rational function that has the given Taylor approximation at the origin but often has a better large $\beta$ behavior. We essentially follow [9, Chapter 16.2] The $[M/N]$-Padé approximant is defined as

$$[M/N](\beta) = \frac{a_0 + a_1 \beta + \cdots + a_M \beta^M}{1 + b_1 \beta + \cdots + b_N \beta^N}. \quad (58)$$

The $M + N + 1$ coefficients $a_0, \ldots, a_M; b_1, \ldots, b_N$ are determined by the equation

$$[M/N](\beta) = \sum_{m=0}^{M+N} h_m \beta^m + O(\beta^{N+M+1}). \quad (59)$$

Multiplication with the denominator of $[M/N]$ and comparison of coefficients leads to the system of equations

$$\begin{align*}
  b_N h_{M-N+1} + b_{N-1} h_{M-N+2} + \cdots + b_0 h_{M+1} &= 0 \\
  \vdots & \quad \vdots & \quad \vdots \\
  b_N h_M + b_{N-1} h_{M+1} + \cdots + b_0 h_{M+N} &= 0
\end{align*} \quad \begin{cases} b_0 = 1 \end{cases} \quad (60)$$

and

$$\begin{align*}
  b_1 h_0 + b_0 h_1 &= a_0 \\
  \vdots & \quad \vdots \\
  b_M h_0 + b_{M-1} h_1 + \cdots + b_0 h_M &= a_M
\end{align*} \quad \begin{cases} b_0 = 1. \end{cases} \quad (61)$$

Solving for $N = M = 1$, as required in Subsection 3.2 one obtains

$$\begin{align*}
  b_1 &= -\frac{h_2}{h_3} \\
  a_0 &= \frac{h_0}{h_1} \\
  a_1 &= h_1 - \frac{h_0 h_2}{h_1}
\end{align*} \quad (62)$$

\(^{18}\) $\hat{G}$ as a graph in $G(M)$ might be disconnected even in $G$ is connected.
With $h_0, h_1, h_2$ as in (42), this leads to

\begin{align*}
 b_1 &= \frac{1}{2} + 2 + 4 + 6 + 4 \\
 a_0 &= 0 \\
 a_1 &= -\left[ 0 + 6 + 4 \right] 
\end{align*}

from which (63) follows.

**Acknowledgements:** During his stay in Bonn, B. Smii received financial support by the DAAD. When this article was prepared, H.G. was supported by the D.F.G. project ”Stochastic methods in quantum field theory”. H. T. received financial support through the SFB 611. The authors gratefully acknowledge this support. The authors also thank Sergio Albeverio for his interest and encouragement while this work was under way. Habib Ouerdiane contributed interesting discussions and comments on the topic during his stay in Bonn in Summer 05.

**References**


