Anomalous statistics in turbulence, financial markets and other complex systems

J. Peinke*, F. Böttcher, and St. Barth
ForWind – Center for Wind Energy Research, Institute of Physics, Carl von Ossietzky University of Oldenburg, 26111 Oldenburg, Germany

Received 3 February 2004, accepted 5 April 2004
Published online 17 June 2004

Key words Stochastic processes, random walks and Levy flights, econophysics, isotropic turbulence.
PACS 02.50.-r, 05.40.Fb, 89.65.Gh, 47.27.Gs

The meaning of non-Gaussian statistics of disordered systems is discussed. The risks of heavy tailed probability distributions of wind gusts are presented together with the discussion of an eventual power law behavior of such heavy tailed distributions. We summarize the argumentation connecting such probability distributions with Levy processes and the meaning of non-existing moments. Based on turbulence and financial market data, stochastic cascade processes are presented, which lead to anomalous statistics with heavy tails, too. Finally we show how for disordered complex systems with an underlying hierarchical structure, like the cascade structure in turbulence, a stochastic equation can be estimated directly from the data. The numerical solution of this reconstructed stochastic process gives back the anomalous statistics.

1 Introduction

Scientists are driven by the desire to understand nature. Features of nature may be divided into two fundamentally different aspects, namely into that of order and that of disorder. In this contribution, we want to focus on the latter problem, namely the characterization of disorder in nature. A big part of our understanding of disorder is based on the research in former centuries. Somehow, the roots may be seen in the 18th century, when scientists were attracted by gambling and thus developed the fundamentals of statistics [1]. For the disordered systems which are governed by Gaussian or normal statistics there is a quite well elaborated understanding. A prominent example of this category is the thermodynamics of systems in equilibrium. Another prominent application of the Gaussian statistics is given in the presentation of experimental results by its mean value and error bar. It is the implicitly supposed Gaussian statistics of the measured variable $x$ with its error, which makes this characterization of an empirical finding clear. Thus, we expect for a repetition of this experiment to find a result within $\langle x \rangle \pm \sigma$ with a probability of 68%, or $\langle x \rangle \pm 2\sigma$ with 95%. Here we should remind the reader that a Gaussian statistics $p(x)$ is uniquely defined by its mean value $\langle x \rangle$ and its variance $\sigma^2 = \langle (x - \langle x \rangle)^2 \rangle$. In this contribution, we discuss situations with pronounced deviations from Gaussian statistics.

In the recent decades, systems driven into non-equilibrium states and complex systems have become of special interest for many physicists. On one hand, nonlinear deterministic dynamics have attracted much interest. New insight into deterministic chaos and spontaneous structure formation has been achieved. On the other hand, it has been realized that there is a large class of disordered systems which are governed

* Corresponding author E-mail: peinke@uni-oldenburg.de
by anomalous statistics, i.e. statistics which deviate from Gaussian statistics. In particular, the research is focused on statistics with heavy tails. As an example, in Fig. 1 the statistics $p(u_\tau)$ of wind gusts is shown. Here, a wind gust is measured as the change of the wind speed $u_\tau = u(t + \tau) - u(t)$ over a short time span, here $\tau = 4$ sec. $u(t)$ denotes the velocity of the wind. Besides the measured probability density, a Gaussian distribution is shown, which has the same mean value and standard deviation. Comparing these two distributions, a huge difference in the probability of large changes in the wind speed can be seen. The higher risk due to the heavy tailed statistics will be discussed in Sect. 2 together with the question, whether the heavy tail or the probability of extreme events, respectively, follows a power law.

In the Sects. 4 and 5 we will discuss the question of what kind of mechanism may lead to such anomalous statistics. First, we will briefly summarize some features of Levy processes. In contrast to the Levy process we will present (Sect. 4) the actual understanding of generating disorder in turbulence by hierarchical or cascade-like processes. A new way to analyze such hierarchical processes will be presented in Sect. 5. As further examples experimental and empirical results from turbulence and from the financial market data are presented and discussed.

## 2 Heavy tailed statistics

Typical features of heavy tailed statistics will be discussed next. For this, let us come back to the probability density of changes of the wind speed over a short time interval, as shown in Fig. 1.\(^1\) Note that the most extreme events correspond to a change of the wind speed by 8 m/sec during 4 sec. Thus, the heavy tail of this probability corresponds to the statistics of wind gusts. Here, we get the impression that wind gusts are not an independent coherent structure of the wind field, which may be well separated from a turbulent background, but the wind gusts seem to emerge naturally from the heavy tailed statistics. If we compare this empirical probability distribution of the wind speed changes with a Gaussian distribution, which is uniquely given by the same mean value and the same variance, as shown by a solid curve in Fig. 1, we see that the Gaussian distribution drastically underestimates the occurrence probability of the large events. One should note that a factor $10^6$, as indicated by an arrow in Fig. 1, corresponds to the number of hours about one century consists of. Thus, the gusts we measured each hour would be expected as an event of the century for a corresponding wind field with Gaussian statistics. We conclude that for heavy tailed statistics the occurrence of extreme events is drastically enhanced, which is an important issue for the estimation of risks (see for example [3]). This effect is often not seen at once. To demonstrate this, we have shown in Fig. 2 the probability distributions of Fig. 1 with a linear $y$-axis. In addition a second Gaussian distribution with

---

\(^1\) In this paper we consider only probability densities $p(x)$, where $p(x)dx$ gives the probability to find a value $x \in [x, x + dx]$. We renounce to discuss properties of the probabilities $P(x > X) = \int_X^\infty p(y)dy$.

\(^2\) This statistics was obtained from a wind measurement at the German North Sea cost line near the town Emden. The data were measured by means of an ultrasonic anemometer at 20 m height. The sampling frequency was 4 Hz. Data were taken over 275 hours in October 1997. For further details see [3].
a too small standard deviation is shown. Two remarks should be added to Fig. 2: first, in this presentation of the probabilities it seems that there are no big problems in the statistics of the tails. Second, it should be noted that it may become a problematic issue to estimate correctly the standard deviation of a distribution, as will be discussed below. In the latter case it becomes obviously difficult to determine the corresponding Gaussian distribution.

For assessment of risk, we ask what will be the probability of an extreme event, which for example should be twice as large as an observed event. In the case of Gaussian distribution we obtain: $p(2x) = (p(x))^4$. Thus, for an event that has a probability of $10^{-3}$, the event, which is twice as large, is $10^{-9}$ times less probable. This means that the probability of more extreme events becomes very unlikely. For our wind gust measurements we find clearly a different behavior. From Fig. 1 it is easily seen that for doubling the size of the wind-speed changes the probability decreases only by a factor $10^{-2}$ at the most. This is due to the much weaker decay of the tails of the distribution. Consequently a larger event than, let us say, an extreme event of a decade is still quite likely to occur.

As a second feature of heavy tailed statistics, we consider the law with which the tails decay. For Gaussian distributions we know that they decay with $\exp(-x^2)$. If it is assumed that the obtained heavy tailed probability distribution is based on an underlying stable stochastic process, which we will define below, then a deviation from Gaussian statistics may belong to Levy statistics. This has the consequence that the tails decay with a power law:

$$p(x) \propto \frac{1}{x^{\beta}} \quad \text{for large } x \text{ and } 1 < \beta < 3.$$  \hfill (1)

Distributions with power law tails will have diverging moments for $n > \beta - 1$, as can be seen easily from the definition of moments of a probability density function:

$$\langle x^n \rangle = \int_{-\infty}^{\infty} x^n p(x) \, dx.$$  \hfill (2)

For central moments, the integration has to be done over $(x - \langle x \rangle)^n$ instead of $x^n$. Only for $\beta \geq 3$ the variance would be defined. If we come back to our example of presenting experimental results by $(\langle x \rangle \pm \sigma)$, we see that if the measured variable $x$ has such a heavy tailed power law statistics, the presentation of the results by a mean value and its error bar $(\langle x \rangle \pm \infty)$ does not make any sense any more. Here, the reader should notice that based on a finite set of measured data, measured during a finite time interval, all moments can be evaluated even for such a heavy tailed distribution with power law decay. The accurate investigation for heavy tailed power law statistics will give non converging results, if the number of data points is increased.

The problem of diverging second moments becomes more explosive, if we take as measured data geophysical data, like for example earth quakes, wind gusts, or flood disasters. If the occurrence probability of these events underlies a heavy tailed power law statistics, it is easy to see how difficult a risk prediction
becomes. The question arises, what kind of information do we get from a typical value for an extreme event occurring each 100 years if the variance is diverging?

Statistics with vanishing or undefined variance do not only cause problems in presenting results but also in modelling such systems. For example, for the problem of modeling extremal events for insurance and finance it was noted by P. Embrechts et. al. in [3]: “... much of the finance literature is based on the notations of volatility and correlations, i.e. finite second moments are necessarily required in such models, and therefore the infinite variance case has gained only marginal popularity. ... (but) it is a fact that most financial data are heavy-tailed!”

At last, we want to comment on the question of how we can decide whether a given statistics belongs to the class of Levy distributions with power law behavior in the tails or not. A simple but not very accurate way is to investigate the decay of the tails in a double-logarithmic plot, as shown for the wind data in Fig. 3. One problem arises from the statement that the power law should hold for large values, see (1). For the data shown in Fig. 3 one may try to see an indication of such a power law, for \( u_\tau > 4\sigma \). Two critical remarks must be made, firstly, the power law extends over less than one decade, secondly, it is based on that part of the probability density which is estimated from a small number of events, where naturally uncertainties increase. To show clear evidence of a power law tail, more careful investigations have to be done (see for example [4]). Finally, we want to note that the straight line in Fig. 3 corresponds to a slope larger than 3, which we take as a hint that no Levy statistics is present for our wind data. Here, we conclude that the wind gust statistics we have presented represents a case of heavy tailed probability which does not obey power law in the tails. In contrast to this wind statistics, there are many known cases where power law tails are reported. Among these are earthquakes, wealth distribution, internet links, just to mention some. Often one also speaks of Pareto distributions instead of statistics with power law tails, for a recent overview see [5].

### 3 Stable Levy processes

Next, a short summary of features of Levy processes will be given. A stochastic process describes the way how, for example in time, random variables \( x_i \) are created. The class of Levy processes belong to stable processes. A stochastic process of a variable \( x \) with its probability density function \( p(x) \) is called stable, if the statistics of a sum of \( x, y = x_1 + \ldots + x_n \) has a pdf \( p(y) \) which can be transformed to the pdf \( p(x) \) by a linear rescaling transformation (cf. [3,5]). A linear transformation does not change the law of the tails. These Levy distributions are characterized by heavy tails with the above mentioned power laws (1), where \( 1 < \beta < 3 \). For \( \beta = 3 \) Gaussian distributions are obtained.

For distributions \( p(x) \) with power law tails and \( \beta > 3 \) the variance is defined. This has the consequence that data following a statistics with a power law tail having \( \beta \geq 3 \) are not stable. These data have finite variance and due to the Central Limit Theorem its sum variable \( y \) converges to a Gaussian distribution.
4 Turbulence and financial market data

In the following, we present some results from fully developed turbulence and from data of the financial market. The explanation of the complex disorder of turbulence will lead us to the idea of a cascade, which may be seen as a hierarchical process. With such hierarchical or cascade processes we will present in Sect. 5 an alternative way to characterize in a quite complete way the disorder of our considered systems, including their heavy tailed statistics.

The profound understanding of turbulence is up to now regarded as an unsolved problem. Although the basic equations of fluid dynamics, namely the Navier Stokes equations, are known for more than 150 years, a general solution of these equations for high Reynolds numbers, i.e. for turbulence, is not known. Even with the use of powerful computers no rigorous solutions can be obtained. Thus for a long time there has been the challenge to understand at least the complexity of idealized turbulence, which is taken to be isotropic and homogeneous. This case will lead us to the well known intermittency problem of turbulence, which is nothing else than the occurrence of heavy tailed statistics. The central question is to understand the mechanism which leads to this anomalous statistics (see [6,7]).

The intermittency problem of turbulence can be reduced to the question about the statistics of the velocity differences over different distances, measured by the increments the mechanism which leads to this anomalous statistics (see [6,7]).

By simple dimensional arguments it follows that

\[ \langle q(l,x)^n \rangle = \left( \epsilon^{n/3} \right)^{n/3}. \]

The simplest ansatz is to take \( \epsilon \) as constant, thus the Kolmogorov scaling \( n/3 \) of 1941 is obtained [8]. Assuming a lognormal distribution for \( \epsilon \), i.e. not \( \epsilon \) but \( l \epsilon \) has a Gaussian distribution, Kolmogorov and Oboukhov proposed for \( \left( \epsilon^{n/3} \right) \) an additional scaling term, leading to the so-called intermittency correction [9]

\[ \langle q(l,x)^n \rangle = l^{\xi_n} \] with \( \xi_n = \frac{n}{3} - \frac{n(n-3)}{18} \) and \( n > 2 \) (4)

Here we want to point out a direct consequence of this nonlinear scaling exponent \( \xi_n \), namely, that the probability densities of \( p(q(l,x)) \) cannot be Gaussian. This point is easily seen, if the definition of eq. (2) of the moments is taken and if it is evaluated only for Gaussian distributions. To fulfill the scaling relation eq. (4), we must take for the scale dependence of the variance of the Gaussian distributions: \( \langle q(l,x)^2 \rangle = l^{2\alpha} \) (\( \alpha \) is any real number, for turbulence close to 2/3.). With these points it is easy to calculate eq. (4) for all \( n \).

Here it should be noted that the term “intermittency” is used frequently in physics for different phenomena, and may cause confusions. This turbulent intermittency is not equal to the intermittency of chaos. There are also different intermittency phenomena introduced for turbulence. There is this intermittency due to the nonlinear scaling, there is the intermittency of switches between turbulent and laminar flows for non local isotropic fully developed turbulent flows, there is the intermittency due to the statistics of small scale turbulence which we discuss here as heavy tailed statistics.
Fig. 4  Autocorrelation of velocity increments $\langle q(l, x) q(l, x + \rho) \rangle / \langle (q(l, x))^2 \rangle$ for the data of Fig. 5; $l = 0.1 L_0$ is about the Taylor length. The corresponding correlation length is $0.05 L_0$.

Fig. 5  Comparison of the numerical solution of the Fokker-Planck equation (solid lines) for the pdfs $p(q(x), l)$ with the pdfs obtained directly from the experimental data (bold symbols). The scales $l$ are (from top to bottom): $l = L_0$, $0.6 L_0$, $0.35 L_0$, $0.2 L_0$ and $0.1 L_0$. The distribution at the largest scale $L_0$ was parameterized (dashed line) and used as initial condition for the Fokker-Planck equation ($L_0$ is the correlation length of the turbulent velocity signal). The pdfs have been shifted in vertical direction for clarity of presentation and all pdfs have been normalized to their standard deviations; after [10].

One finds that $\xi_n = n \alpha / 2$, i.e., there is no nonlinear intermittency correction. Saying this in another way, the nonlinear $n$-dependence of $\xi_n$ means nothing else than that the probabilities $p(q(l, x))$ must change their form with the scale parameter $l$, a finding heavily discussed in the last decades. We see in Fig. 5 that for large scales ($l \approx L_0$) the distributions are nearly Gaussian. As the scale is decreased the probability densities become more and more heavy tailed.

Based on this finding, we can argue why a Levy process is not consistent with the statistics of velocity increments of turbulence. As already mentioned, the Levy statistics is based on stability. Thus, the sum of the variable should have the same statistics. The sum of two successive increments $q(l, x) + q(l, x + l) = (u(x + l) - u(x)) + (u(x + 2l) - u(x + l))$ is nothing else than the increment $q(2l, x)$ of a scale twice as large. We have just seen that $p(q(2l, x))$ has a different form, i.e. the process is not stable. It is shown in Fig. 4, that two successive increments are statistically independent.

Next, we present some anomalous statistical features of data from the financial market. These features are astonishingly similar to the just discussed intermittency of turbulence [12]. The following analysis is based on a data set $Y(t)$, which consists of 1 472 241 quotes for US dollar-German Mark exchange rates from the years 1992 and 1993. Many of the features we will discuss here are also found in other financial data like for quotes of stocks. One central issue is the understanding of the statistics of price changes over a certain time interval $\tau$ which determines losses and gains. The changes of a time series of quotations $Y(t)$ are commonly measured by returns $r(\tau, t) := Y(t + \tau) / Y(t)$, logarithmic returns or by increments $q(\tau, t) := Y(t + \tau) - Y(t)$. [13] The moments of these quantities often display power law behavior similar to the just discussed Kolmogorov scaling for turbulence, cf. [14–16]. In addition one finds

© 2004 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
for the probability distributions an increasing tendency to heavy tailed probability distributions for small $\tau$ (see Fig. 6). This represents the high frequency dynamics of the financial market. The identification of the underlying process leading to these heavy tailed probability density functions of price changes is a prominent puzzle (see [3,15–19]), like it is for turbulence. One should also remark that the scaling properties of the moments and the discussion of the form of the probabilities can be set in close connection with multifractal features of disordered systems. It is the nonlinearity of the scaling exponents $\xi_n$ which is closely related to the multifractal spectrum of fractal dimensions.

5 Stochastic process for scale dependent complexity

In Sect. 2 we discussed the meaning of the heavy tailed probability distributions mainly based on power law tails and Levy statistics. Here, we want to present an alternative, non conventional way to characterize anomalous statistics, which is based on the idea of a cascade process as discussed in the preceding section. In particular we show how from given data a stochastic process can be estimated. This makes it possible to model the measured statistics quite accurately. Guided by the finding that the statistics changes with scale, as shown in Figs. 5 and 6, we consider the evolution of the quantity $q(l,x)$, or $q(\tau,t)$ with the variable $l$ or $\tau$ respectively. Thus, we will consider processes which evolve in the scale variable. This we take as a stochastic cascade process.

For the following, we present a generalized discussion. Let us denote the given complex structure by $Y(x)$. With respect to the preceding sections, $Y(x)$ denotes the velocity of a turbulent flow at the location $x$, or the price of a stock at the time “$x$”. To characterize the disorder of the structure $Y(x)$ a local scale resolving quantity $q(l,x)$ is considered. This quantity $q$ is here defined as an increment $q(l,x) := Y(x+l) - Y(x)$, but may also be a wavelet, a local measure, or the sum of the square of derivatives of $Y$, just to mention some possibilities. For time series, like the financial market data, $l$ stands for the time lag.

Note, the moments of $p(q)$ are called structure functions if $q$ is an increment. It is easily seen that the second moment corresponds to the autocorrelation function $\langle Y(x)Y(x+l) \rangle$, which is related by the Wiener–Khintchine theorem to the power spectrum of $Y(x)$. Higher order moments correspond to higher order two point correlations.

For one fixed scale $l$ we get the scale dependent disorder by the statistics of $q(l,x)$. The complete stochastic information of the disorder on all length scales is given by the joint probability density function

$$p(q_1, \ldots, q_n),$$

Fig. 6 Probability densities (pdf) $p(q(t), \tau)$ of price changes $q(\tau,t) = Y(t + \tau) - Y(t)$ for the time delays $\tau = 5120, 10240, 20480, 40960$ s (from bottom to top). Symbols: results obtained from the analysis of middle prices of bit-ask quotes for the US dollar-German Mark exchange rates from October 1st, 1992 until September 30th, 1993. Full lines: results from a numerical iteration of the Fokker-Planck equation (13); the probability distribution for $\tau = 40960$ s (dashed lines) was taken as the initial condition. The pdfs have been shifted in vertical direction for clarity of presentation and all pdfs have been normalized to their standard deviations; after [11].
where we set \( q_i = q(l_i, x) \). Without loss of generality we take \( l_i < l_{i+1} \). This joint probability may be seen in analogy to joint probabilities of statistical mechanics (thermodynamics), describing in the most general way the occupation probabilities of the microstates of \( n \) particles, where \( q \) is a six-dimensional state vector.

Next, the question is discussed, whether it is possible to simplify the joint probability. In general such joint probabilities can be expressed by conditional probabilities:

\[
p(q_1, \ldots, q_n) = p(q_1|q_2, \ldots, q_n)p(q_2|q_3, \ldots, q_n) \cdots p(q_{n-1}|q_n).
\]

A simplification is given, if for the multiple conditioned probabilities

\[
p(q_i|q_{i+1}, \ldots, q_n) = p(q_i|q_{i+1})
\]

holds. Eq. (7) is known as the condition for a Markov process evolving from state \( q_{i+1} \) to the state \( q_i \), i.e. from scale \( l_{i+1} \) to \( l_i \).

A “single particle approximation” would correspond to the condition:

\[
p(q_1, \ldots, q_n) = p(q_1)p(q_2) \cdots p(q_n).
\]

Based on eqs. (6) and (7), eq. (8) holds if \( p(q_i|q_{i+1}) = p(q_i) \). Only for the last case, the knowledge of \( p(q_i) \) is sufficient to characterize the complexity of the whole system. Otherwise an analysis of only \( p(q_i) \) is a non unique projection of the whole complexity of the system. Such a projection is done frequently like for the investigations of the moments as presented in eqs. (3) and (4). This is also the case for the characterization of complex structures by means of fractality or multifluctuality (cf. for multifluctuality [20], for turbulence [6,7], for financial market [14]). The scaling analysis of moments as indicated for turbulence in Eq. (4) provides a complete knowledge of any joint \( n \)-scale probability density only if eq. (8) is valid. One gets a much more general characterization of the disorder if the transition probabilities from one scale to a smaller one is investigated and if eq. (7) is fulfilled.

In the following, we focus on the proposed more complete statistical characterization. The basic idea is to consider the evolution of \( q(l, x) \) at one fixed point \( x \) as a stochastic process in scale \( l \). (The extensive discussion of the analysis of financial and turbulent data can be found in [10,11]).

To prove that a Markov process is given, it has to be shown that (7) holds. For a given data set \( Y(x) \) it is easy to evaluate the corresponding conditional probabilities by the calculation of \( q(l, x) \) at different points \( x \) for several fixed scales \( l \). Having shown that the multiconditioned probabilities are equal to the single conditioned probabilities, the evolution of the conditional probability density \( p(q, l|q_0, l_0) \) starting from the initial scale \( l_0 \) follows

\[
-l \frac{\partial}{\partial l} p(q, l|q_0, l_0) = \sum_{k=1}^{\infty} \left( - \frac{\partial}{\partial q} \right)^k D^{(k)}(q, l) p(q, l|q_0, l_0).
\]

(The minus sign on the left side is introduced, because we consider processes running to smaller scales \( l \), furthermore we multiply the stochastic equation by \( l \), which leads to a new parameterization of the cascade by the variable \( ln(1/l) \), a simplification for a process with scaling law behavior of its moments.) This equation is known as the Kramers-Moyal expansion [23]. The Kramers-Moyal coefficients \( D^{(k)}(q, l) \) are defined as the limit \( \Delta l \rightarrow 0 \) of the conditional moments \( M^{(k)}(q, l, \Delta l) \):

\[
D^{(k)}(q, l) = \lim_{\Delta l \rightarrow 0} M^{(k)}(q, l, \Delta l),
\]

\[
M^{(k)}(q, l, \Delta l) := \frac{l}{k!\Delta l} \int_{-\infty}^{+\infty} (q_\Delta - q)^k p(q_\Delta, l - \Delta l|q, l) \, dq_\Delta.
\]

Thus, for the estimation of the \( D^{(k)} \) coefficients it is only necessary to estimate the conditional probabilities \( p(q_\Delta, l - \Delta l|q, l) \). For a general stochastic process, all Kramers-Moyal coefficients are different from...
zero. According to Pawula’s theorem, however, the Kramers-Moyal expansion stops after the second term, provided that the fourth order coefficient \( D^{(4)} \) vanishes. In that case the Kramers-Moyal expansion is reduced to a Fokker-Planck equation (also known as Kolmogorov equation [24]):

\[
-l \frac{\partial}{\partial l} p(q, l|q_0, l_0) = \left\{ -\frac{\partial}{\partial q} D^{(1)}(q, l) + \frac{\partial^2}{\partial q^2} D^{(2)}(q, l) \right\} p(q, l|q_0, l_0). \tag{12}
\]

\( D^{(1)} \) is denoted as drift term, \( D^{(2)} \) as diffusion term. The probability density function \( p(q, l) \) has to obey the same equation:

\[
-l \frac{\partial}{\partial l} p(q, l) = \left\{ -\frac{\partial}{\partial q} D^{(1)}(q, l) + \frac{\partial^2}{\partial q^2} D^{(2)}(q, l) \right\} p(q, l). \tag{13}
\]

We remind the reader that the Fokker-Planck equation describes the probability density function of a stochastic process generated by a corresponding Langevin equation (we use Itô’s definition)

\[
-l \frac{\partial}{\partial l} q(l) = D^{(1)}(q, l)/l + \sqrt{D^{(2)}(q, l)/l} f(l) \tag{14}
\]

where \( f(l) \) is white noise (Gaussian distributed and delta correlated). This feature provides an approach to simulate numerically typical scale dependent data by the knowledge of the process coefficients, as it was proposed for example in [25] and shown in [26].

Let us return to the discussion of scale dependent complexity. We have indicated a way to characterize this complexity by means of stochastic processes. The new point is that we consider processes running in the scale variable and not in time as usually. The presented features of the stochastic process have been well known for decades [24]. A strength of this analysis is that it does not depend on assumptions. By the estimation of the conditional probabilities, the Markovian property (7) as well as the Kramers-Moyal coefficients can be evaluated [21, 22]. Knowing the evolution equation (9), the n-increment statistics \( p(q_1, \ldots, q_n) \) is known, too. Definitely, the information like scaling behavior of the moments of \( q(l, x) \) can also be extracted from the knowledge of the process equations. Multiplying (9) by \( q^n \) and successively integrating over \( q_0 \) and \( q \), an equation for the moments is obtained:

\[
-l \frac{\partial}{\partial l} \langle q^n \rangle = \sum_{k=1}^{n} \left( -\frac{\partial}{\partial q} \right)^k \frac{n!}{(n-k)!} \langle D^{(k)}(q, l) q^{n-k} \rangle. \tag{15}
\]

Scaling, i.e. multi affinity, is obtained if \( D^{(k)}(q, l) \propto q^k \), see [27].

Based on this procedure we were able to reconstruct directly from the given data the corresponding stochastic processes. Knowing these processes one can perform numerical solutions (see [10, 11]). In Figs. 5 and 6 the numerical solutions are shown by solid curves.

We see that the heavy tailed structure of the probabilities is well described by this approach over a Fokker-Planck equation, which may be taken as a verification of the above described way to reconstruct this process from pure data analysis.

For financial as well as for the turbulence data it was found that the diffusion term is quadratic in the state space of the scale resolved variable. With respect to the corresponding Langevin equation, the multiplicative nature of the noise term becomes evident, which causes heavy tailed probability densities and eventually multifractal scaling. The scale dependency of drift and diffusion term corresponds to a non-stationary process in scale variables \( \tau \) and \( l \), respectively. From this point we conclude again that a Levy statistics for one fixed scale, i.e. for the statistics of \( q(x, l) \) for fixed \( l \) is not the adequate class of statistics.

Comparing the tips of the distributions at small scales in Figs. 5 and 6 one finds a less sharp tip for the turbulence data. This finding is in accordance with a larger additive term in the diffusion term \( D^{(2)} \) for the turbulence data. Knowing that \( D^{(2)} \) has an additive and quadratic \( q \) dependence it is clear that for small \( q \) values, i.e. for the tips of the distribution, the additive term dominates. Taking this result in combination with the Langevin equation, we see that for small \( q \) values Gaussian noise is present, which leads to a Gaussian tip of the probability distribution, as found in Fig. 5.
6 Conclusion

A discussion of anomalous statistics has been presented. Features of heavy tailed statistics and Levy processes have been outlined, especially the meaning of diverging moments like the divergence of the variance. The discussion of heavy tailed probability distributions in turbulence and financial data brought us to another class of disordered systems, which is governed by cascade or hierarchical processes leading to anomalous statistics of small scale increments. A method for reconstructing the equations of stochastic processes directly from given data sets has been presented. The reconstruction is based on the estimation of conditional probabilities, from which the drift and diffusion coefficient could be estimated.

Knowing the form of the drift and diffusion term, further quantities of the data can be evaluated. For example, it is easy to derive the equation for the moments (2) (which are often used to characterize the complexity of financial and turbulent data).

In addition to the above mentioned traditional analysis of the moments, the Markovian process provides information on the n-point probability distribution for n different \((q_i, l_i)\) (see eqs. (6) and (7)). From this point of view, a complete statistical characterization of the scale resolved complexity can be provided by the proposed method.

As a last comment we want to point out that it is evident that this method can be applied to time series. In this case, nonlinear dynamic equations can be reconstructed from experimental data even if dynamical noise is present. Although one may find in the actual literature several suggestions for how to reconstruct dynamic equations from empirical data sets, it is interesting to note that the mathematical frame of the method we have discussed here and which works very well [28–30] has been given by Kolmogorov already in 1931 [24].

Acknowledgements We acknowledge helpful discussions with Ch. Renner, St. Lück, M. Siefert, M. Wächter and R. Friedrich.

References