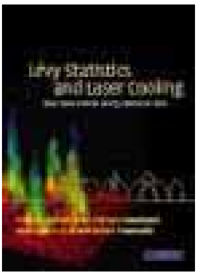

Lévy Statistics and Laser Cooling

How Rare Events Bring Atoms to Rest

François Bardou, Jean-Philippe Bouchaud, Alain Aspect, Claude Cohen-Tannoudji

Book DOI: http://dx.doi.org/10.1017/CBO9780511755668

Online ISBN: 9780511755668

Hardback ISBN: 9780521808217

Paperback ISBN: 9780521004220

Chapter

4 - Broad distributions and Lévy statistics: a brief overview pp. 42-59

Chapter DOI: http://dx.doi.org/10.1017/CBO9780511755668.005

Cambridge University Press

4

Broad distributions and Lévy statistics: a brief overview

In this chapter, we introduce the main concepts and tools of Lévy statistics that will be used in subsequent chapters in the context of laser cooling. In Section 4.1, we show how statistical distributions with slowly decaying power-law tails can appear in a physical problem. Then, in Section 4.2, we introduce the generalized Central Limit Theorem enabling one to handle statistically 'Lévy sums', i.e. sums of independent random variables, the distributions of which have power-law tails. We also sketch, in a part that can be skipped at first reading, the proof of the theorem and present a few mathematical properties concerning distributions with power-law tails and Lévy distributions. In Section 4.3, we present some properties of Lévy sums which will turn out to be crucial for the physical discussion presented in subsequent chapters: the scaling behaviour, the hierarchy and fluctuation problems. These properties are illustrated using numerical simulations. Finally, in Section 4.4, we present the distribution $S(t)$, called the 'sprinkling distribution'. This distribution presents unexpected features and will play an essential role in the following chapters.

4.1 Power-law distributions. When do they occur?

Situations where broad distributions appear and where rare events play a dominant role are more and more frequently encountered in physics, as well as in many other fields, such as geology, economy and finance. The term 'broad distributions' usually refers to distributions decaying very slowly for large deviations, typically as a power law, implying that some moments of the distribution are formally infinite.

The paradigm problem concerning these types of random variables is the behaviour of the sum of a large number of them. For example, in the problem of interest here, the total experimental time can be decomposed into a sum of the time intervals corresponding to the trapping region and to the external region. Precise theorems govern the properties of these sums, generalizing the well known

(Gaussian) Central Limit Theorem (CLT). We shall not state these results in full generality (the reader can consult [GnK54, BoG90]), but rather focus on the case relevant to our purpose. We shall thus restrict our discussion to *positive* random variables τ (representing random *times*), distributed for large τ as:

$$P(\tau) \underset{\tau\to\infty}{\simeq} \frac{\mu \tau_{\mathrm{b}}^{\mu}}{\tau^{1+\mu}} \tag{4.1}$$

where τ_{b} sets the scale of the phenomenon, and μ is an exponent describing how fast the distribution decays to 0. (The extra factor μ in the numerator is included for later convenience.) To normalize the distribution P, $\mu > 0$ is required. All the moments $\langle \tau^q \rangle = \int_0^\infty \mathrm{d}\tau\, \tau^q P(\tau)$ such that $q \geq \mu$ are divergent. The most interesting case, as we shall see below, is the case where $\mu \leq 1$, for which the mean value $\langle \tau \rangle$ of τ is infinite.

When do such power-law distributions occur? They sometimes result from the highly complex underlying dynamics of the physical system, as in chaotic systems [KSZ96, Zas99], and models of avalanches or earthquakes [Bak96, BoC97].

Another frequent scenario for creating power-law distributions is a change of variable. A first variable a, which is naturally sampled by the physical process, is distributed according to a law which may be of any type (Gaussian, exponential, uniform, ...), but the distribution of a related physical quantity $b = f(a)$ turns out to be a power law for certain types of (non-linear) functions $f(a)$. A first example of such a situation was given in Section 3.3. While the probability of reaching a small momentum p is approximately uniform, the lifetime $\tau \propto p^{-\alpha}$ of the corresponding p states is distributed according to a power law, thus leading to eq. (4.1) with $\mu = D/\alpha$.

Another interesting example arising from a change of variable is thermal activation out of a deep potential valley [Shl88, BoD95]. The Arrhenius law states that the average exit time τ is proportional to $\tau_0 \exp(E/k_{\mathrm{B}}T)$, where E is the energy barrier, T the temperature and τ_0 a typical time. In disordered systems, the barriers E are themselves random variables which are often distributed according to an exponential law: $\Pi(E) = E_0^{-1} \exp(-E/E_0)$. The resulting distribution $P(\tau)$ of exit times τ, which is given by $P(\tau)\mathrm{d}\tau = \Pi(E)\mathrm{d}E$ with $\tau = \tau_0 \exp(E/k_{\mathrm{B}}T)$, is thus equal to

$$\begin{aligned} P(\tau) &= \frac{1}{E_0} \exp\left(-\frac{E}{E_0}\right) \frac{k_{\mathrm{B}}T}{\tau_0} \exp\left(-\frac{E}{k_{\mathrm{B}}T}\right) \\ &= \frac{k_{\mathrm{B}}T}{E_0}\, \frac{\tau_0^{\mu}}{\tau^{1+\mu}}. \end{aligned} \tag{4.2}$$

We get an expression similar to (4.1) with $\mu = k_{\mathrm{B}}T/E_0$ and $\tau_{\mathrm{b}} = \tau_0$. Interestingly, for $k_B T < E_0$, the average relaxation time is infinite, leading to strongly anomalous dynamics (see below, and [Bou92, BoD95, BCK97, Bou00]).

Notice that the above derivation of eq. (4.2) assumes that the exit times τ are deterministically fixed by the height E of the barrier. In parallel with the results of Chapter 3 (Section 3.3.1), the result (4.2) is not dramatically altered if the exit times are distributed as an exponential with an average given by the Arrhenius law.

4.2 Generalized Central Limit Theorem

4.2.1 Lévy sums. Asymptotic behaviour and Lévy distributions

Let T_N be the sum of N independent positive random variables, all distributed according to the distribution $P(\tau)$ of eq. (4.1):

$$T_N = \sum_{i=1}^{N} \tau_i. \tag{4.3}$$

When $\mu > 2$, the usual form of the CLT is valid since both the mean value $\langle \tau \rangle$ and the variance $\sigma^2 = \langle \tau^2 \rangle - \langle \tau \rangle^2$ exist. Defining a new variable ξ by

$$T_N = \langle \tau \rangle N + \sigma \sqrt{N}\, \xi, \tag{4.4}$$

the CLT then says that, for large N, ξ tends to a dimensionless Gaussian random variable with zero mean value and unit variance, i.e. it is distributed according to $G(\xi) = (2\pi)^{-1/2} \exp(-\xi^2/2)$ ('normal' distribution). More precisely, one has, independently of the shape of $P(\tau)$:

$$\lim_{N\to\infty} \mathcal{P}\left(\xi_1 \le \frac{T_N - \langle \tau \rangle N}{\sigma\sqrt{N}} \le \xi_2\right) = \int_{\xi_1}^{\xi_2} \mathrm{d}\xi\, G(\xi). \tag{4.5}$$

We note that the second (fluctuating) term in eq. (4.4) is negligible compared to the first one when $N \to \infty$.

For $\mu < 2$, the mean value $\langle \tau \rangle$ and/or the variance σ^2 diverge and eq. (4.5) is no longer valid. The CLT has been generalized by Lévy and Gnedenko, and gives results which are *independent of the detailed shape* of $P(\tau)$ and which depend only on the long time behaviour described by eq. (4.1). The sums T_N are called 'Lévy sums'. We now state a few important results concerning the asymptotic behaviour of these Lévy sums (for large N). A sketch of the proof of these results will be presented in the next section, using the properties of the Laplace transforms of functions with power-law tails.

The generalized CLT takes two different forms for $1 < \mu < 2$ and for $\mu < 1$[1].

[1] Logarithmic corrections appear in the cases $\mu = 1$ and $\mu = 2$, requiring a separate discussion (see Appendix C).

Consider first the case $1 < \mu < 2$, where τ has a finite mean value $\langle\tau\rangle$ but an infinite variance. If we introduce a new variable ξ by

$$1 < \mu < 2: \qquad T_N = \langle\tau\rangle N + \xi\tau_{\mathrm{b}} N^{1/\mu}, \tag{4.6}$$

then the generalized CLT states that ξ is a random variable of order one, distributed for large N according to a function $L_\mu(\xi)$ which depends only on μ and which is called the 'completely asymmetric' Lévy distribution of index μ[2]. More precisely, we can write

$$\lim_{N\to\infty} \mathcal{P}\left(\xi_1 \le \frac{T_N - \langle\tau\rangle N}{\tau_{\mathrm{b}} N^{1/\mu}} \le \xi_2\right) = \int_{\xi_1}^{\xi_2} \mathrm{d}\xi\, L_\mu(\xi). \tag{4.7}$$

Note that the second (fluctuating) term in eq. (4.6) is still negligible compared with the first when $N \to \infty$. The Lévy distributions $L_\mu(\xi)$ have simple Laplace transforms[3]:

$$\mathcal{L}L_\mu(u) = \int_0^\infty \mathrm{d}\xi\, L_\mu(\xi)\mathrm{e}^{-u\xi} = \exp(-b_\mu u^\mu) \quad \text{with} \quad b_\mu = \frac{(\mu-1)\Gamma(1-\mu)}{\mu}. \tag{4.8}$$

In the case $\mu < 1$, both the mean value and the variance of τ diverge and one finds that T_N grows faster than the number of terms N. Equation (4.6) has to be replaced by

$$\mu < 1: \qquad T_N = \xi\tau_{\mathrm{b}} N^{1/\mu}, \tag{4.9}$$

and one finds that ξ is again a random variable of order one, distributed for large N according to a Lévy distribution $L_\mu(\xi)$, whose Laplace transform is now:

$$\mathcal{L}L_\mu(u) = \int_0^\infty \mathrm{d}\xi\, L_\mu(\xi)\mathrm{e}^{-u\xi} = \exp(-b_\mu u^\mu) \quad \text{with} \quad b_\mu = \Gamma(1-\mu). \tag{4.10}$$

The analogue of eq. (4.7) is:

$$\lim_{N\to\infty} \mathcal{P}\left(\xi_1 \le \frac{T_N}{\tau_{\mathrm{b}} N^{1/\mu}} \le \xi_2\right) = \int_{\xi_1}^{\xi_2} \mathrm{d}\xi\, L_\mu(\xi). \tag{4.11}$$

4.2.2 Sketch of the proof of the generalized CLT

We try now to give an idea of the mathematical properties leading to the very simple forms (4.8) and (4.10) for the Laplace transforms of the Lévy distributions $L_\mu(\xi)$.

The fact that Laplace transforms play an important role in this problem is easy

[2] Since τ is positive, L_μ is actually a particular case ('completely asymmetric') of more general Lévy distributions, which arise when the random variable involved in the summation has power-law tails both at $+\infty$ and at $-\infty$.

[3] We denote the Laplace transform of f by $\mathcal{L}f$.

to understand. Let $\Pi_N(T_N)$ be the probability distribution of the Lévy sum T_N. It can be written:

$$\Pi_N(T_N) = \int \mathrm{d}\tau_1 \ldots \mathrm{d}\tau_N \; P(\tau_1) \ldots P(\tau_N)\, \delta\left(\sum_{i=1}^{N} \tau_i - T_N\right) \tag{4.12}$$

where the constraint on the value of the sum is imposed through a δ-function. In fact, the right-hand side of eq. (4.12) is a convolution product of N functions $P(\tau)$, so that the Laplace transform $\mathcal{L}\Pi_N(s)$ of $\Pi_N(T_N)$ is nothing but the N^{th} power of the Laplace transform $\mathcal{L}P(s)$ of $P(\tau)$[4]:

$$\mathcal{L}\Pi_N(s) = \left[\int_0^\infty \mathrm{d}\tau \;\; P(\tau)\mathrm{e}^{-s\tau}\right]^N = [\mathcal{L}P(s)]^N. \tag{4.13}$$

We now use the fact that $P(\tau)$ is a probability distribution, i.e. takes positive values and is normalized to one. This implies that $\mathcal{L}P(s) \leq 1$ for any $s \geq 0$, the upper bound being obtained for $s = 0$. Since $\mathcal{L}P(s)$ is raised to a high power N in eq. (4.13), one expects that $\mathcal{L}\Pi_N(s)$, which is equal to one for $s = 0$, will be appreciable only in the neighbourhood of $s = 0$. This explains the importance in this problem of the small-s behaviour of $\mathcal{L}P(s)$, which is itself determined by the long-τ behaviour of $P(\tau)$.

We will focus here on distributions (4.1) with $\mu < 1$. We suppose in addition that the subleading corrections to eq. (4.1) decay faster than τ^{-2} for large τ. One can then show that the small-s behaviour of their Laplace transforms $\mathcal{L}P(s)$ is given by

$$\mathcal{L}P(s) \underset{s\to 0}{=} 1 - \Gamma(1-\mu)\,(\tau_{\mathrm{b}}s)^\mu - A_0\tau_{\mathrm{b}}s + \cdots \tag{4.14}$$

where A_0 is a constant. In view of its importance here, a brief proof of this result will be given in point (ii) of Section 4.2.3.

Using eq. (4.13), one gets:

$$\mathcal{L}\Pi_N(s) \underset{s\to 0}{=} \left[1 - \Gamma(1-\mu)\,(\tau_{\mathrm{b}}s)^\mu + O(\tau_{\mathrm{b}}s)\right]^N. \tag{4.15}$$

Setting $\hat{s} = s\,\tau_{\mathrm{b}}N^{1/\mu}$, one obtains

$$\mathcal{L}\Pi_N\left(s = \frac{\hat{s}}{\tau_{\mathrm{b}}N^{1/\mu}}\right) \underset{s\to 0}{=} \left(1 - \Gamma(1-\mu)\,\frac{\hat{s}^\mu}{N} + \frac{O(\hat{s})}{N^{1/\mu}}\right)^N. \tag{4.16}$$

[4] Note that in eq. (4.13) s is conjugate to a time variable, T_N or τ, so that it has the dimension of the inverse of time, whereas in eqs. (4.8) and (4.10) the conjugate variables ξ and u are both dimensionless.

Taking the limit $N \to \infty$ and $s \to 0$, with $\hat{s}$ fixed, gives

$$\begin{aligned}
\mathcal{L}\Pi_N\left(s = \frac{\hat{s}}{\tau_b N^{1/\mu}}\right) &\underset{s\to 0}{=} \exp\left[N \ln\left(1 - \Gamma(1-\mu)\,\frac{\hat{s}^\mu}{N} + \frac{O(\hat{s})}{N^{1/\mu}}\right)\right] \\
&\underset{N\to\infty}{\longrightarrow} \exp\left(-\Gamma(1-\mu)\hat{s}^\mu + \frac{O(\hat{s})}{N^{1/\mu-1}}\right) \\
&\underset{N\to\infty}{\longrightarrow} \exp\left[-\Gamma(1-\mu)\hat{s}^\mu\right], \qquad (4.17)
\end{aligned}$$

since $\mu < 1$. Using the definition of $\mathcal{L}\Pi_N(s) = \int e^{-sT_N}\,\Pi_N(T_N)\,dT_N$, the change of variable $\xi = T_N/\tau_b N^{1/\mu}$ and the relation $P(\xi)\,d\xi = \Pi_N(T_N)\,dT_N$, the above calculation directly shows that $\mathcal{L}L_\mu(u)$ given by eq. (4.10) is indeed the Laplace transform of the distribution of ξ at large N.

4.2.3 A few mathematical results

We gather in this subsection a few useful mathematical results which are referred to in this chapter. This part can be skipped at first reading.

A few properties of the Laplace transforms of functions with power-law tails

(i) Suppose first that $\mu > 1$ so that $\langle\tau\rangle$ is finite. For $s \to 0$, one can then write

$$\begin{aligned}
\mathcal{L}P(s) &= \int_0^\infty d\tau\; P(\tau) e^{-s\tau} \\
&\underset{s\to 0}{\simeq} \int_0^\infty d\tau\; P(\tau)(1 - s\tau) = 1 - s\langle\tau\rangle. \qquad (4.18)
\end{aligned}$$

We will come back to the higher-order terms of the small-s expansion of $\mathcal{L}P(s)$ (see eq. (4.23)).

(ii) If $\mu < 1$, the previous expression is no longer valid because $\langle\tau\rangle$ is infinite. We rewrite $e^{-s\tau}$ in the first line of eq. (4.18) as $1 + e^{-s\tau} - 1$, so that

$$\begin{aligned}
\mathcal{L}P(s) &= \int_0^\infty d\tau\; P(\tau)(1 + e^{-s\tau} - 1) \\
&= 1 + \int_0^\infty d\tau\; P(\tau)(e^{-s\tau} - 1). \qquad (4.19)
\end{aligned}$$

Let τ^* be the value of τ beyond which the asymptotic expression (4.1) is correct. The integral of the last line of eq. (4.19) from 0 to ∞ can be split into an integral from 0 to τ^* and an integral from τ^* to ∞. Since $\left|e^{-s\tau} - 1\right| < s\tau$, one has:

$$\begin{aligned}
\left|\int_0^{\tau^*} d\tau\; P(\tau)(e^{-s\tau} - 1)\right| &< s\int_0^{\tau^*} d\tau\; \tau\, P(\tau) < \tau^* s \int_0^{\tau^*} d\tau\, P(\tau) \\
&< \tau^* s \int_0^{\infty} d\tau\, P(\tau) < \tau^* s. \qquad (4.20)
\end{aligned}$$

Thus, when s tends to 0, more precisely when $s \ll 1/\tau^*$, the integral from 0 to τ^* is at most of order $O(\tau^* s)$. In the integral from τ^* to ∞, we replace $P(\tau)$ by its asymptotic form (4.1) and we perform integration by parts. This gives, putting $x = s\,\tau$:

$$\int_{\tau^*}^{\infty} \mathrm{d}\tau\, P(\tau)(\mathrm{e}^{-s\tau} - 1) = \mu(\tau_\mathrm{b} s)^\mu \int_{\tau^* s}^{\infty} \mathrm{d}x\, x^{-(1+\mu)}(\mathrm{e}^{-x} - 1)$$
$$= (\tau_\mathrm{b} s)^\mu (\mathrm{e}^{-\tau^* s} - 1)(\tau^* s)^{-\mu} - (\tau_\mathrm{b} s)^\mu \int_{s\tau^*}^{\infty} \mathrm{d}x\, x^{-\mu}\mathrm{e}^{-x}. \tag{4.21}$$

Combining the last line of eq. (4.21) with eq. (4.20), we obtain

$$\mathcal{L}P(s) \underset{s\to 0}{\simeq} 1 - \Gamma(1-\mu)\,(\tau_\mathrm{b} s)^\mu - A_0 \tau_\mathrm{b} s + \cdots \tag{4.22}$$

where A_0 is a constant depending on the detailed shape of $P(\tau)$. This is nothing but eq. (4.14).

If we subtract from $P(\tau)$ its asymptotic behaviour (4.1), we are left with a new function $\widetilde{P}(\tau)$ which decays faster than $\tau^{-(1+\mu)}$ at large τ. If it decays faster than τ^{-2}, the integral $\int_0^\infty \mathrm{d}\tau\, \tau\, \widetilde{P}(\tau)$ converges and a calculation similar to that of eq. (4.18) gives a term of order $O(s\,\tau^*)$ when $s \to 0$. Combined with similar contributions of the same order from eq. (4.20) and eq. (4.21), this gives the last term of the right-hand side of eq. (4.22).

(iii) If μ had been larger than one, but different from any integer[5], the small-s expansion of $\mathcal{L}P(s)$ would have taken the following form:

$$\mathcal{L}P(s) = 1 - M_1 s + \frac{M_2}{2!} s^2 + \cdots + (-1)^n \frac{M_n}{n!} s^n - C_\mu s^\mu - \cdots \tag{4.23}$$

where n is the integer value of μ, and the M_i are the moments of $P(\tau)$ (for example, $M_1 = \langle \tau \rangle$ is the mean value of τ). In other words, the small-s expansion is regular up to its n^{th} term, until the power-law singularity is met. Conversely, the knowledge of $\mathcal{L}P(s)$ for small s allows one to extract the power-law behaviour of $P(\tau)$ for large τ.

(iv) Actually, one should also note that eq. (4.14) can be extended to the case where $P(\tau)$ is not a normalizable probability density and varies as $C\tau^{-(1-\nu)}$ at large τ with $\nu > 0$. Such a case was encountered in Chapter 3, Section 3.4.2: the probability of an atom being present at the starting point of a three-dimensional random walk decays as $\tau^{-1/2}$, corresponding to $\nu = 1/2$. In this case, calculations similar to the previous ones show that the leading term of $\mathcal{L}P(s)$ for small s reads:

$$\mathcal{L}P(s) = \Gamma(\nu) C s^{-\nu} + A + \cdots \tag{4.24}$$

where A is a constant, again depending on the detailed shape of $P(\tau)$.

A few properties of Lévy distributions

We now list without proofs a few important properties of $L_\mu(\xi)$, defined in eqs. (4.8) and (4.10), remembering that we are restricting ourselves to the case of positive random variables.

5 Again, if μ is an integer, logarithmic corrections appear, see Appendix C.

(i) For $\mu = 2$, $L_\mu(\xi)$ reduces to the usual Gaussian distribution $G(\xi) = (2\pi)^{-1/2} \exp(-\xi^2/2)$.

(ii) For $0 < \mu < 2$ and $\xi \to \infty$, $L_\mu(\xi)$ decays as a power law with the same exponent as $P(\tau)$:

$$L_\mu(\xi) \underset{\xi \to +\infty}{\simeq} \frac{\mu}{\xi^{1+\mu}} + O\left(\frac{1}{\xi^{1+2\mu}}\right). \tag{4.25}$$

(iii) For $\mu < 1$, $L_\mu(\xi)$ is obviously 0 for $\xi < 0$ and has an essential singularity for $\xi \to 0$:

$$L_\mu(\xi) \underset{\xi \to 0}{\simeq} A\, \xi^{\frac{\mu-2}{2(1-\mu)}} \exp\left(-B\, \xi^{\frac{\mu}{\mu-1}}\right) \tag{4.26}$$

where A and B are prefactors.

(iv) For $\mu = 1/2$, an explicit expression can be given for all ξ:

$$L_{1/2}(\xi) = Y(\xi) \frac{1}{2\xi^{3/2}} \exp\left(-\frac{\pi}{4\xi}\right) \tag{4.27}$$

where $Y(\xi)$ is the Heaviside function. The variations of $L_{1/2}(\xi)$ with ξ are represented in figure 4.1. All functions $L_\mu(\xi)$ with $\mu < 1$ have qualitatively similar variations. Note that the maximum of $L_{1/2}(\xi)$ is reached for $\xi = \pi/6$, which clearly shows that the dimensionless random variable ξ is of the order of one.

(v) For $1 < \mu < 2$, $L_\mu(\xi)$ describes the fluctuations of T_N around the mean value $N\langle\tau\rangle$, and thus extends from $-\infty$ to $+\infty$. The decay of $L_\mu(\xi)$ for $\xi \to -\infty$ is however much faster than the power law (4.25), and is given by:

$$L_\mu(\xi) \underset{\xi \to -\infty}{\simeq} C\, \xi^{\frac{\mu-2}{2(1-\mu)}} \exp\left(-D\, |\xi|^{\frac{\mu}{\mu-1}}\right) \tag{4.28}$$

where C and D are prefactors.

(vi) Only the moments of order $q < \mu$ of $L_\mu(\xi)$ exist. For $\mu < 1$, an explicit calculation leads to:

$$\langle \xi^q \rangle \equiv \int_0^\infty \mathrm{d}\xi\, \xi^q L_\mu(\xi) = b_\mu^{q/\mu} \frac{\Gamma(-q/\mu)}{\mu\Gamma(-q)} \tag{4.29}$$

where $b_\mu = \Gamma(1-\mu)$ (see eq. (4.10)).

4.3 Qualitative discussion of some properties of Lévy sums

4.3.1 Dependence of a Lévy sum on the number of terms for $\mu < 1$

One of the most important results of the generalized CLT is that a Lévy sum T_N scales as $N^{1/\mu}$ when $\mu < 1$ (see eq. (4.9)). For example, for $\mu = 1/2$, T_N scales as N^2; for $\mu = 1/4$, as N^4. The smaller μ, the greater the exponent of the power-law dependence of T_N on N. Such behaviour is quite different from that of usual random variables τ which have a finite mean value $\langle\tau\rangle$ and for which T_N scales as $N\langle\tau\rangle$ (usual law of large numbers).

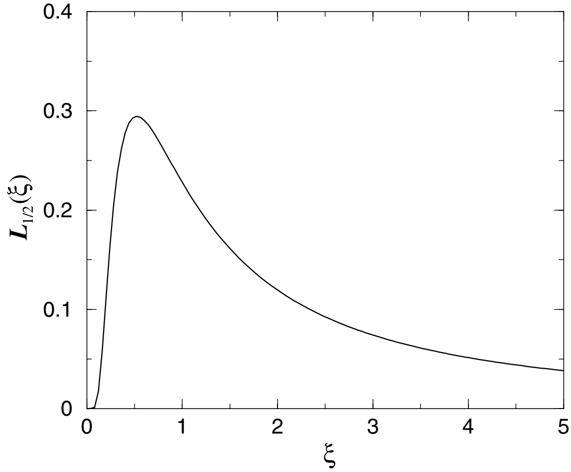

Fig. 4.1. Function $L_{1/2}(\xi)$. After a very slow increase near $\xi = 0$, $L_{1/2}(\xi)$ reaches a maximum for $\xi = \pi/6$ and then decreases as a power law at large ξ (as $1/(2\xi^{3/2})$).

Such a result is linked to the fact that the probability of having a very large value of τ in a drawing of the random variable is not negligible when $P(\tau)$ decreases slowly at large τ. When one increases the number N of trials, larger and larger values of τ can be obtained, and this explains why the sum T_N can grow faster than N.

4.3.2 Hierarchical structure in a Lévy sum

We now want to address the following questions. Suppose that one orders the sequence $\{\tau_1, \tau_2, \ldots, \tau_N\}$ of the various terms of a Lévy sum T_N from the largest one to the smallest. Let $\tau^{(1)}$ be the first one (the largest), $\tau^{(2)}$ the next one, ..., $\tau^{(n)}$ the n^{th} one. What are the orders of magnitude of these various terms? How do they scale with N and with n? In other words, is there a hierarchy between these terms? Is $\tau^{(n)}$ much larger than $\tau^{(n+1)}$?

To answer these questions, we first determine the most probable value of $\tau^{(n)}$. Let $\Pi(\tau^{(n)})\mathrm{d}\tau$ be the probability of finding the n^{th} term between $\tau^{(n)}$ and $\tau^{(n)} + \mathrm{d}\tau$. We have (see also Section 2.1.1 in [Gum58]):

$$\Pi(\tau^{(n)}) = N \binom{N-1}{n-1} P(\tau^{(n)}) \left[\int_{\tau^{(n)}}^{\infty} \mathrm{d}\tau\, P(\tau)\right]^{n-1} \left[1 - \int_{\tau^{(n)}}^{\infty} \mathrm{d}\tau\, P(\tau)\right]^{N-n}. \tag{4.30}$$

The first term, N, corresponds to the N possible positions of $\tau^{(n)}$ in the sequence

$\tau_1, \tau_2, \ldots, \tau_N$. The second term, $\binom{N-1}{n-1}$, counts the different possible ways of obtaining $n-1$ drawings larger than $\tau^{(n)}$ and $N-n$ smaller than $\tau^{(n)}$. Finally, the last three terms are the probabilities of drawing values of τ equal to, larger or smaller than $\tau^{(n)}$, respectively, raised to the appropriate power. Using eq. (4.1), one gets

$$\int_{\tau^{(n)}}^{\infty} \mathrm{d}\tau \; P(\tau) = \left(\frac{\tau_{\mathrm{b}}}{\tau^{(n)}}\right)^{\mu} \tag{4.31}$$

and a simple calculation shows that the most probable value of $\tau^{(n)}$, which maximizes eq. (4.30), is given by:

$$\begin{aligned} \tau^{(n)} &= \tau_{\mathrm{b}} \left[\frac{1+\mu N}{1+\mu n}\right]^{1/\mu} \\ &\simeq \tau_{\mathrm{b}} \left(\frac{N}{n}\right)^{1/\mu} \qquad \text{if} \qquad N, n \gg 1/\mu. \end{aligned} \tag{4.32}$$

A first important result expressed by eq. (4.32) is that the largest term of a Lévy sum, $\tau^{(1)}$, scales with N as $\tau_{\mathrm{b}} N^{1/\mu}$. This result is valid for any value of $\mu > 0$, in the limit $N \to \infty$. Interestingly, for $\mu < 1$, one has $T_N \simeq \tau_{\mathrm{b}} N^{1/\mu}$ according to eq. (4.9) so that the largest term $\tau^{(1)}$ is of the order of the sum itself. A single term of the Lévy sum can be of the order of the total sum! This is the most important qualitative property of the Lévy sums for $\mu < 1$: a significant fraction of the total 'time' T_N is spent in the 'deepest trap'. This is precisely the situation encountered in the Monte Carlo simulations described in Chapter 2.

The n-dependence of $\tau^{(n)}$ is also very interesting. As soon as n becomes larger than $1/\mu$, $\tau^{(n)}$ scales with n as $n^{-1/\mu}$. For example, for $\mu = 1/2$, $\tau^{(10)}$ is $2^2 = 4$ times larger than $\tau^{(20)}$, $3^2 = 9$ times larger than $\tau^{(30)}$, and so on. In other words, there is a strong hierarchy between the various terms of a Lévy sum with $\mu < 1$. Such a sum is 'dominated' by a very small number of terms. If one plots $\ln \tau^{(n)}$ versus $\ln n$, one expects, according to eq. (4.32), to get a straight line with a slope $-1/\mu$. Conversely, when one analyses a set of independent random numbers, it may be useful to order them and to plot $\ln \tau^{(n)}$ versus $\ln n$. If one gets a straight line with a slope $-1/\mu$, this is a good indication that the corresponding random variable is distributed according to a power-law distribution such as eq. (4.1)[6].

It is interesting to compare the previous results, typical of Lévy statistics, with those corresponding to usual Gaussian statistics where $P(\tau)$ is a 'narrow' distribution for which the CLT is applicable. Take, for example, the exponential

[6] A more precise 'maximum likelihood' procedure to estimate the exponent μ is known as the Hill estimator, see [Hil75].

distribution

$$P(\tau) = \frac{1}{\tau_b} e^{-\tau/\tau_b} \tag{4.33}$$

leading to a simple analytical expression for the value of $\tau^{(n)}$ which maximizes eq. (4.30):

$$\tau^{(n)} = \tau_b \ln\left(\frac{N}{n}\right). \tag{4.34}$$

Instead of power-law variations with N and n, we obtain now logarithmic variations which are extremely slow. In other words, there is now no hierarchy between the various terms of the sum which are all of the same order. An increase of the size N of the statistical sample leads only to a modest increase of the typical size of the largest term $\tau^{(1)}$.

4.3.3 Large fluctuations

For usual statistics obeying the standard CLT (finite $\langle\tau\rangle$ and $\langle\tau^2\rangle$), the sample to sample fluctuations of the sum T_N vanish when the size of the sample, i.e. the number of terms N, increases. More precisely, let us consider the relative fluctuations $\sigma_r(N)$ of the average value for a sample of size N defined by[7]:

$$\sigma_r(N) = \frac{\langle|(T_N/N) - \langle\tau\rangle|\rangle}{\langle\tau\rangle} = \frac{\langle|T_N - N\langle\tau\rangle|\rangle}{N\langle\tau\rangle}. \tag{4.35}$$

According to eqs. (4.4) and (4.5), the variable $(T_N - N\langle\tau\rangle)/(\sigma\sqrt{N})$ is of order one when $N \gg 1$ and a simple calculation leads to

$$\langle\tau^2\rangle < \infty: \quad \sigma_r(N) \simeq \frac{\sigma}{\langle\tau\rangle\sqrt{N}}. \tag{4.36}$$

These fluctuations tend to zero when N tends to infinity. This guarantees an asymptotically perfect repeatability of average values in the limit of large samples. In other words, average values can be accurately predicted for large samples, even if individual values fluctuate a lot. This is the origin of the traditional success of statistical methods in both natural and social sciences.

For Lévy statistics, the situation can be radically different. For $1 < \mu < 2$, a simple calculation using eq. (4.6) leads to

$$1 < \mu < 2: \quad \sigma_r(N) \simeq \tau_b/(\langle\tau\rangle N^{1-1/\mu}). \tag{4.37}$$

The relative fluctuations of the average value again vanish[8] at large N, although

[7] We take an absolute value instead of a root mean square to avoid divergencies for the case $1 < \mu < 2$ considered below.

[8] But the fluctuations of the second moment would not vanish.

more slowly than $1/\sqrt{N}$. But for $0 < \mu < 1$, this is no longer true. Since we can no longer define the relative fluctuations $\sigma_r(N)$ by (4.35) ($\langle\tau\rangle$ is infinite), we use the following argument: as the largest term $\tau^{(1)}$ is of the order of the sum T_N, the sum T_N fluctuates as much as a single term. Therefore the relative fluctuations from sample to sample are the same as the fluctuations from term to term, i.e. they are of order one whatever the size of the sample:

$$\mu < 1: \quad \sigma_r(N) \simeq 1. \tag{4.38}$$

As a consequence, the value of the sum T_N is not repeatable from one sample to another sample. The accuracy of the statistical prediction is not improved by increasing the sample size.

It thus appears that Lévy statistics lead, when $\mu < 1$, to a behaviour which is radically different from that deduced from the usual CLT [Man82, Man96]. The usual CLT describes how the fluctuations vanish at large N, whereas the generalized CLT (for $\mu < 1$) shows that the fluctuations continue to play an essential role however large N may be.

Repeatability is unavoidably lost when $\mu < 1$. However, the generalized CLT still allows *some predictability*. It predicts the typical, i.e. most probable, values for the sums T_N. Such an order of magnitude prediction is the best that statistical tools can offer when $\mu < 1$.

It is worth pointing out that the presence in a physical phenomenon of a sum T_N undergoing large fluctuations does not necessarily imply that the phenomenon is on the whole unrepeatable. Other quantities related to, but different from, T_N can still be accurately predicted even when $\mu < 1$. The physically relevant quantities calculated in the following chapters are of this kind.

4.3.4 Illustration with numerical simulations

All the spectacular features of Lévy statistics analysed in the previous section clearly appear in numerical simulations. These numerical simulations are performed in the following way. One makes successive drawings $\tau_1, \tau_2, \ldots, \tau_N, \ldots$ of the random variable τ distributed according to eq. (4.1), and one plots $T_N = \sum_{i=1}^{N} \tau_i$ versus N, for different values of μ. These sequences are generated using the *same*[9] sequence $x_1, x_2, \ldots, x_N, \ldots$ of random numbers uniformly distributed between 0 and 1, and then defining:

$$\tau_i = \tau_b x_i^{-1/\mu}. \tag{4.39}$$

[9] The use of the same sequence of x_i enables one to see the effects of different μ values not blurred by the statistical fluctuations.

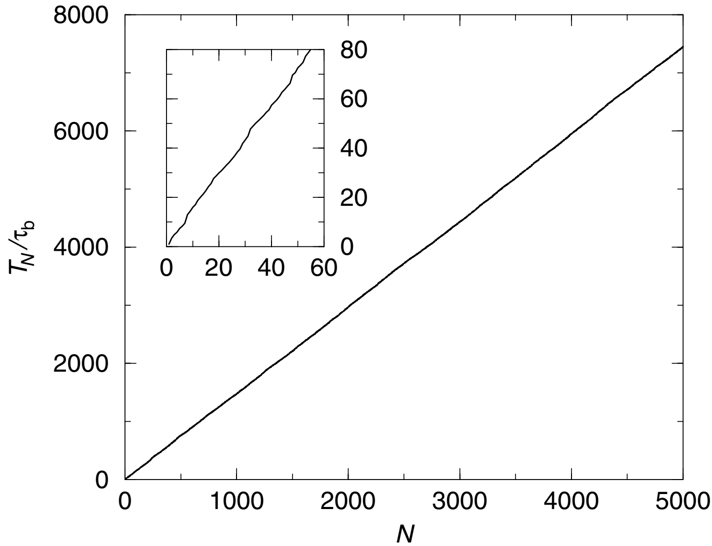

Fig. 4.2. Plot of T_N (in units of τ_b) versus N for $\mu = 3$. The inset shows a zoom of a small portion of the plot.

One can check that this transformation produces τ_i values that are distributed according to eq. (4.1).

Figure 4.2 shows T_N (in units of τ_b) versus N for $\mu = 3$. In this case, $\langle\tau\rangle$ is finite and equal to $\mu\,\tau_\mathrm{b}/(\mu - 1) = 1.5\,\tau_\mathrm{b}$ (see eq. (3.35)), and one obtains a plot which looks like a straight line with a slope $\mu/(\mu - 1) = 1.5$. In fact, there are $N = 5000$ vertical steps in such a plot, but each individual step is so small that it cannot be distinguished in the full scale figure. Zooming in on a small portion of the figure reveals these individual steps which appear to be all of the same order (see inset of figure 4.2).

For $\mu < 1$, when $\langle\tau\rangle$ is infinite, the plot has a radically different shape. It looks like a 'devil's staircase' where a small number of individual large steps are clearly visible and are of the order of the total sum itself (see, for example, figure 4.3 corresponding to $\mu = 1/2$). When μ is still smaller, for example when $\mu = 0.1$, one nearly sees only a single huge step (see figure 4.4). Between two large steps, T_N remains nearly constant. This is due to the strong hierarchy between the individual steps (see eq. (4.32)). A few of them are so large that the others can hardly be distinguished. Note the difference of the vertical scales from figure 4.2 to figure 4.3 and figure 4.4, which reflects the $N^{1/\mu}$ dependence of T_N when $\mu < 1$. Zooming in on a small portion of figure 4.3 and figure 4.4 reveals a structure which

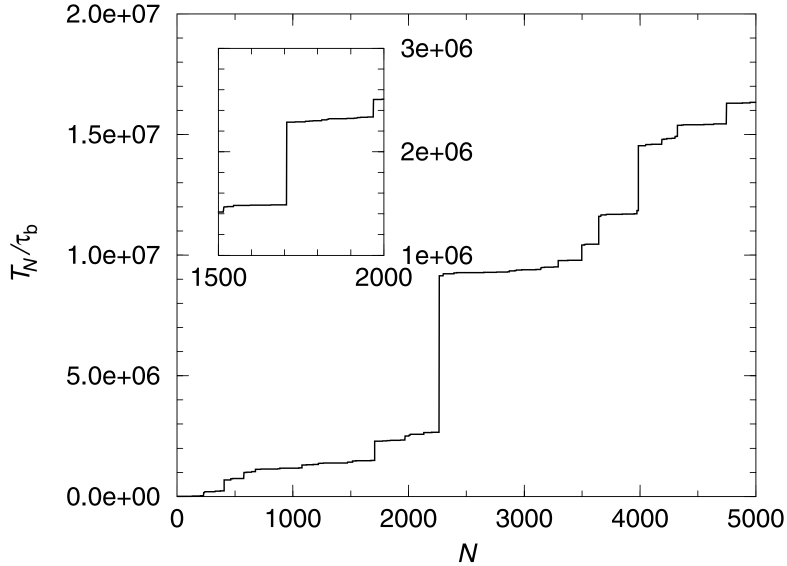

Fig. 4.3. As figure 4.2, but with $\mu = 1/2$. Note the difference of vertical scale. Contrary to figure 4.2, a few large steps are clearly visible and are of the order of the total sum. The same general behaviour appears in the zoom shown in the inset.

has the same shape as the full scale figures (see the insets): one still gets a kind of 'devil's staircase' dominated by a small number of large steps. In other words, the behaviour of T_N versus N is self-similar at all scales.

The hierarchical structure of the various terms of a Lévy sum also appears in rank ordered histograms where one plots $\ln \tau^{(n)}$ versus $\ln n$. Figure 4.5 shows such plots for $\mu = 3$ and $\mu = 1/2$. As expected from the calculations of Section 4.3.2, one obtains a decrease which is well represented by a straight line with a slope equal to $-1/\mu$. These straight lines are shown as interrupted lines in the figure. Note that for $\mu = 1/2$ there are about six orders of magnitude between the largest term and the smallest term of the sequence.

4.4 Sprinkling distribution

4.4.1 Definition. Laplace transform

In this section, we introduce a probability distribution which will be useful for the calculations presented in Chapters 5 and 6. Suppose that one makes successive drawings $\tau_1, \tau_2, \ldots, \tau_n, \ldots$ of the random variable τ distributed according to eq. (4.1), and let us define a random sequence of events $M_1, M_2, \ldots, M_n, \ldots$

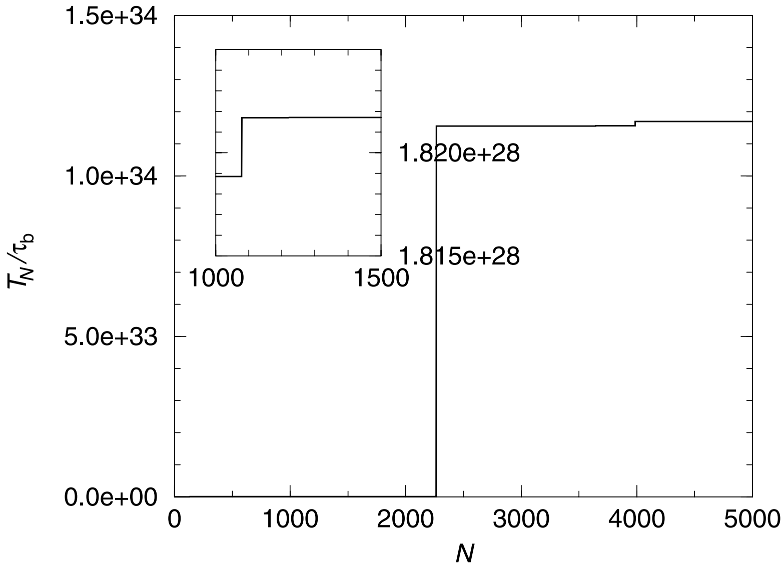

Fig. 4.4. As figure 4.2, but with $\mu = 0.1$. The hierarchical structure in the Lévy sum is still more pronounced than in figure 4.3 and the largest term is huge (note the new vertical scale) and dominates all the others. Here also, the same behaviour appears at all scales (see the inset).

occurring at times $t_1, t_2, \ldots, t_n, \ldots$ such that

$$t_1 = \tau_1, \quad t_2 = \tau_1 + \tau_2, \ldots, \quad t_n = t_{n-1} + \tau_n, \ldots. \tag{4.40}$$

In other words, we introduce a random set of events such that the time intervals between two *successive* events is distributed according to $P(\tau)$. This is illustrated in Fig. 4.6. Averaging over several different realizations of such a random sequence, one can then ask the following question: what is the probability density $S(t)$ of finding an event at time t, disregarding the number of previous events? We shall call such a distribution the 'sprinkling distribution' associated with $P(t)$. It represents the mean density at time t of the random sequence of events M_1, M_2, $\ldots$, $M_n, \ldots$ introduced above.

It is easy to find an equation satisfied by $S(t)$. Either the event observed at time t is the first one which appears, with probability $P(t)$; or an arbitrary number of events have already occurred before this event, the last one happening at $t_l < t$. Hence, one has:

$$S(t) = P(t) + \int_0^t \mathrm{d}t_l \, P(t - t_l) S(t_l). \tag{4.41}$$

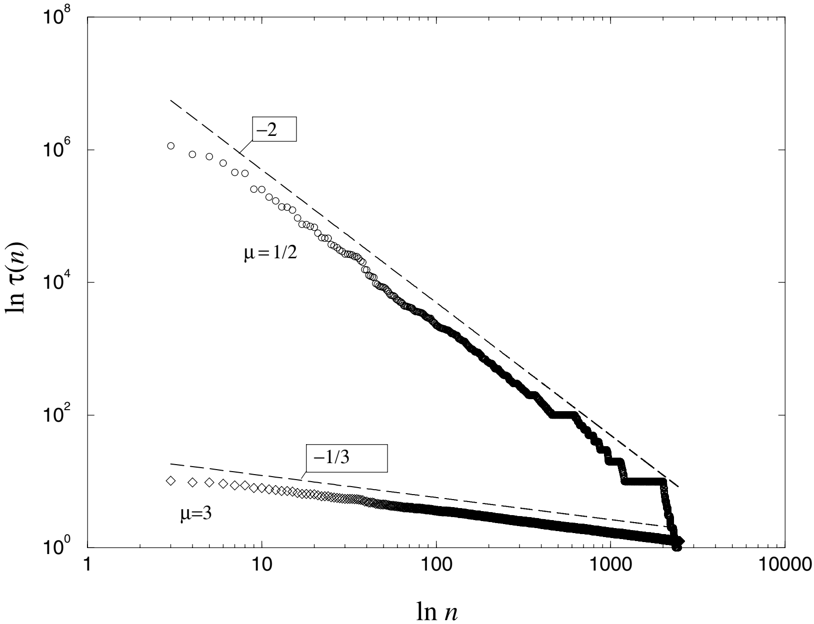

Fig. 4.5. Rank ordered histograms giving $\ln \tau^{(n)}$ versus $\ln n$ for two different values of μ: $\mu = 3$ and $\mu = 1/2$. The interrupted straight lines give the theoretically predicted behaviour of a linear decrease with a slope equal to $-1/\mu$.

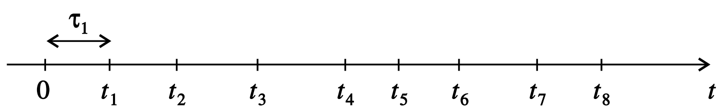

Fig. 4.6. Random set of events M_i occurring at times t_i, with a distribution $P(\tau)$ for the time intervals between two successive events.

This equation is readily solved using Laplace transforms, which converts the convolution into a simple product and one gets:

$$\mathcal{L}S(s) = \frac{\mathcal{L}P(s)}{1 - \mathcal{L}P(s)}. \tag{4.42}$$

4.4.2 Examples taken from other fields

In this book, the events that will be considered in the next sections, and which will be characterized by the sprinkling distribution $S(t)$, are the successive entries of the atom in the trapping zone $p \leq p_{\text{trap}}$ during its random walk in momentum space.

In the general theory of stochastic processes, the sprinkling distribution is known as the density of a renewal process, or the 'renewal density'[10]. Schematically, a renewal process is a statistical process in which a device, say a lightbulb, is installed at time $t = 0$, until it fails and is replaced at time $t = \tau_1$ (random variable), until the new device fails and is replaced at time $t = \tau_1 + \tau_2, \dots$. The renewal density indicates (statistically) when the devices must be replaced.

Renewal processes are ubiquitous in quantum optics although they are not usually named as such. Consider, for example, the sequence of fluorescence photons emitted by a single atom excited by a resonant laser field (single atom resonance fluorescence). A very important quantity characterizing such a sequence is the so-called second-order correlation function

$$G_2(t) = \langle E^-(0)E^-(t)E^+(t)E^+(0)\rangle \tag{4.43}$$

where E^- and E^+ are the negative and positive frequency parts, respectively, of the electric field operator. It can be shown that $G_2(t)$ is the probability of having a spontaneous emission at time $t = 0$ and another one at time t, not necessarily the next one. It is clear that $G_2(t)$ is a renewal density. More recently, attention has also been paid to the waiting time distribution $W(\tau)$ (or 'delay function', see Section 2.3.3), giving the distribution of the time intervals between two *successive* spontaneous emissions. The two distributions $G_2(t)$ and $W(\tau)$ are related by an equation

$$G_2(t) = W(t) + \int_0^t \mathrm{d}t_l\, W(t - t_l)G_2(t_l), \tag{4.44}$$

which is identical to the renewal equation (4.41) giving the sprinkling distribution, with the correspondence $W \to P$ and $G_2 \to S$ (see, for example, eq. (6.19) in [Rey83], eq. (4.13) in [RDC88] or eq. (45) in [PlK98]).

4.4.3 Asymptotic behaviour. Broad versus narrow distributions

We now investigate the long time behaviour of $S(t)$. Suppose first that $\mu > 1$ so that $\langle\tau\rangle$ is finite. Using eq. (4.18), which states that the small-s expansion of $\mathcal{L}P(s)$ is $1 - \langle\tau\rangle s + \cdots$, one gets:

$$\mathcal{L}S(s) \underset{s\to 0}{\simeq} \frac{1}{\langle\tau\rangle}\frac{1}{s} - 1 \cdots . \tag{4.45}$$

This shows that, for large times, $S(t)$ is constant, equal to $1/\langle\tau\rangle$. We thus find an *a priori* obvious result. For large times, the probability of finding a particular event between t and $t + \mathrm{d}t$ is a constant equal to the inverse of the average time interval $\langle\tau\rangle$ between two successive events. In other words, the set of events M_1, $M_2, \dots, M_n, \dots$ has a constant density equal to $1/\langle\tau\rangle$.

[10] This connection was made in [BaB00], see Section 10.2.1.

Equation (4.45) is in fact valid for any probability distribution $P(\tau)$ with a finite average value. This is in particular the case of the sprinkling distribution $G_2(t)$ associated with single atom resonance fluorescence, since $W(t)$, which plays the role of $P(t)$, is a sum of exponentials with a finite mean value. Another simple example is provided by the Poisson process which enables us to check eq. (4.45). In this case, the waiting time distribution is $P(\tau) = \Gamma e^{-\Gamma\tau}$ and one expects the rate $S(t)$ of occurrence of the events to be constant, for any time including small times. Using eq. (4.42) and $\mathcal{L}P(s) = \Gamma/(\Gamma + s)$, one obtains $\mathcal{L}S(s) = \Gamma/s$. This agrees with the first term of eq. (4.45) but, in this case, the relation is exact, for any s. We get $S(t) = \Gamma$ for all t, as expected for a Poisson process.

Such a result is no longer valid when $\mu < 1$. We must now use eq. (4.14) which, inserted into eq. (4.42), gives[11]

$$\mathcal{L}S(s) \underset{s\to 0}{\simeq} \frac{1}{\Gamma(1-\mu)} (\tau_{\mathrm{b}} s)^{-\mu} + \text{subleading terms}. \tag{4.46}$$

Using eq. (4.24) and the identity

$$\Gamma(\mu)\Gamma(1-\mu) = \frac{\pi}{\sin(\pi\mu)}, \tag{4.47}$$

one finally gets[12]

$$S(t) \underset{t\to\infty}{\simeq} \frac{\sin(\pi\mu)}{\pi} \frac{1}{\tau_{\mathrm{b}}} \left(\frac{\tau_{\mathrm{b}}}{t}\right)^{1-\mu} + O\left[(\tau_{\mathrm{b}}/t)^{2-2\mu}\right]. \tag{4.48}$$

Note that $S(t)$ has the dimension of the inverse of a time, but goes to zero when $t \to \infty$. This is related to the fact that, as the time t increases, the probability of drawing a large value of τ, of the order of t itself, remains constant, so that the mean density of events decreases. In such a process, *the rate of events decreases at long times due to a purely statistical property* ($\langle\tau\rangle = \infty$) *while the distribution* $P(\tau)$ *of the increments* τ_i *is perfectly stationary*. The identification of this unusual feature in laser cooling is one of the most salient results of the presented statistical approach.

This also means that the observation of $S(t)$ allows one to infer the starting 'date' ($t = 0$) of the process – which would of course be impossible to do for $\mu > 1$. In other words, time translation invariance is broken for $\mu < 1$ and the process 'ages'. Such a scenario was discussed in the context of glassy dynamics in [Bou92, BoD95, BCK97]. The sprinkling distribution $S(t)$ associated with a broad distribution $P(t)$ therefore exhibits interesting new features compared with the usual case where $P(t)$ has a finite mean value.

[11] If $\mu < 1/2$, the subleading terms of eq. (4.46) are constant terms plus terms in $(\tau_{\mathrm{b}} s)^{1-2\mu}$. If $\mu > 1/2$, these corrections are in $(\tau_{\mathrm{b}} s)^{1-2\mu}$.

[12] If $\mu = 1$, logarithmic terms appear.