Experimental Investigation of Nonergodic Effects in Subrecoil Laser Cooling

B. Saubaméa, M. Leduc, and C. Cohen-Tannoudji

Collège de France and Laboratoire Kastler Brossel, Ecole Normale Supérieure, 24 rue Lhomond, F-75231 Paris Cedex 05, France

(Received 25 June 1999)

We present the first detailed investigation of the line shape of the momentum distribution of atoms cooled far below the recoil limit. This distribution is deduced from a direct measurement of the atomic spatial correlation function of metastable helium atoms cooled by velocity-selective coherent population trapping. The measured line shape is then compared to the prediction of an analytical model of this cooling based on Lévy statistics. A very good agreement is found between experiment and theory and fundamental features such as self-similarity and nonergodicity are identified.

PACS numbers: 32.80.Pj, 42.50.Vk

Laser cooling techniques are commonly used to lower the temperature of neutral atoms and ions, making their momentum distribution as narrow as possible. Most often, the details of the shape of this distribution do not matter as much as its width, which gives the momentum spread of the atoms. Nonetheless, there is a lot to learn from this shape as it reflects the dynamics and essential features of the evolution of the atoms during the cooling. Using a matter wave interferometric method described in a previous paper [1], we present in this Letter a detailed experimental investigation of the momentum distribution of metastable Helium atoms, cooled well below the recoil limit in the nanokelvin range. As in other physical problems like micelles or spin glasses as well as in finances or geology, "broad distributions" naturally occur in subrecoil cooling. By broad, we mean a distribution with power-law tails decreasing too slowly to have its mean and/or its variance defined. Lévy statistics are known to be the relevant tool to deal with these distributions [2,3]. Here, we compare the measured momentum distribution of the atoms with a theoretical model based on Lévy statistics and we show how our measurements reveal some important features of the cooling process such as self-similarity and nonergodicity.

In most cooling schemes, the ensemble of atoms reaches a dynamical equilibrium resulting from a competition between the cooling process, a friction force which damps the atomic momentum, and a heating process, the momentum diffusion due to spontaneous emission. The motion of an atom interacting with laser light is thus very similar to a classical Brownian motion, and the resulting momentum distribution is found to be Gaussian, as expected for a normal random walk.

The situation is quite different for subrecoil cooling. In this case, the cooling mechanism is not based on a friction force but on an inhomogeneous random walk. More precisely, the steps of the random walk in momentum space have still the size $\hbar k$, the photon momentum, but the jump rate $R(p)$ is now momentum dependent and vanishes when the atomic momentum p tends to zero. If, after a spontaneous emission, the atom reaches a state with a very low momentum $p \simeq 0$, the mean time $\tau(p) = [R(p)]^{-1}$ it remains there before undergoing the next jump can be very long. The random walk in momentum space therefore slows down as $|p| \to 0$, leading to an accumulation of atoms about $p = 0$, with a momentum distribution $\mathcal{P}(p)$ whose width δp can become much smaller than $\hbar k$ (subrecoil regime). Up to now, the vanishing of $R(p)$ for $p = 0$ has been implemented by using two different methods: velocity selective coherent population trapping (VSCPT), using a velocity selective destructive quantum interference between different transition amplitudes leading to the same excited state [4] and velocity selective Raman cooling [5].

One of the most important features of subrecoil cooling is the absence of steady state. Even for an arbitrarily long interaction time θ, there are always atoms with a small enough p that their characteristic evolution time $\tau(p)$ is longer than θ and even approaches infinity as $|p| \to 0$. A first consequence of this situation is that $\mathcal{P}(p)$ never stops to evolve when θ increases, with a width δp decreasing as $1/\sqrt{\theta}$ (for VSCPT). This has been experimentally checked with a great accuracy [1]. Nonergodic features also appear in the shape of $\mathcal{P}(p)$ and their investigation is the subject of this Letter. Before describing our experimental results, we briefly sketch the main steps in the derivation of $\mathcal{P}(p)$. More details can be found in [6,7].

We first introduce the probability distribution $P(\tau_S \mid p)$ of the time τ_S spent in a state p before the next spontaneous jump. This distribution is nothing but the well-known "delay function" or "waiting time distribution" [8,9]. In the long time regime, $P(\tau_S \mid p)$ can be approximated by [10]:

$$P(\tau_S \mid p) = R(p) e^{-R(p)\tau_S}, \tag{1}$$

where the jump rate is given by

$$R(p) = \begin{cases} \frac{1}{\tau_0} \left(\frac{p}{p_0}\right)^2 & \text{if } |p| < p_0 , \\ \frac{1}{\tau_0} & \text{if } |p| \geq p_0 . \end{cases} \tag{2}$$

Here we neglect the slow decrease of $R(p)$ at large p due to the Doppler effect. This simplification is valid if the

 0031-9007/99/83(19)/3796(4)$15.00

interaction time is short enough so that the momentum diffusion during θ does not allow p to reach large values.

We now introduce a value p_{trap} such that all atoms with $|p| \leq p_{\text{trap}}$ have a small enough jump rate to be considered trapped in momentum space [11]. An atom initially trapped will eventually emit a spontaneous photon which, if we assume $p_{\text{trap}} \ll \hbar k$, will certainly kick it out of the trap. Outside the trap, it will undergo a series of spontaneous emissions, which will eventually bring it back to the trap, and so on. The time evolution of the atom therefore consists of a succession of trapping and diffusing periods. We call τ_i the duration of the ith trapping period, and $\hat{\tau}_i$ the duration of the ith diffusing period ($\hat{\tau}_i$ is a "first return time" in the trap).

It then can be shown from Eqs. (1) and (2), that the two random variables τ and $\hat{\tau}$ are distributed as $P(x) \sim 1/x^{3/2}$ for large x. We recognize here a typical power law for the tails of a "broad distribution" with no mean value or variance. Therefore, if we are interested, for example, in the total time an atom has spent in the trap, namely $T = \sum_i \tau_i$, we can no longer use the central limit theorem. Instead we must resort to the generalized central limit theorem which states that T is distributed not as a Gauss law but as a Lévy law [3]. In fact, the broad tails of $P(\tau)$ and $\hat{P}(\hat{\tau})$ are one of the key points of the cooling process since they reflect the fact that the evolution of the system is mainly dominated by the very long times.

Another important quantity for the calculation of $\mathcal{P}(p)$ is the "sprinkling distribution" $S(t)$: $S(t)dt$ is the probability that an atom enters the trap between t and $t + dt$, whatever the number of entries and exits it made before. This function can be calculated from $P(\tau)$ and $\hat{P}(\hat{\tau})$ and one finds $S(t) \propto t^{-1/2}$, which shows that the sprinkling of the trap becomes weaker and weaker as t increases, an indication of the "aging" of the system. Finally, one can say that an atom with $|p| < p_{\text{trap}}$ at time θ entered the trap for the last time at some time $t \leq \theta$ and did not undergo a spontaneous jump between t and θ. The final momentum distribution is therefore given by

$$\mathcal{P}(p) \propto \int_0^\theta S(t)\Pi(p)\Psi(p, \theta - t)\, dt\,, \qquad (3)$$

where $\Pi(p)$ is the probability that an atom entering the trap has a momentum p with $|p| \leq p_{\text{trap}}$, and $\Psi(p, \theta - t)$ is the probability that its next spontaneous jump would occur after a delay $\tau_S \geq \theta - t$. With the assumption $p_{\text{trap}} \ll \hbar k$, one can consider $\Pi(p)$ as uniform and equal to $1/2p_{\text{trap}}$. In addition, one has $\Psi(p, \theta - t) = \int_{\theta - t}^{\infty} P(\tau_S|p)\, d\tau_S = e^{-R(p)(\theta - t)}$. A simple calculation then gives

$$\mathcal{P}(p) \propto \mathcal{N}\, \frac{1}{2\pi p}\, e^{-p^2/p_\theta^2} \int_0^{p^2/p_\theta^2} x^{-1/2} e^x\, dx\,, \qquad (4)$$

where $\mathcal{N}$ is a normalization constant and the parameter p_θ is defined as

$$R(p_\theta)\theta = 1 \;\Rightarrow\; p_\theta = \frac{1}{4}\, \frac{\hbar\Gamma}{E_R}\, \frac{\tilde{\Omega}}{\tilde{\theta}^{1/2}} \times \hbar k\,. \qquad (5)$$

In this expression, $\hbar\Gamma$ is the natural width of the transition, $E_R = \hbar^2 k^2/2M$ the recoil energy, $\tilde{\Omega} = \Omega/\Gamma$ and $\tilde{\theta} = \theta\Gamma$ the dimensionless Rabi frequency and interaction time [12]. The expression above shows that $\mathcal{P}(p)$ can always be transformed into the same law $G(\eta)$, independent of p_θ, by a simple rescaling $\eta = p/p_\theta$. The momentum distribution thus displays a self-similar behavior: it remains time invariant in the scale parameters used.

It is easy to prove from Eq. (4) that the tails of $\mathcal{P}(p)$ decrease as $1/p^2$, and this can be given a simple physical explanation. For a momentum $|p| \gg p_\theta$, the mean time $\tau(p) = [R(p)]^{-1}$ spent in the state p is, according to (5), much shorter than θ. A trapping event with momentum $|p| \gg p_\theta$ can thus be sampled many times during the observation time θ, and we expect time and ensemble averages to be equivalent. The population of a state p is therefore proportional to the mean time spent in this state leading to $\mathcal{P}(p) \propto [R(p)]^{-1} \propto 1/p^2$. However, this behavior in $1/p^2$ cannot be extrapolated for $|p| < p_\theta$. First, it would lead to a nonnormalizable distribution. A second, more physical argument is that $|p| < p_\theta$ implies $\tau(p) > \theta$ so that the corresponding trapping event has a high probability not to be sampled many times during θ. Time and ensemble averages can thus no longer coincide and the deviations of $\mathcal{P}(p)$ from a $1/p^2$ law at small p reflect the resulting nonergodicity. However, in the presence of a weak loss process adding a nonzero jump rate R_0 varying slowly with p about $p = 0$, the distribution of the delay τ_S would be truncated beyond the value R_0^{-1}. In the regime $\theta \gg R_0^{-1}$, all trapping events could then be sampled many times during θ and the system would show ergodicity with a Lorentzian momentum distribution $\mathcal{P}(p) \propto [R(p)]^{-1} \propto 1/(R_0 + p^2/\tau_0 p_0^2)$.

The theoretical line shape (4), valid for $p > p_\theta$ as well as $p < p_\theta$, results from an ensemble average. The corresponding rescaled distribution $G(\eta)$ is represented by the solid line of Fig. 1. For $\eta < 1$, it is much flatter than the normalized Lorentzian $\mathcal{L}(\eta)$ having the same tails (dotted line of Fig. 1). This arises from the fact that, whatever the interaction time θ, atomic states with $|p| < p_\theta$ have even longer trapping times and thus cannot be discriminated by the experiment. This effect is even more dramatic if we use, instead of the "probabilistic" form (1), the "deterministic" form $P(\tau_S|p) = \delta[\tau_S - \tau(p)]$ which means that an atom trapped with momentum p remains in the trap during a well-defined time $\tau(p) = [R(p)]^{-1}$. In this case, $G(\eta)$ displays a plateau for $0 \leq \eta \leq 1$ and then decreases with the same asymptotic behavior in $1/\eta^2$ as seen before (dashed line of Fig. 1). With the more correct form (1) of $P(\tau_S|p)$, the plateau disappears but a clear flattening remains present.

It is challenging to obtain an experimental evidence for the nonergodicity of the cooling process by observing the

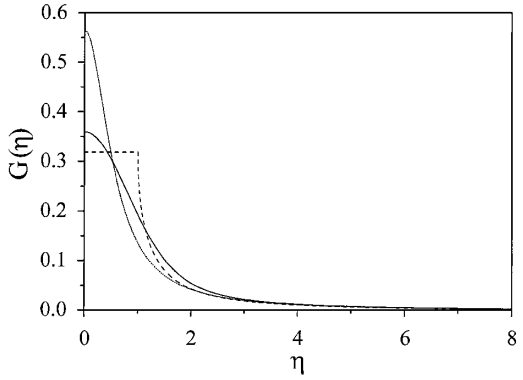

FIG. 1. Theoretical normalized momentum distribution as a function of the dimensionless parameter η with the probabilistic form of the delay function (solid line). Around $\eta = 0$, $G(\eta)$ is much flatter than the Lorentzian $\mathcal{L}(\eta)$ having the same asymptotic behavior (dotted line) and this is related to the nonergodicity of the VSCPT cooling. The effect is even stronger with the deterministic model (dashed line).

predicted shape (4) of $\mathcal{P}(p)$. With long interaction times, VSCPT reaches the nanokelvin range and it becomes hopeless to measure accurately $\mathcal{P}(p)$ with the usual time-of-flight technique. We rather use a new method, described in detail in [1], which gives the Fourier transform of $\mathcal{P}(p)$ with a high degree of accuracy. After a VSCPT cooling stage, an atom is left in a coherent superposition of two completely overlapping wave packets with opposite mean momenta $\pm\hbar k$. As soon as the VSCPT laser beams are switched off, the two wave packets freely fly apart. After a "dark period" of duration t_D, they are separated by $\xi = 2(\hbar k/M)t_D$. We then switch on the VSCPT light again for a short pulse of duration $\tau_A = 8\ \mu$s. During the dark period, the overlap of the two wave packets decreased and so did the destructive interference which inhibits the fluorescence for completely overlapping wave packets. A fraction of atoms therefore absorbs light and diffuses away in momentum space. We measure the remaining fraction Π_{NC} which is not coupled to the light. Since the overlap is simply given by the Fourier transform of $\mathcal{P}(p)$, we derive [13]:

$$\Pi_{\mathrm{NC}}(t_D) = \frac{1}{2} + \frac{1}{2}\int dp\ \mathcal{P}(p)\cos\left(\frac{2kpt_D}{M}\right). \quad (6)$$

Like $\mathcal{P}(p)$, the function $\Pi_{\mathrm{NC}}(t_D)$ is self-similar and can be rescaled yielding the function $\mathcal{F}(\zeta)$ where $\zeta = \tilde{t}_D/\tilde{\tau}_\theta$, $\tilde{t}_D = t_D\Gamma$, and $\tilde{\tau}_\theta = M\Gamma/(2kp_\theta) = \tilde{\theta}^{1/2}\tilde{\Omega}^{-1}$.

The experiment was performed after a careful cancellation of the magnetic field to 0.3 mG using the Hanle effect [14]. We measured $\tilde{\Omega}$ with a high degree of accuracy at the exact location of the atomic cloud, by use of the radiative broadening of the Hanle effect signal [15]. We took 12 sets of data with $\tilde{\Omega} = 0.72(2)$ and different values of $\tilde{\theta}$ ranging between 2000 and 15 000. We then rescaled each set according to $\tilde{t}_D \rightarrow \tilde{t}_D\tilde{\theta}^{-1/2} = \zeta/\tilde{\Omega}$ and removed the data points corresponding to $\tilde{t}_D\tilde{\theta}^{-1/2} < 0.8$ as explained in [13]. For the raw data (Fig. 2a) the 12 measurements do not have the same decay time since $\mathcal{P}(p)$ and thus $\Pi_{\mathrm{NC}}(\tilde{t}_D)$ depend upon $\tilde{\theta}$. After rescaling (Fig. 2b), the data collapse and lie very well on the theoretical curve corresponding to $\tilde{\Omega} = 0.72$, directly demonstrating the self-similarity of the process.

As explained above, a loss process could introduce ergodicity in the system and change $\mathcal{P}(p)$ into a Lorentzian. It is thus interesting to see if the data are well fitted by the expression (4) and if this fit is significantly better than the one using a Lorentzian momentum distribution. Since we measure the Fourier transform of $\mathcal{P}(p)$, we thus seek deviations of the measured $\mathcal{F}(\tilde{t}_D\tilde{\theta}^{-1/2})$ from an exponential. The average of the 12 sets of rescaled data is presented in Fig. 3 along with a fit by the calculated function $\mathcal{F}$ with $\tilde{\Omega}$ as the only adjustable parameter (solid

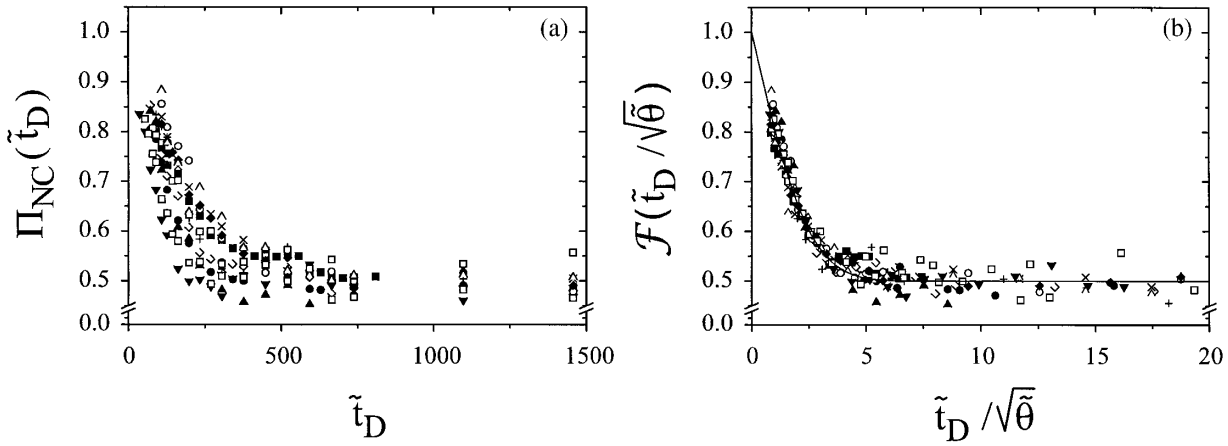

FIG. 2. Measured $\Pi_{\mathrm{NC}}(t_D)$ (data marks) before (a) and after (b) rescaling. These 12 data sets are obtained with the same Rabi frequency $\tilde{\Omega} = 0.72(2)$ but different values of the interaction time $\tilde{\theta}$ ranging from 2000 to 15 000. Since the final temperature depends upon $\tilde{\theta}$, the 12 curves of raw data do not have the same decay time. After rescaling, all the data collapse on the theoretical function (solid line).

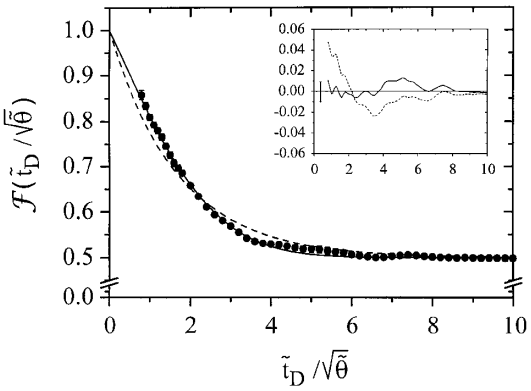

FIG. 3. Fit of the averaged data with the calculated function $\mathcal{F}$ (solid line) or a simple exponential (dashed line). Residues are presented in the inset. The vertical bar on the left is the largest error bar of the data. The data are well adjusted by the theoretical function ($\chi^2 = 0.0049$) and the fit gives $\tilde{\Omega} = 0.70$ in very good agreement with the measured value $\tilde{\Omega} = 0.72(2)$. The exponential fit leads to $\chi^2 = 0.0285$.

line) and a simple exponential fit (dashed line). The data fit well to the function derived from Lévy statistics and give $\tilde{\Omega} = 0.70(0)$, in very good agreement with the experimental value 0.72(2). On the contrary, the best exponential fit fails to reproduce the signal, especially for $\tilde{t}_D\tilde{\theta}^{-1/2} < 4$ where the discrepancies are 3 or 4 times larger than the standard deviation σ represented as error bars on the plot. This clearly shows that we are not in an ergodic regime limited by some loss process. It should be noted that the data exhibit a weak but statistically significant deviation from the theory around $\tilde{t}_D\tilde{\theta}^{-1/2} \simeq 5$. In order to check if a residual magnetic field could produce such an effect, we performed the same experiment in the presence of a controlled transverse magnetic field and found no modification of the signal as long as $B < 2$ mG. Since we compensate B with a precision of about 0.3 mG, we ruled out the possibility that this deviation could be due to a stray magnetic field. The largest deviation occurring at a time much larger than the decay time of $\mathcal{F}$, this could be related to some details of $\mathcal{P}(p)$ in the close vicinity of $p = 0$, which are not identified in our model.

In conclusion, we have used an interferometric method to determine the line shape of the momentum distribution of ultracold atoms with a great accuracy. Our results confirm the predictions of a theoretical analysis of VSCPT based on Lévy statistics and give clear evidence for fundamental features of subrecoil cooling such as self-similarity and nonergodicity. This not only gives a new lightening on subrecoil cooling, but also is one of the very few experimental evidences of anomalous diffusion processes [16] in laser cooling.

We thank F. Bardou, J. P. Bouchaud, A. Aspect, K. Madison, C. Salomon, and J. Dalibard for fruitful discussions. Laboratoire Kastler-Brossel is a Laboratoire de l'ENS et de l'Université Paris VI, associé au CNRS.

[1] B. Saubaméa, T. W. Hijmans, S. Kulin, E. Rasel, E. Peik, M. Leduc, and C. Cohen-Tannoudji, Phys. Rev. Lett. **79**, 3146 (1997).
[2] C. Tsallis, Phys. World **10**, 42 (1997).
[3] J. P. Bouchaud and A. Georges, Phys. Rep. **195**, 125 (1990).
[4] A. Aspect, E. Arimondo, R. Kaiser, N. Vansteenkiste, and C. Cohen-Tannoudji, Phys. Rev. Lett. **61**, 826 (1988).
[5] M. Kasevich and S. Chu, Phys. Rev. Lett. **69**, 1741 (1992).
[6] F. Bardou, J. P. Bouchaud, O. Emile, A. Aspect, and C. Cohen-Tannoudji, Phys. Rev. Lett. **72**, 203 (1994).
[7] F. Bardou, J. P. Bouchaud, A. Aspect, and C. Cohen-Tannoudji (to be published).
[8] C. Cohen-Tannoudji and J. Dalibard, Europhys. Lett. **1**, 441 (1986).
[9] P. Zoller, M. Marte, and D. F. Walls, Phys. Rev. A **35**, 198 (1987).
[10] Strictly speaking, $P(\tau_S|p)$ is not a single exponential, but the square of a sum of exponentials with time constants related to the relaxation modes of the system. Here we keep only the exponential with the longest time constant which determines the long time behavior of the system. Keeping the other exponentials would change some prefactors but not the shape of the momentum distribution.
[11] The exact value of p_{trap} is irrelevant since for θ long enough, $\delta p \ll p_{\text{trap}}$ and $\mathcal{P}(p)$ is independent of p_{trap}.
[12] For a detuning $\delta = 0$, as in our experiment, one has $\tau_0 = 2\Gamma/\Omega^2$ and $p_0 = (M\Omega^2)/(2\sqrt{2}k\Gamma)$, from which the expression of p_θ given in Eq. (5) is derived.
[13] During the probe pulse of duration τ_A, the atoms with $R(p)\tau_A > 1$ absorb light. The atomic momentum distribution is thus truncated beyond a momentum much larger than p_θ if τ_A is short enough. Correlatively, the shape of $\Pi_{\text{NC}}(t_D)$ is also changed, but only in a small interval about 0, much smaller than the decay time of $\Pi_{\text{NC}}(t_D)$. The derivation of Eq. (6), explicitly given in [1], does not include this effect. To analyze our data, we therefore removed the very first values of $\Pi_{\text{NC}}(t_D)$. A more detailed analysis, including this perturbation, is presented in [15].
[14] R. Kaiser, N. Vansteenkiste, A. Aspect, E. Arimondo, and C. Cohen-Tannoudji, Z. Phys. D **18**, 17 (1991).
[15] B. Saubaméa, thesis, Université Paris VI, 1998.
[16] B. G. Klappauf, W.H. Oskay, D.A. Steck, and M. G. Raizen, Phys. Rev. Lett. **81**, 4044 (1998).