Laser cooling and trapping of neutral atoms*

William D. Phillips
National Institute of Standards and Technology, Physics Laboratory, Atomic Physics Division, Gaithersburg, Maryland 20899

INTRODUCTION

In 1978, while I was a postdoctoral fellow at MIT, I read a paper by Art Ashkin (1978) in which he described how one might slow down an atomic beam of sodium using the radiation pressure of a laser beam tuned to an atomic resonance. After being slowed, the atoms would be captured in a trap consisting of focused laser beams, with the atomic motion being damped until the temperature of the atoms reached the microkelvin range. That paper was my first introduction to laser cooling, although the idea of laser cooling (the reduction of random thermal velocities using radiative forces) had been proposed three years earlier in independent papers by Hänisch and Schawlow (1975) and Wineland and Dehmelt (1975). Although the treatment in Ashkin's paper was necessarily over-simplified, it provided one of the important inspirations for what I tried to accomplish for about the next decade. Another inspiration appeared later that same year: Wineland, Drullinger and Walls (1978) published the first laser cooling experiment, in which they cooled a cloud of Mg ions held in a Penning trap. At essentially the same time, Neuhauser, Hohenstatt, Toschek and Dehmelt (1978) also reported laser cooling of trapped Ba⁺ ions.

Those laser cooling experiments of 1978 were a dramatic demonstration of the mechanical effects of light, but such effects have a much longer history. The understanding that electromagnetic radiation exerts a force became quantitative only with Maxwell's theory of electromagnetism, even though such a force had been conjectured much earlier, partly in response to the observation that comet tails point away from the sun. It was not until the turn of the century, however, that experiments by Lebedev (1901) and Nichols and Hull (1901, 1903) gave a laboratory demonstration and quantitative measurement of radiation pressure on macroscopic objects. In 1933 Frisch made the first demonstration of light pressure on atoms, deflecting an atomic sodium beam with resonance radiation from a lamp. With the advent of the laser, Ashkin (1970) recognized the potential of intense, narrow-band light for manipulating atoms and in 1972 the first "modern" experiments demonstrated the deflection of atomic beams with lasers (Picqué and Vialle, 1972; Schieder et al., 1972). All of this set the stage for the laser cooling proposals of 1975 and for the demonstrations in 1978 with ions.

Comet tails, deflection of atomic beams and the laser cooling proposed in 1975 are all manifestations of the radiative force that Ashkin has called the "scattering force," because it results when light strikes an object and is scattered in random directions. Another radiative force, the dipole force, can be thought of as arising from the interaction between an induced dipole moment and the gradient of the incident light field. The dipole force was recognized at least as early as 1962 by Askar'yan, and in 1968, Letokhov proposed using it to trap atoms—even before the idea of laser cooling! The trap proposed by Ashkin in 1978 relied on this "dipole" or "gradient" force as well. Nevertheless, in 1978, laser cooling, the reduction of random velocities, was understood to involve only the scattering force. Laser trapping, confinement in a potential created by light, which was still only a dream, involved both dipole and scattering forces. Within 10 years, however, the dipole force was seen to have a major impact on laser cooling as well.

Without understanding very much about what difficulties lay in store for me, or even appreciating the exciting possibilities of what one might do with laser cooled atoms, I decided to try to do for neutral atoms what the groups in Boulder and Heidelberg had done for ions: trap them and cool them. There was, however, a significant difficulty: we could not first trap and then cool neutral atoms. Ion traps were deep enough to easily trap ions having temperatures well above room temperature, but none of the proposed neutral atom traps had depths of more than a few kelvin. Significant cooling was required before trapping would be possible, as Ashkin had outlined in his paper (1978), and it was with this idea that I began.

Before describing the first experiments on the deceleration of atomic beams, let me digress slightly and discuss why laser cooling is so exciting and why it has attracted so much attention in the scientific community: When one studies atoms in a gas, they are typically moving very rapidly. The molecules and atoms in air at room temperature are moving with speeds on the order of 300 m/s, the speed of sound. This thermal velocity can be reduced by refrigerating the gas, with the velocity varying as the square root of the temperature, but even at 77 K, the temperature at which N₂ condenses into a liquid, the nitrogen molecules are moving at about 150 m/s. At 4 K, the condensation temperature of helium, the He atoms have 90 m/s speeds. At temperatures for which atomic thermal velocities would be below 1 m/s, any gas in equilibrium (other than spin-polarized atomic hydro-

*The 1997 Nobel Prize in Physics was shared by Steven Chu, Claude N. Cohen-Tannoudji, and William D. Phillips. This text is based on Dr. Phillips's address on the occasion of the award.
LASER COOLING OF ATOMIC BEAMS

In 1978 I had only vague notions about the excitement that lay ahead with laser cooled atoms, but I concluded that slowing down an atomic beam was the first step. The atomic beam was to be slowed using the transfer of momentum that occurs when an atom absorbs a photon. Figure 1 shows the basic process underlying the “scattering force” that results. An atomic beam with velocity $v$ is irradiated by an opposing laser beam. For each photon that a ground-state atom absorbs, it is slowed by $v_{rec} = \frac{\hbar k}{m}$. In order to absorb again the atom must return to the ground state by emitting a photon. Photons are emitted in random directions, but with a symmetric average distribution, so their contribution to the atom’s momentum averages to zero. The randomness results in a “heating” of the atom, discussed below.

For sodium atoms interacting with the familiar yellow resonance light, $v_{rec} = 3 \text{ cm/s}$, while a typical beam velocity is about $10^5 \text{ cm/s}$, so the absorption-emission process must occur about $3 \times 10^4$ times to bring the Na atom to rest. In principle, an atom could radiate and absorb photons at half the radiative decay rate of the excited state (a 2-level atom in steady state can spend at most half of its time in the excited state). For Na, this implies that a photon could be radiated every $32 \text{ ns}$ on average, bringing the atoms to rest in about $1 \text{ ms}$. Two problems, optical pumping and Doppler shifts, can prevent this from happening. I had an early indication of the difficulty of decelerating an atomic beam shortly after reading Ashkin’s 1978 paper. I was then working with a sodium atomic beam at MIT, using tunable dye lasers to study the scattering properties of optically excited sodium. I tuned a laser to be resonant with the Na transition from $3S_{1/2} \rightarrow 3P_{3/2}$, the D2 line, and directed its beam opposite to the atomic beam. I saw that the atoms near the beam source were fluorescing brightly as they absorbed the laser light, while further away from the source, the atoms were relatively dim. The problem, I concluded, was optical pumping, illustrated in Fig. 2.

Sodium is not a two-level atom, but has two ground hyperfine levels ($F = 1$ and $F = 2$ in Fig. 2), each of which consists of several, normally degenerate, states. Laser excitation out of one of the hyperfine levels to the excited state can result in the atom radiating to the other hyperfine level. This optical pumping essentially shuts off the absorption of laser light, because the linewidths of the transition and of the laser are much smaller than the separation between the ground state hyperfine com-

FIG. 1. (a) An atom with velocity $v$ encounters a photon with momentum $\hbar k = \hbar/\lambda$; (b) after absorbing the photon, the atom is slowed by $\hbar k/m$; (c) after re-radiation in a random direction, on average the atom is slower than in (a).

FIG. 2. (a) The optical pumping process preventing cycling transitions in alkalis like Na; (b) use of a repumping laser to allow many absorption-emission cycles.
ponents. Even for atoms excited on the $3S_{1/2} \ (F=2) \rightarrow 3P_{3/2} \ (F'=3)$ transition, where the only allowed decay channel is to $F=2$, off-resonant excitation of $F'=2$ (the linewidth of the transition is 10 MHz, while the separation between $F'=2$ and $F'=3$ is 60 MHz) leads to optical pumping into $F=1$ after only about a hundred absorptions. This optical pumping made the atoms “dark” to my laser after they traveled only a short distance from the source.

An obvious solution [Fig. 2(b)] is to use a second laser frequency, called a repumper, to excite the atoms out of the “wrong” ($F=1$) hyperfine state so that they can decay to the “right” state ($F=2$) where they can continue to cool. Given the repumper, another problem becomes apparent: the Doppler shift. In order for the laser light to be resonantly absorbed by a counterpropagating atom moving with velocity $v$, the frequency $\omega$ of the light must be $kv$ lower than the resonant frequency for an atom at rest. As the atom repeatedly absorbs photons, slowing down as desired, the Doppler shift changes and the atom goes out of resonance with the light. The natural linewidth $\Gamma/2\pi$ of the optical transition in Na is 10 MHz (full width at half maximum). A change in velocity of 6 m/s gives a Doppler shift this large, so after absorbing only 200 photons, the atom is far enough off resonance that the rate of absorption is significantly reduced. The result is that only atoms with the “proper” velocity to be resonant with the laser are slowed, and they are only slowed by a small amount.

Nevertheless, this process of atoms being slowed and pushed out of resonance results in a cooling or narrowing of the velocity distribution. In an atomic beam, there is typically a wide spread of velocities around $v_\text{th} = 3k_B T/m$. Those atoms with the proper velocity will absorb rapidly and decelerate. Those that are too fast will absorb more slowly, then more rapidly as they come into resonance, and finally more slowly as they continue to decelerate. Atoms that are too slow to begin with will absorb little and decelerate little. Thus atoms from a range of velocities around the resonant velocity are pushed into a narrower range centered on a lower velocity. This process was studied theoretically by Minogin (1980) and in 1981, at Moscow’s Institute for Spectroscopy, was used in the first experiment clearly demonstrating laser cooling of neutral atoms (Andreev et al., 1981).

Figure 3 shows the velocity distribution after such cooling of an atomic beam. The data was taken in our laboratory, but is equivalent to what had been done in Moscow. The characteristic of this kind of beam cooling is that only a small part of the total velocity distribution (the part near resonance with the laser beam) is slowed by only a small amount (until the atoms are no longer resonant). The narrow peak, while it represents true cooling in that its velocity distribution is narrow, consists of rather fast atoms.

One solution to this problem had already been outlined in 1976 by Letokhov, Minogin, and Pavlik. They suggested a general method of changing the frequency (chirping) of the cooling laser so as to interact with all the atoms in a wide distribution and to stay in resonance with the atoms as they are cooled. The Moscow group applied the technique to decelerating an atomic beam (Balykin et al., 1979) but without clear success (Balykin, 1980). [Later, in 1983, John Prodan and I obtained the first clear deceleration and cooling of an atomic beam with this “chirp-cooling” technique (Phillips and Prodan, 1983, 1984; Phillips, Prodan, and Metcalf, 1983a; Prodan and Phillips, 1984). Those first attempts failed to bring the atoms to rest, something that was finally achieved by Ertmer, Blatt, Hall and Zhu (1985).] The chirp-cooling technique is now one of the two standard methods for decelerating beams. The other is “Zeeman cooling.”

By late 1978, I had moved to the National Bureau of Standards (NBS), later named the National Institute of Standards and Technology (NIST), in Gaithersburg. I was considering how to slow an atomic beam, realizing that the optical pumping and Doppler shift problems would both need to be addressed. I understood how things would work using the Moscow chirp-cooling technique and a repumper. I also considered using a broadband laser, so that there would be light in resonance with the atoms, regardless of their velocity. [This idea was refined by Hoffnagle (1988) and demonstrated by Hall’s group (Zhu, Oates, and Hall, 1991).] Finally I considered that instead of changing the frequency of the laser to stay in resonance with the atoms (chirping), one could use a magnetic field to change the energy level separation in the atoms so as to keep them in resonance with the fixed-frequency laser (Zeeman cooling). All of these ideas for cooling an atomic beam, along with vari-

![FIG. 3. Cooling an atomic beam with a fixed frequency laser.](image-url)
ous schemes for avoiding optical pumping, were contained in a proposal (Phillips, 1979) that I submitted to the Office of Naval Research in 1979. Around this time Hal Metcalf, from the State University of New York at Stony Brook, joined me in Gaithersburg and we began to consider what would be the best way to proceed. Hal contended that all the methods looked reasonable, but we should work on the Zeeman cooler because it would be the most fun! Not only was Hal right about the fun we would have, but his suggestion led us to develop a technique with particularly advantageous properties. The idea is illustrated in Fig. 4.

The atomic beam source directs atoms, which have a wide range of velocities, along the axis ($z$ direction) of a tapered solenoid. This magnet has more windings at its entrance end, near the source, so the field is higher at that end. The laser is tuned so that, given the field-induced Zeeman shift and the velocity-induced Doppler shift of the atomic transition frequency, atoms with velocity $v_0$ are resonant with the laser when they reach the point where the field is maximum. Those atoms then absorb light and begin to slow down. As their velocity changes, their Doppler shift changes, but is compensated by the change in Zeeman shift as the atoms move to a point where the field is weaker. At this point, atoms with initial velocities slightly lower than $v_0$ come into resonance and begin to slow down. The process continues with the initially fast atoms decelerating and staying in resonance while initially slower atoms come into resonance and begin to be slowed as they move further down the solenoid. Eventually all the atoms with velocities lower than $v_0$ are brought to a final velocity that depends on the details of the magnetic field and laser tuning.

The first tapered solenoids that Hal Metcalf and I used for Zeeman cooling of atomic beams had only a few sections of windings and had to be cooled with air blown by fans or with wet towels wrapped around the coils. Shortly after our initial success in getting some substantial deceleration, we were joined by my first postdoc, John Prodan. We developed more sophisticated solenoids, wound with wires in many layers of different lengths, so as to produce a smoothly varying field that would allow the atoms to slow down to a stop while remaining in resonance with the cooling laser.

These later solenoids were cooled with water flowing over the coils. To improve the heat transfer, we filled the spaces between the wires with various heat-conducting substances. One was a white silicone grease that we put onto the wires with our hands as we wound the coil on a lathe. The grease was about the same color and consistency as the diaper rash ointment I was then using on my baby daughters, so there was a period of time when, whether at home or at work, I seemed to be up to my elbows in white grease.

The grease-covered, water-cooled solenoids had the annoying habit of burning out as electrolytic action attacked the wires during operation. Sometimes it seemed that we no sooner obtained some data than the solenoid would burn out and we were winding a new one.

On the bright side, the frequent burn-outs provided the opportunity for refinement and redesign. Soon we were embedding the coils in a black, rubbery resin. While it was supposed to be impervious to water, it did not have good adhesion properties (except to clothing and human flesh) and the solenoids continued to burn out. Eventually, an epoxy coating sealed the solenoid against the water that allowed the electrolysis, and in more recent times we replaced water with a fluorocarbon liquid that does not conduct electricity or support electrolysis. Along the way to a reliable solenoid, we learned how to slow and stop atoms efficiently (Phillips and Metcalf, 1982; Prodan, Phillips, and Metcalf, 1982; Phillips, Prodan, and Metcalf, 1983a, 1983b, 1984a, 1984b, 1985; Metcalf and Phillips, 1985).

The velocity distribution after deceleration is measured in a detection region some distance from the exit end of the solenoid. Here a separate detection laser beam produces fluorescence from atoms having the correct velocity to be resonant. By scanning the frequency of the detection laser, we were able to determine the velocity distribution in the atomic beam. Observations with the detection laser were made just after turning off the cooling laser, so as to avoid any difficulties with having both lasers on at the same time. Figure 5 shows the velocity distribution resulting from Zeeman cooling: a large fraction of the initial distribution has been swept down into a narrow final velocity group.

One of the advantages of the Zeeman cooling technique is the ease with which the optical pumping problem is avoided. Because the atoms are always in a strong axial magnetic field (that is the reason for the “bias” windings in Fig. 4), there is a well-defined axis of quantization that allowed us to make use of the selection rules for radiative transitions and to avoid the undesirable optical pumping. Figure 6 shows the energy levels of Na in a magnetic field. Atoms in the $3S_{1/2} (m_F=2)$ state, irradiated with circularly polarized $\sigma^+$ light, must increase their $m_F$ by one unit, and so can go only to the $3P_{3/2} (m_F=3)$ state. This state in turn can decay only to $3S_{1/2} (m_F=2)$, and the excitation process can be repeated indefinitely. Of course, the circular polarization is not perfect, so other excitations are possible, and these may lead to decay to other states. Fortunately, in a high magnetic field, such transitions are highly unlikely.

![FIG. 4. Upper: Schematic representation of a Zeeman slower. Lower: Variation of the axial field with position.](image-url)
either they involve a change in the nuclear spin projection $m_I$, which is forbidden in the high field limit, or they are far from resonance. These features, combined with high purity of the circular polarization, allowed us to achieve, without a "wrong transition," the $3 \times 10^4$ excitations required to stop the atoms. Furthermore, the circular polarization produced some "good" optical pumping; atoms not initially in the $3S_{1/2} (m_F=2)$ state were pumped into this state, the "stretched" state of maximum projection of angular momentum, as they absorbed the angular momentum of the light. These various aspects of optical selection rules and optical pumping allowed the process of Zeeman cooling to be very efficient, decelerating a large fraction of the atoms in the beam.

In 1983 we discussed a number of these aspects of laser deceleration, including our early chirp-cooling results, at a two-day workshop on "Laser-Cooled and Trapped Atoms" held at NBS in Gaithersburg (Phillips, 1983). I view this as an important meeting in that it and its proceedings stimulated interest in laser cooling. In early 1984, Stig Stenholm, then of the University of Helsinki, organized an international meeting on laser cooling in Tvarminne, a remote peninsula in Finland. Figure 7 shows the small group attending (I was the photographer), and in that group, only some of the participants were even active in laser cooling at the time. Among these were Stig Stenholm [who had done pioneering work in the theory of laser cooling and the mechanical effects of light on atoms (Stenholm, 1978a, 1978b, 1985, 1986; Javanainen and Stenholm, 1980a, 1980b, 1980c, 1981a, 1981b)] along with some of his young colleagues; Victor Balykin and Vladimir Minogin from the Moscow group; and Claude Cohen-Tannoudji and Jean Dalibard from Ecole Normale Superieure (ENS) in Paris, who had begun working on the theory of laser cooling and trapping. Also present were Jurgen Mlynek and Wolfgang Ertmer, both of whom now lead major research groups pursuing laser cooling and atom optics. At that time, however, only our group and the Moscow group had published any experiments on cooling of neutral atoms.

Much of the discussion at the Tvarminne meeting involved the techniques of beam deceleration and the problems with optical pumping. I took a light-hearted attitude toward our trials and tribulations with optical pumping, often joking that any unexplained features in our data could certainly be attributed to optical pumping. Of course, at the Ecole Normale, optical pumping had a long and distinguished history. Having been pioneered by Alfred Kastler and Jean Brossel, optical pumping had been the backbone of many experiments in the Laboratoire de Spectroscopie Hertzienne (now the Laboratoire Kastler-Brossel). After one discussion in which I had joked about optical pumping, Jean Dalibard privately mentioned to me, "You know, Bill, at the Ecole Normale, optical pumping is not a joke." His gentle note of caution calmed me down a bit, but it turned out to be strangely prophetic as well. As we saw a few years later, optical pumping had an important, beautiful, and totally unanticipated role to play in laser cooling, and it was surely no joke.

STOPPING ATOMS

As successful as Zeeman cooling had been in producing large numbers of decelerated atoms as in Fig. 5, we
had not actually observed the atoms at rest, nor had we trapped them. In fact, I recall a conversation with Steve Chu that took place during the International Conference on Laser Spectroscopy in Interlaken in 1983 in which I had presented our results on beam deceleration (Phillips, Prodan, and Metcalf, 1983a). Steve was working on positronium spectroscopy but was wondering whether there still might be something interesting to be done with laser cooling of neutral atoms. I offered the opinion that there was still plenty to do, and in particular, that trapping of atoms was still an unrealized goal. It wasn’t long before each of us achieved that goal, in very different ways.

Our approach was first to get some stopped atoms. The problem had been that, in a sense, Zeeman cooling worked too well. By adjusting the laser frequency and magnetic field, we could, up to a point, choose the final velocity of the atoms that had undergone laser deceleration. Unfortunately, if we chose too small a velocity, no slow atoms at all appeared in the detection region. Once brought below a certain velocity, about 200 m/s, the atoms always continued to absorb enough light while traveling from the solenoid to the detection region so as to stop before reaching the detector. By shutting off the cooling laser beam and delaying observation until the slow atoms arrived in the observation region, we were able to detect atoms as slow as 40 m/s with a spread of 10 m/s, corresponding to a temperature (in the atoms’ rest frame) of 70 mK (Prodan, Phillips, and Metcalf, 1982).

The next step was to get these atoms to come to rest in our observation region. We were joined by Alan Migdall, a new postdoc, Jean Dalibard, who was visiting from ENS, and Ivan So, Hal Metcalf’s student. We decided that we needed to proceed as before, shutting off the cooling light, allowing the slow atoms to drift into the observation region, but then to apply a short pulse of additional cooling light to bring the atoms to rest. The sequence of laser pulses required to do this—a long pulse of several milliseconds for doing the initial deceleration, followed by a delay and then another pulse of a few hundred microseconds, followed by another delay before detection—was provided by a rotating wheel with a series of openings corresponding to the places where the laser was to be on. Today we accomplish such pulse sequences with acousto-optic modulators under computer control, but in those days it required careful construction and balancing of a rapidly rotating wheel.

The result of this sequence of laser pulses was that we had atoms at rest in our observation region with a velocity spread corresponding to <100 mK (Prodan et al., 1985). Just following our 1985 paper reporting this in Physical Review Letters was a report of the successful stopping of atoms by the chirp-cooling method in Jan Hall’s group (Ertmer, Blatt, Hall, and Zhu, 1985). At last there were atoms slow enough to be trapped, and we decided to concentrate first on magnetostatic trapping.

MAGNETIC TRAPPING OF ATOMS

The idea for magnetic traps had first appeared in the literature as early as 1960 (Heer, 1960, 1963; Vladimirskii, 1960), although Wolfgang Paul had discussed them in lectures at the University of Bonn in the mid 1950s, as a natural extension of ideas about magnetic focusing of...
atomic beams (Vauthier, 1949; Friedburg, 1951; Friedburg and Paul, 1951). Magnetic trapping had come to our attention particularly because of the successful trapping of cold neutrons (Kugler et al., 1978). We later learned that in unpublished experiments in Paul’s laboratory, there were indications of confining sodium in a magnetic trap (Martin, 1975).

The idea of magnetic trapping is that in a magnetic field, an atom with a magnetic moment will have quantum states whose magnetic or Zeeman energy increases with increasing field and states whose energy decreases, depending on the orientation of the moment compared to the field. The increasing-energy states, or low-field-seekers, can be trapped in a magnetic field configuration having a point where the magnitude of the field is a relative minimum. [No dc field can have a relative maximum in free space (Wing, 1984), so high-field-seekers cannot be trapped.] The requirement for stable trapping, besides the kinetic energy of the atom being low enough, is that the magnetic moment move adiabatically in the field. That is, the orientation of the magnetic moment with respect to the field should not change.

We considered some of the published designs for trapping neutrons, including the spherical hexapole (Golub and Pendlebury, 1979), a design comprising three current loops, but we found them less than ideal. Instead we decided upon a simpler design, with two loops, which we called a spherical quadrupole. The trap, its magnetic field lines and equipotentials are shown in Fig. 8. Although we thought that we had discovered an original trap design, we later learned that Wolfgang Paul had considered this many years ago, but had not given it much attention because atoms were not harmonically bound in such a trap. In fact, the potential for such a trap is linear in the displacement from the center and has a cusp there.

With a team consisting of Alan Migdall, John Prodan, Hal Metcalf and myself, and with the theoretical support of Tom Bergeman, we succeeded in trapping atoms in the apparatus shown in Fig. 9 (Migdall et al., 1985). As in the experiments that stopped atoms, we start with Zeeman slowing, decelerating the atoms to 100 m/s in the solenoid. The slowing laser beam is then extinguished, allowing the atoms to proceed unhindered for 4 ms to the magnetic trap. At this point, only one of the two trap coils has current; it produces a magnetic field that brings the atoms into resonance with the cooling laser when it is turned on again for 400 μs, bringing the atoms to rest. Once the atoms are stopped, the other coil is energized, producing the field shown in Fig. 8, and the trap is sprung. The atoms are held in the trap until released, or until collisions with the room-temperature background gas molecules in the imperfect vacuum knock them out. After the desired trapping time, we turn off the magnetic field, and turn on a probe laser, so as to see how many atoms remain in the trap. By varying the frequency of this probe on successive repetitions of the process, we could determine the velocity distribution of the atoms, via their Doppler shifts.

The depth of our trap was about 17 mK (25 mT), corresponding to Na atoms with a velocity of 3.5 m/s. In the absence of trapping fields, atoms that fast would escape from the region of the trap coils in a few milliseconds. Figure 10 shows a section of chart paper with spectra of the atoms remaining after 35 ms of trapping time. If the trap had not been working, we would have seen essentially nothing after that length of time, but the signal, noisy as it was, was unmistakable. It went away when the trap was off, and it went away when we did not provide the second pulse of cooling light that stops the atoms before trapping them. This was just the signature we were looking for, and Hal Metcalf expressed his characteristic elation at good results with his exuberant “WAHOO!!” at the top of the chart.

As the evening went on, we were able to improve the signal, but we found that the atoms did not stay very long in the trap, a feature we found a bit frustrating. Finally, late in the evening we decided to go out and get some fast food, talk about what was happening and attack the problem afresh. When we returned a little later that night, the signal had improved and we were able to trap atoms for much longer times. We soon realized that during our supper break the magnetic trap had cooled down, and stopped outgassing, so the vacuum just in the
vicinity of the trap improved considerably. With this insight we knew to let the magnet cool off from time to time, and we were able to take a lot of useful data. We continued taking data until around 5:00 am, and it was probably close to 6:00 am when my wife Jane found Hal and me in our kitchen, eating ice cream as she prepared to leave for work. Her dismay at the lateness of our return and our choice of nourishment at that hour was partially assuaged by Hal’s assurance that we had accomplished something pretty important that night.

Figure 11(a) presents the sequence of spectra taken after various trapping times, showing the decrease in signal as atoms are knocked out of the trap by collisions with the background gas molecules. Figure 11(b) shows that the loss of atoms from the trap is exponential, as expected, with a lifetime of a bit less than one second, in a vacuum of a few times $10^{-6}$ pascals. A point taken when the vacuum was allowed to get worse illustrates that poor vacuum made the signal decay faster. In more recent times, we and others have achieved much longer trapping times, mainly because of an improved vacuum. We now observe magnetic trap lifetimes of one minute or longer in our laboratory.

Since our demonstration (Migdall et al., 1985) of magnetic trapping of atoms in 1985, many different kinds of magnetic atom traps have been used. At MIT, Dave Pritchard’s group trapped (Bagnato et al., 1987) and cooled (Helmerson et al., 1992) Na atoms in a linear quadrupole magnetic field with an axial bias field, similar to the trap first discussed by Ioffe and collaborators (Gott, Ioffe, and Telkovsky, 1962) in 1962, and later by others (Pritchard, 1983; Bergeman et al., 1987). Similar traps were used by the Kleppner-Greytak group to trap (Hess et al., 1987) and evaporatively cool (Masuhura et al., 1988) atomic hydrogen, and by Walraven’s group to trap (van Roijen et al., 1988) and laser-cool hydrogen (Setija et al., 1994). The Ioffe trap has the advantage of having a non-zero magnetic field at the equilibrium point, in contrast to the spherical quadrupole, in which the field is zero at the equilibrium point. The zero field allows the magnetic moment of the atom to flip (often called Majorana flopping), so that the atom is in an untrapped spin state. While this problem did not cause difficulties in our 1985 demonstration, for colder atoms, which spend more time near the trap center, it can be a quite severe loss mechanism (Davis, Mewes, Joffe et al., 1995; Petrich et al., 1995). In 1995, modifications to the simple quadrupole trap solved the problem of spins flips near the trap center, and allowed the achievement of Bose-Einstein condensation (Anderson et al., 1995; Davis, Mewes, Andrews et al., 1995).

OPTICAL MOLASSES

At the same time that we were doing the first magnetic trap experiments in Gaithersburg, the team at Bell Labs, led by Steve Chu, was working on a different and extremely important feature of laser cooling. After a beautiful demonstration in 1978 of the use of optical forces to focus an atomic beam (Bjorkholm et al., 1978), the Bell Labs team had made some preliminary attempts to decelerate an atom beam, and then moved on to
other things. Encouraged by the beam deceleration experiments in Gaithersburg and in Boulder, Steve Chu reassembled much of that team and set out to demonstrate the kind of laser cooling suggested in 1975 by Hänsch and Schawlow. [The physical principles behind the Hänsch and Schawlow proposal are, of course, identical to those expressed in the 1975 Wineland and Dehmelt laser cooling proposal. These principles had already led to the laser cooling of trapped ions (Neuhauser et al., 1978; Wineland et al., 1978). The foci of Hänsch and Schawlow (1975) and Wineland and Dehmelt (1975), however, has associated the former with neutral atoms and the latter with ions.] In fact, the same physical principle of Doppler cooling results in the compression of the velocity distribution associated with laser deceleration of an atomic beam [see sections 2 and 3 of Phillips (1992)]. Nevertheless, in 1985, laser cooling of a gas of neutral atoms at rest, as proposed in Hänsch and Schawlow (1975), had yet to be demonstrated.

The idea behind the Hänsch and Schawlow proposal is illustrated in Fig. 12. A gas of atoms, represented here in one dimension, is irradiated from both sides by laser beams tuned slightly below the atomic resonance frequency. An atom moving toward the left sees that the laser beam opposing its motion is Doppler shifted toward the atomic resonance frequency. It sees that the laser beam directed along its motion is Doppler shifted further from its resonance. The atom therefore absorbs more strongly from the laser beam that opposes its motion, and it slows down. The same thing happens to an atom moving to the right, so all atoms are slowed by this arrangement of laser beams. With pairs of laser beams added along the other coordinate axes, one obtains cooling in three dimensions. Because of the role of the Doppler effect in the process, this is now called Doppler cooling.

Later treatments (Letokhov et al., 1977; Neuhauser et al., 1978; Stenholm, 1978a; Wineland et al., 1978; Wineland and Itano, 1979; Javanainen, 1980; Javanainen and Stenholm, 1980b) recognized that this cooling process leads to a temperature whose lower limit is on the order of $\hbar \Gamma$, where $\Gamma$ is the rate of spontaneous emission of the excited state ($\Gamma^{-1}$ is the excited state lifetime). The temperature results from an equilibrium between laser cooling and the heating process arising from the random nature of both the absorption and emission of photons. The random addition to the average momentum transfer produces a random walk of the atomic momentum and an increase in the mean square atomic momentum. This heating is countered by the cooling force $F$ opposing atomic motion. The force is proportional to the atomic velocity, as the Doppler shift is proportional to velocity. In this, the cooling force is similar to the friction force experienced by a body moving in a viscous fluid. The rate at which energy is removed by cooling is $F \cdot v$, which is proportional to $v^2$, so the cooling rate is proportional to the kinetic energy. By contrast the heating rate, proportional to the total photon scattering rate, is independent of atomic kinetic energy for low velocities. As a result, the heating and cooling come to equi-
librium at a certain value of the average kinetic energy. This defines the temperature for Doppler cooling, which is

\[ m \langle v_i^2 \rangle = k_B T = \frac{\hbar \Gamma}{4} \left( \frac{\Gamma}{2 \delta} + \frac{2 \delta}{\Gamma} \right), \]

where \( \delta \) is the angular frequency of the detuning of the lasers from atomic resonance and \( v_i \) is the velocity along some axis. This expression is valid for 3D Doppler cooling in the limit of low intensity and when the recoil energy \( \hbar k_i^2 v_i^2 / 2m \ll \hbar \Gamma \). Interestingly, the equilibrium velocity distribution for Doppler cooling is the Maxwell-Boltzmann distribution. This follows from the fact that the Fokker-Planck equation describing the damping and heating in laser cooling is identical in form to the equation that describes collisional equilibrium of a gas (Stenholm, 1986). Numerical simulations of real cases, where the recoil energy does not vanish, show that the distribution is still very close to Maxwellian (Lett et al., 1989).

The minimum value of this temperature is called the Doppler cooling limit, occurring when \( \delta = -\Gamma/2 \),

\[ k_B T_{Dopp} = \frac{\hbar \Gamma}{2}. \]

The first rigorous derivation of the cooling limit appears to be by Letokhov, Minogin and Pavilik (1977) [although the reader should note that Eq. (32) is incorrectly identified with the rms velocity]. Wineland and Itano (1979) give derivations for a number of different situations involving trapped and free atoms and include the case where the recoil energy is not small but the atoms are in collisional equilibrium.

The Doppler cooling limit for sodium atoms cooled on the resonance transition at 589 nm where \( \Gamma/2\pi = 10 \text{ MHz} \), is 240 \( \mu \text{K} \), and corresponds to an rms velocity of 30 cm/s along a given axis. The limits for other atoms and ions are similar, and such low temperatures were quite appealing. Before 1985, however, these limiting temperatures had not been obtained in either ions or neutral atoms.

A feature of laser cooling not appreciated in the first treatments was the fact that the spatial motion of atoms in any reasonably sized sample would be diffusive. For example, a simple calculation (Lett et al., 1989) shows that a sodium atom cooled to the Doppler limit has a “mean free path” (the mean distance it moves before its initial velocity is damped out and the atom is moving with a different, random velocity) of only 20 \( \mu \text{m} \), while the size of the laser beams doing the cooling might easily be one centimeter. Thus, the atom undergoes diffusive, Brownian-like motion, and the time for a laser cooled atom to escape from the region where it is being cooled is much longer than the ballistic transit time across that region. This means that an atom is effectively “stuck” in the laser beams that cool it. This stickiness, and the similarity of laser cooling to viscous friction, prompted the Bell Labs group (Chu et al., 1985) to name the intersecting laser beams “optical molasses.” At NBS (Phillips, Prodan, and Metcalf, 1985), we independently used the term “molasses” to describe the cooling configuration, and the name “stuck.” Note that an optical molasses is not a trap. There is no restoring force keeping the atoms in the molasses, only a viscous inhibition of their escape.

Using the techniques for chirp cooling an atomic beam developed at NBS-JILA (Ertmer et al., 1985) and a novel pulsed beam source, Chu’s team at Bell Labs succeeded in loading cold sodium atoms into an optical molasses (Chu et al., 1985). They observed the expected long “lifetime” (the time required for the atoms to diffuse out of the laser beams) of the molasses, and they developed a method, now called “release-and-recapture,” for measuring the temperature of the atoms. The method is illustrated in Fig. 13. First, the atoms are captured and stored in the molasses, where for short periods of time they are essentially immobile due to the strong damping of atomic motion [Fig. 13(a)]. Then, the molasses laser beams are switched off, allowing the atoms to move ballistically away from the region to which they had originally been viscously confined [Fig. 13(b)]. Finally the laser beams are again turned on, recapturing the atoms that remain in the intersection (molasses) region [Fig. 13(c)]. From the fraction of atoms remaining after various periods of ballistic expansion one can determine the velocity distribution and therefore the temperature of the atoms at the time of release. The measured temperature at Bell Labs was \( 240^{+200}_{-60} \mu \text{K} \). [Now, one would expect a much lower temperature; the high temperature observed in this experiment has since been ascribed to the presence of a stray magnetic field from an ion pump (Chu, 1997).] The large uncertainty is due to the sensitive dependence of the analysis on the size and density distribution of atoms in the molasses, but the result was satisfyingly consistent with the predicted Doppler cooling limit.

By the end of 1986, Phil Gould and Paul Lett had joined our group and we had achieved optical molasses in our laboratory at NBS, loading the molasses directly from a decelerated beam. [Today it is also routine to load atoms directly into a magneto-optical trap (MOT) (Raab et al., 1987) from an uncooled vapor (Cable et al., 1990; Monroe et al., 1990) and then into molasses.] We repeated the release-and-recapture temperature measurements, found them to be compatible with the reported measurements of the Bell Labs group, and we proceeded with other experiments. In particular, with Paul Julienne, Helen Thorsheim and John Wiener, we made a 2-focus laser trap and used it to perform the first measurements of a specific collision process (associative ionization) with laser cooled atoms (Gould et al., 1988).

[Earlier, Steve Chu and his colleagues had used optical...}

---

**FIG. 13. Release-and-recapture method for temperature measurement.**
molasses to load a single-focus laser trap—the first demonstration of an optical trap for atoms (Chu et al., 1986). In a sense, our collision experiment represented a sort of closure for me because it realized the two-focus trap proposed in Ashkin’s 1978 paper, the paper that had started me thinking about laser cooling and trapping. It also was an important starting point for our group, because it began a new and highly productive line of research into cold collisions, producing some truly surprising and important results (Lett et al., 1991; Lett et al., 1993; Ratliff et al., 1994; Lett et al., 1995; Walhout et al., 1995; Jones et al., 1996; Tiesinga et al., 1996). In another sense, though, that experiment was a detour from the road that was leading us to a new understanding of optical molasses and of how laser cooling worked.

SUB-DOPPLER LASER COOLING

During 1987 Gould, Lett and I investigated the behavior of optical molasses in more detail. Because the temperature was hard to measure and its measurement uncertainty was large, we concentrated instead on the molasses lifetime, the time for the atoms to diffuse out of the intersecting laser beams. We had calculated, on the basis of the Doppler cooling theory, how the lifetime would vary as a function of the laser frequency detuning and the laser intensity. We also calculated how the lifetime should change when we introduced a deliberate imbalance between the two beams of a counterpropagating pair. Now we wanted to compare experimental results with our calculations. The results took us somewhat by surprise.

Figure 14 shows our measurements (Lett et al., 1989) of the molasses lifetime as a function of laser frequency along with the predicted behavior according to the Doppler cooling theory. The 1-D theory did not quantitatively reproduce the observed 3-D diffusion times, but that was expected. The surprise was the qualitative differences: the experimental lifetime peaked at a laser detuning above 3 linewidths, while the theory predicted a peak below one linewidth. We did not know how to reconcile this difficulty, and the results for the drift induced by beam imbalance were also in strong disagreement with the Doppler theory. In our 1987 paper, we described our failed attempts to bring the Doppler cooling theory into agreement with our data and ended saying (Gould et al., 1987): “It remains to consider whether the multiple levels and sublevels of Na, multiple laser frequencies, or a consideration of the detailed motion of the atoms in 3-D can explain the surprising behavior of optical molasses.” This was pure guesswork, of course, but it turned out to have an element of truth, as we shall see below.

Having seen such a clear discrepancy between the Doppler cooling theory and the experimental results, with no resolution in sight, we, as experimentalists, decided to take more data. Paul Lett argued that we should measure the temperature again, this time as a function of the detuning, to see if it, too, would exhibit behavior different from that predicted by the theory. We felt, however, that the release-and-recapture method, given the large uncertainty associated with it in the past, would be unsuitable. Hal Metcalf suggested a different approach, illustrated in Fig. 15.

In this time-of-flight (TOF) method, the atoms are first captured by the optical molasses, then released by switching off the molasses laser beams. The atom cloud expands ballistically, according to the distribution of atomic velocities. When atoms encounter the probe laser beam, they fluoresce, and the time distribution of fluorescence gives the time-of-flight distribution for atoms arriving at the probe. From this the temperature can be deduced. Now, with a team that included Paul...
FIG. 16. The experimental TOF distribution (points) and the predicted distribution curves for 40 μK and 240 μK (the predicted lower limit of Doppler cooling). The band around the 40 μK curve reflects the uncertainty in the measurement of the geometry of the molasses and probe.

Lettt, Rich Watts, Chris Westbrook, Phil Gould, as well as Hal Metcalf and myself, we implemented the TOF temperature measurement. In our experiment, the probe was placed as close as 1 cm from the center of the molasses, which had a radius of about 4.5 mm. At the lowest expected temperature, the Doppler cooling limit of 240 μK for Na atoms, a significant fraction of the atoms would have been able to reach the probe, even with the probe above the molasses. For reasons of convenience, we did put the probe beam above the molasses, but we saw no fluorescence from atoms reaching the probe after the molasses was turned off. We spent a considerable time testing the detection system to be sure that everything was working properly. We deliberately “squirted” the atoms to the probe beam by heating them with a pair of laser beams in the horizontal plane, and verified that such heated atoms reached the probe and produced the expected time-of-flight signal.

Finally, we put the probe under the molasses. When we did, we immediately saw the TOF signals, but were reluctant to accept the conclusion that the atoms were colder than the Doppler cooling theory predicted, until we had completed a detailed modeling of the TOF signals. Figure 16 shows a typical TOF distribution for one of the colder observed temperatures, along with the model predictions. The conclusion was inescapable: Our atoms had a temperature of about 40 μK, much colder than the Doppler cooling limit of 240 μK. They had had insufficient kinetic energy to reach the probe when it was placed above the molasses. As clear as this was, we were apprehensive. The theory of the Doppler limit was simple and compelling. In the limit of low intensity, one could derive the Doppler limit with a few lines of calculations (see for example, Lett et al., 1989); the most complete theory for cooling a two-level atom (Gordon and Ashkin, 1980) did not predict a cooling limit any lower. Of course, everyone recognized that sodium was not a two-level atom, but it had seemed unlikely that it made any significant difference (our speculation in Gould et al., 1987, notwithstanding). At low laser intensity the temperature depends on the laser detuning and the line-width of the transition. Since the linewidth is identical for all possible transitions in the Na D2 manifold, and since the cooling transition [3S1/2 (F = 2) → 3P3/2 (F = 3)] was well separated from nearby transitions, and all the Zeeman levels were degenerate, it seemed reasonable that the multilevel structure was unimportant in determining the cooling limit.

As it turned out, this was completely wrong. At the time, however, the Doppler limit seemed to be on firm theoretical ground, and we were hesitant to claim that it was violated experimentally. Therefore, we sought to confirm our experimental results with other temperature measurement methods. One of these was to refine the “release-and-recapture” method described above. The large uncertainties in the earlier measurements (Chu et al., 1985) arose mainly from uncertainties in the size of the molasses and the recapture volume. We addressed that problem by sharply aperturing the molasses laser beams so the molasses and recapture volumes were well defined. We also found that it was essential to include the effect of gravity in the analysis (as we had done already for the TOF method). Because released atoms fall, the failure to recapture atoms could be interpreted as a higher temperature if gravity is not taken into account.

Another method was the “fountain” technique. Here we exploited our initial failure to observe a TOF signal with the probe above the molasses. By adjusting the height of the probe, we could measure how high the atoms could go before falling back under the influence of gravity. Essentially, this allowed us to measure the atoms’ kinetic energy in terms of their gravitational potential energy, a principle very different from the TOF method. Finally, we used the “shower” method. This determined how far the atoms spread in the horizontal direction as they fell following release from the molasses. For this, we measured the fluorescence from atoms reaching the horizontal probe laser beam at different positions along that beam. From this transverse position distribution, we could get the transverse velocity distribution and therefore the temperature.

(The detailed modeling of the signals expected from the various temperature measurement methods was an essential element in establishing that the atomic temperature was well below the Doppler limit. Rich Watts, who had come to us from Hal Metcalf’s lab and had done his doctoral dissertation with Carl Wieman, played a leading role in this modeling. Earlier, with Wieman, he had introduced the use of diode lasers in laser cooling. With Metcalf, he was the first to laser cool rubidium, the element with which Bose-Einstein condensation was first achieved. He was a pioneer of laser cooling and continued a distinguished scientific career at NIST after completing his postdoctoral studies in our group. Rich died in 1996 at the age of 39, and is greatly missed.)

While none of the additional methods proved to be as accurate as the TOF technique (which became a standard tool for studying laser cooling temperatures), each of them showed the temperature to be significantly below the Doppler limit. Sub-Doppler temperatures were
not the only surprising results we obtained. We also (as Paul Lett had originally suggested) measured the temperature as a function of the detuning from resonance of the molasses laser. Figure 17 shows the results, along with the prediction of the Doppler cooling theory. The dependence of the temperature on detuning is strikingly different from the Doppler theory prediction, and recalls the discrepancy evident in Fig. 14. Our preliminary study indicated that the temperature did not depend on the laser intensity [although later measurements (Lett et al., 1989; Phillips et al., 1989; Salomon et al., 1990) showed that the temperature actually had a linear dependence on intensity]. We observed that the temperature depended on the polarization of the molasses laser beams, and was highly sensitive to the ambient magnetic field. Changing the field by 0.2 mT increased the temperature from 40 μK to 120 μK when the laser was detuned 20 MHz from resonance [later experiments (Lett et al., 1989) showed even greater effects]. This field dependence was particularly surprising, considering that transitions were being Zeeman shifted on the order of 14 MHz/mT, so the Zeeman shifts were much less than either the detuning or the 10 MHz transition linewidth.

Armed with these remarkable results, in the early spring of 1988 we sent a draft of the paper (Lett et al., 1988) describing our measurements to a number of experimental and theoretical groups working on laser cooling. I also traveled to a few of the leading laser cooling labs to describe the experiments in person and discuss them. Many of our colleagues were skeptical, as well they might have been, considering how surprising the results were. In the laboratories of Claude Cohen-Tannoudji and of Steve Chu, however, the response was: “Let’s go into the lab and find out if it is true.” Indeed, they soon confirmed sub-Doppler temperatures with their own measurements and they began to work on an understanding of how such low temperatures could come about. What emerged from these studies was a new concept of how laser cooling works, an understanding that is quite different from the original Hänsch-Schawlow and Wineland-Dehmelt picture.

During the spring and summer of 1988 our group was in close contact with Jean Dalibard and Claude Cohen-Tannoudji as they worked out the new theory of laser cooling and we continued our experiments. Their thinking centered on the multilevel character of the sodium atom, since the derivation of the Doppler limit was rigorous for a two-level atom. The sensitivity of temperature to magnetic field and to laser polarization suggested that the Zeeman sublevels were important, and this proved to be the case. Steve Chu (now at Stanford) and his colleagues followed a similar course, but the physical image that Dalibard and Cohen-Tannoudji developed has dominated the thinking about multilevel laser cooling. It involves a combination of multilevel atoms, polarization gradients, light shifts and optical pumping. How these work together to produce laser cooling is illustrated in simple form in Fig. 18, but the reader should see the Nobel Lectures of Cohen-Tannoudji and Chu along with the more detailed papers (Dalibard and Cohen-Tannoudji, 1989; Ungar et al., 1989; Cohen-Tannoudji and Phillips, 1990; Cohen-Tannoudji, 1992).

Figure 18(a) shows a 1-D set of counterpropagating beams with equal intensity and orthogonal, linear polarizations. The interference of these beams produces a standing wave whose polarization varies on a subwavelength distance scale. At points in space where the linear polarizations of the two beams are in phase with each other, the resultant polarization is linear, with an axis that bisects the polarization axes of the two individual beams. Where the phases are in quadrature, the resultant polarization is circular and at other places the polarization is elliptical. An atom in such a standing wave experiences a fortunate combination of light shifts and optical pumping processes.

Because of the differing Clebsch-Gordan coefficients governing the strength of coupling between the various ground and excited sublevels of the atom, the light shifts of the different sublevels are different, and they change

![FIG. 17. Dependence of molasses temperature on laser detuning (points) compared to the prediction of Doppler cooling theory (curve). The different symbols represent different molasses-to-probe separations.](image1)

![FIG. 18. (a) Interfering, counterpropagating beams having orthogonal, linear polarizations create a polarization gradient. (b) The different Zeeman sublevels are shifted differently in light fields with different polarizations; optical pumping tends to put atomic population on the lowest energy level, but nonadiabatic motion results in “Sisyphus” cooling.](image2)
with polarization (and therefore with position). Figure 18(b) shows the sinusoidal variation of the ground-state energy levels (reflecting the varying light shifts or dipole forces) of a hypothetical $J_z = 1/2 \rightarrow J_z = 3/2$ atomic system. Now imagine an atom to be at rest at a place where the polarization is circular $\sigma^-$ as at $z = \lambda/8$ in Fig. 18(a). As the atom absorbs light with negative angular momentum and radiates back to the ground states, it will eventually be optically pumped into the $m_z = -1/2$ ground state, and simply cycle between this state and the excited $m_z = -3/2$ state. For low enough intensity and large enough detuning we can ignore the time the atom spends in the excited state and consider only the motion of the atom on the ground state potential. In the $m_z = -1/2$ state, the atom is in the lower energy level at $z = \lambda/8$, as shown in Fig. 18(b). As the atom moves, it climbs the potential hill of the $m_z = -1/2$ state, but as it nears the top of the hill at $z = 3\lambda/8$, the polarization of the light becomes $\sigma^+$ and the optical pumping process tends to excite the atom in such a way that it decays to the $m_z = +1/2$ state. In the $m_z = +1/2$ state, the atom is now again at the bottom of a hill, and it again must climb, losing kinetic energy, as it moves. The continual climbing of hills recalls the Greek myth of Sisyphus, so this process, by which the atom rapidly slows down while passing through the polarization gradient, is called Sisyphus cooling. Dalibard and Cohen-Tannoudji (1985) had already described another kind of Sisyphus cooling, for two-level atoms, so the mechanism and the name were already familiar. In both kinds of Sisyphus cooling, the radiated photons, in comparison with the absorbed photons, have an excess energy equal to the light shift. By contrast, in Doppler cooling, the energy excess comes from the Doppler shift.

The details of this theory were still being worked out in the summer of 1988, the time of the International Conference on Atomic Physics, held that year in Paris. The sessions included talks about the experiments on sub-Doppler cooling and the new ideas to explain them. Beyond that, I had lively discussions with Dalibard and Cohen-Tannoudji about the new theory. One insight that emerged from those discussions was an understanding of why we had observed such high sensitivity of temperature to magnetic field: It was not the size of the Zeeman shift compared to the linewidth or the detuning that was important. Rather, when the Zeeman shift was comparable to the much smaller ($\approx 1$ MHz) light shifts and optical pumping rates, the cooling mechanism, which depended on these phenomena, would be disturbed. We now suggested a crucial test: the effect of the magnetic field should be reduced if the light intensity were higher. From Paris, I telephoned back to the lab in Gaithersburg and urged my colleagues to perform the appropriate measurements.

The results were as we had hoped. Figure 19 shows temperature as a function of magnetic field for two different light intensities. At magnetic fields greater than 100 $\mu$T (1 gauss), the temperature was lower for higher light intensity, a reversal of the usual linear dependence of temperature and intensity (Lett et al., 1989; Salomon et al., 1990). We considered this to be an important early confirmation of the qualitative correctness of the new theory, confirming the central role played by the light shift and the magnetic sublevels in the cooling mechanism. Joined by Steve Rolston and Carol Tanner we (Paul Lett, Rich Watts, Chris Westbrook and myself) carried out additional studies of the behavior of optical molasses, providing qualitative comparisons with the predictions of the new theory. Our 1989 paper (Lett et al., “Optical Molasses” summarized these results and contrasted the predictions of Doppler cooling with the new theory. Steve Chu’s group also published additional measurements at the same time (Weiss et al., 1989). Other, even more detailed measurements in Paris (Salomon et al., 1990) (where I was very privileged to spend the academic year of 1989–1990) left little doubt about the correctness of the new picture of laser cooling. In those experiments we cooled Cs atoms to 2.5 $\mu$K. It was a truly exciting time, when the developments in the theory and the experiments were pushing each other to better understanding and lower temperatures. Around this time, Jan Hall [whose pioneering work in chirp cooling (Ertmer et al., 1985) had done so much to launch the explosive activity a few years before] commented that being in the field of laser cooling was an experience akin to being in Paris at the time of the Impressionists. Figure 20 symbolizes the truth of that comment.

**OPTICAL LATTICES**

In 1989 we began a different kind of measurement on laser cooled atoms, a measurement that was to lead us to a new and highly fruitful field of research. We had always been a bit concerned that all of our temperature measurements gave us information about the velocity distribution of atoms after their release from the optical molasses and we wanted a way to measure the temperature in situ. Phil Gould suggested that we measure the spectrum of the light emitted from the atoms while they were being cooled. For continuous, single frequency ir-
radiation at low intensity and large detuning, most of the fluorescence light scattered from the atoms should be "elastically" scattered, rather than belonging to the "Mollow triplet" of high-intensity resonance fluorescence (Mollow, 1969). This elastically scattered light will be Doppler shifted by the moving atoms and its spectrum should show a Doppler broadening characteristic of the temperature of the atomic sample. The spectrum will also contain the frequency fluctuations of the laser itself, but these are relatively slow for a dye laser, so Gould suggested a heterodyne method of detection, where the fluorescent light is mixed on a photodiode with local oscillator light derived from the molasses laser, producing a beat signal that is free of the laser frequency fluctuations.

The experiment was not easy, and it worked mainly because of the skill and perseverance of Chris Westbrook. An example of the surprising spectrum we obtained (Westbrook et al., 1990) is shown in Fig. 21. The broad pedestal corresponded well to what we expected from the time-of-flight temperature measurement on a similar optical molasses, but the narrow central peak was a puzzle. After rejecting such wild possibilities as the achievement of Bose-Einstein condensation (Fig. 21 looks remarkably similar to velocity distributions in partially Bose-condensed atomic gases) we realized that the answer was quite simple: we were seeing line-narrowing from the Lamb-Dicke effect (Dicke, 1953) of atoms localized to less than a wavelength of light.

Atoms were being trapped by the dipole force in periodically spaced potential wells like those of Fig. 18(b). We knew from both theory and experiments that the thermal energy of the atoms was less than the light shifts producing the potential wells, so it was quite reasonable...
that the atoms should be trapped. Confined within a region much less than a wavelength of light, the emitted spectrum shows a suppression of the Doppler width, the Lamb-Dicke effect, which is equivalent to the Mössbauer effect. This measurement (Westbrook et al., 1990) marked the start of our interest in what are now called optical lattices: spatially periodic patterns of light-shift-induced potential wells in which atoms are trapped and well localized. It also represents a realization of the 1968 proposal of Letokhov to reduce the Doppler width by trapping atoms in a standing wave.

Joined by Poul Jessen, who was doing his Ph.D. research in our lab, we refined the heterodyne technique and measured the spectrum of Rb atoms in a 1-D laser field like that of Fig. 18(a). Figure 22 shows the results (Jessen et al., 1992), which display well-resolved sidebands around a central, elastic peak. The sidebands are separated from the elastic peak by the frequency of vibration of atoms in the 1-D potential wells. The sideband spectrum can be interpreted as spontaneous Raman scattering, both Stokes and anti-Stokes, involving transitions that begin on a given quantized vibrational level for an atom bound in the optical potential and end on a higher vibrational level (the lower sideband), the same level (elastic peak) or a lower level (the higher sideband). We did not see sidebands in the earlier experiment in a 3-D, six-beam optical molasses (Westbrook et al., 1990) at least in part because of the lack of phase stability among the laser beams (Westbrook et al., 1993). We have seen well-resolved sidebands in a 3-D, four-beam lattice (Gatzke et al., 1997).

The spectrum of Fig. 22 gives much information about the trapping of atoms in the potential wells. The ratio of sideband intensity to elastic peak intensity gives the degree of localization, the ratio of the two sideband intensities gives the temperature, and the spacing of the sidebands gives the potential well depth. Similar, but in many respects complementary, information can be obtained from the absorption spectrum of such an optical lattice, as illustrated by the experiments performed earlier in Paris (Verkerk et al., 1992). The spectrum of Fig. 22 can be calculated from first principles (Marte et al., 1993) and the comparison of the experimental and theoretical spectra shown provides one of the most detailed confirmations of our ability to predict theoretically the behavior of laser cooled atoms.

In our laboratory, we have continued our studies of optical lattices, using adiabatic expansion to achieve temperatures as low as 700 nK (Kastberg et al., 1995), applying Bragg scattering to study the dynamics of atomic motion (Birkl et al., 1995; Phillips, 1997; Raithel, Birkl, Kastberg et al., 1997; Raithel, Birkl, Phillips, and Rolston, 1997), and extending heterodyne spectral measurements to 3-D (Gatzke et al., 1997). The Paris group has also continued to perform a wide range of experiments on optical lattices (Louis et al., 1993; Meacher et al., 1994; Verkerk et al., 1994; Meacher et al., 1995), as have a number of other groups all over the world.

The optical lattice work has emphasized that a typical atom is quite well localized within its potential well, implying a physical picture rather different from the Sisyphus cooling of Fig. 18, where atoms move from one well to the next. Although numerical calculations give results in excellent agreement with experiment in the case of lattice-trapped atoms, a physical picture with the simplicity and power of the original Sisyphus picture has not yet emerged. Nevertheless, the simplicity of the experimental behavior makes one think that such a picture should exist and remains to be found. The work of Castin (1992) and Castin et al. (1994) may point the way to such an understanding.

CONCLUSION

I have told only a part of the story of laser cooling and trapping at NIST in Gaithersburg, and I have left out most of the work that has been done in other laboratories throughout the world. I have told this story from my
personal vantage point as an experimentalist in Gaithersburg, as I saw it unfold. The reader will get a much more complete picture by also reading the Nobel lectures of Steve Chu and Claude Cohen-Tannoudji. For the work in my lab, I have tried to follow the thread that leads from laser deceleration and cooling of atomic beams (Phillips and Metcalf, 1982; Prodan et al., 1982; Phillips and Prodan, 1984; Prodan et al., 1985) to magnetic trapping (Migdall et al., 1985), the discovery of sub-Doppler cooling (Lett et al., 1988; Lett et al., 1989), and the beginnings of optical lattice studies (Westbrook et al., 1990; Jessen et al., 1992). Topics such as later studies of lattices, led by Steve Rolston, and collisions of cold atoms, led by Paul Lett, have only been mentioned, and other areas such as the optical tweezer work (Mammen et al., 1996; Helmerson et al., 1997) led by Kris Helmerson have been left out completely.

The story of laser cooling and trapping is still rapidly unfolding, and one of the most active areas of progress is in applications. These include “practical” applications like atomic clocks, atom interferometers, atom lithography, and optical tweezers, as well as “scientific” applications such as collision studies, atomic parity non-conservation, and Bose-Einstein condensation (BEC). (The latter is a particularly beautiful and exciting outgrowth of laser cooling and trapping. Since the 1997 Nobel festivities, our laboratory has joined the growing number of groups having achieved BEC, as shown in Fig. 23.) Most of these applications were completely unanticipated when laser cooling started, and many would have been impossible without the unexpected occurrence of sub-Doppler cooling.

Laser cooling and trapping has from its beginnings been motivated by a blend of practical applications and basic curiosity. When I started doing laser cooling, I had firmly in mind that I wanted to make better atomic clocks. On the other hand, the discovery of sub-Doppler cooling came out of a desire to understand better the basic nature of the cooling process. Nevertheless, without sub-Doppler cooling, the present generation of atomic fountain clocks would not have been possible.

I hesitate to predict where the field of laser cooling and trapping will be even a few years from now. Such predictions have often been wrong in the past, and usually too pessimistic. But I firmly believe that progress, both in practical applications and in basic understanding, will be best achieved through research driven by both aims.

ACKNOWLEDGMENTS

I owe a great debt to all of the researchers in the many laboratories around the world who have contrib-
uted so much to the field of laser cooling and trapping of neutral atoms. Their friendly competition and generous sharing of understanding and insights has inspired me and educated me in an invaluable way. Very special thanks go to those researchers with whom I have been privileged to work here in Gaithersburg: to Hal Metcalf, who was part of the laser cooling experiments from the start, through most of the work described in this paper; to postdocs John Prodan, Alan Migdall, Phil Gould, Chris Westbrook, and Rich Watts, whose work led our group to the discovery of sub-Doppler cooling, and who moved on to distinguished careers elsewhere; to Paul Lett, Steve Rolston, and Kris Helmerson who also were pivotal figures in the development of laser cooling and trapping in Gaithersburg, who have formed the nucleus of the present Laser Cooling and Trapping Group (and who have graciously provided considerable help in the preparation of this manuscript); and to all the other postdocs, visitors and students who have so enriched our studies here. To all of these, I am thankful, not only for scientific riches but for shared friendship.

I know that I share with Claude Cohen-Tannoudji and with Steve Chu the firm belief that the 1997 Nobel Prize in Physics honors not only the three of us, but all those other researchers in this field who have made laser cooling and trapping such a rewarding and exciting subject.

I want to thank NIST for providing and sustaining the intellectual environment and the resources that have nurtured a new field of research and allowed it to grow from a few rudimentary ideas into a major branch of modern physics. I also thank the U.S. Office of Naval Research, which provided crucial support when I and my ideas were unproven, and which continues to provide invaluable support and encouragement.

There are many others, friends, family and teachers who have been of great importance. I thank especially my wife and daughters who have supported and encouraged me and provided that emotional and spiritual grounding that makes achievement worthwhile. Finally I thank God for providing such a wonderful and intriguing ground for learning some new things about it, and for allowing me to do so in the company of such good friends and colleagues.

REFERENCES


Chu, S., 1997, personal communication.


