Raman Cooling to High Phase Space Density

by

Zachary Vendeiro

B.A., University of California, Berkeley (2014)

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2021

Author .....
Department of Physics
August 19, 2021

Certified by .....
Vladan Vuletić
Lester Wolfe Professor of Physics
Thesis Supervisor

Accepted by .....
Deepto Chakrabarty
Associate Head of Physics

Raman Cooling to High Phase Space Density

by

Zachary Vendeiro

Submitted to the Department of Physics
on August 19, 2021, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Experiments on quantum degenerate gasses have become widespread since the first experimental observations of Bose-Einstein condensates (BECs) a few decades ago. Traditionally such experiments have relied heavily on evaporation to produce their quantum degenerate Bose or Fermi gases. Although evaporation has proven to be effective for many atomic species, it leads to the loss of many trapped atoms and is generally slow. The work presented in this thesis aims to improve upon the performance of evaporation by using Raman cooling. First it is shown that Raman cooling alone can produce (impure) BECs, cooling clouds all the way to condensation without evaporation. This is the first demonstration of direct laser cooling to a true three-dimensional BEC. Next it is shown that when evaporation and Raman cooling are both used in the same sequence, pure BECs can be produced in as little as 575 ms. This is the fastest BEC production time known to the author and it is achieved with a much simpler apparatus than other sub-second BEC experiments.

Raman sideband cooling in a 3D optical lattice may provide a way to prepare BECs even faster and potentially with very few collisions between atoms. Some preliminary work along these lines is also presented. Results from Raman sideband cooling in 1D and 2D optical lattices are also presented. Raman sideband cooling in a 1D lattice is shown to produce clouds with phase space densities of about 0.1, which is significantly larger than that achieved in previous work, but still shy of quantum degeneracy. Raman sideband cooling in a 2D lattice is shown to lead to non-thermalized clouds when the atom number and trap frequencies are sufficiently large. These unusual clouds are significantly hotter along the loosely-confined direction of the trap than in the tightly-confined direction, indicating the lack of thermalization in these effectively one-dimensional systems.

Along a separate line of research, the design of a Rydberg cavity quantum electrodynamics (cQED) experiment is discussed. The apparatus is designed to house two optical cavities and two imaging systems. One optical cavity is asymmetric in mirror transmission, and a method is demonstrated for measuring the offset between the trap and probe light standing waves very precisely using only frequency measurements. The imaging systems have moderately large numerical aperture (NA) and are

designed to be flexible, cost effective, relatively simple, cause only small aberrations, and require little in-vacuum alignment. The system will be capable of creating atom arrays in the cavities. These arrays could be used to implement quantum logic gates locally using Rydberg interactions between nearby atoms, and the cavity modes could be used as a bus to meditate gates between distant atoms.

Thesis Supervisor: Vladan Vuletić
Title: Lester Wolfe Professor of Physics

Acknowledgments

I have been fortunate enough to work with many great people throughout my Ph.D., and have be equally fortunate with who I've met outside of research as well. First and foremost I must thank my advisor Vladan Vuletić for his patience and guidance. Vladan gave me the time and freedom to dig deep into problems in the lab, allowing me to learn as much as possible while working on projects rather than pressuring me to finish them in minimal time. This approach helped me gain experimental expertise rather than just experimental experience, so I deeply appreciate his patience. His patience also often shone when answering my questions. I can assure the reader that I've asked Vladan plenty of dumb questions over the course of my time here, but you would never have been able to tell by Vladan's reaction. He would always provide a clear response in a kind tone, no matter what I asked.

In addition to Vladan, I also had the pleasure of working with many other colleagues on my experiment. Hao Zhang and Jiazhong Hu, the preceding grad students, got me started in AMO. I did not come from an AMO background, so this was no small task and I thank them for the time they invested in me. In particular I spent a lot of time designing a new cavity experiment apparatus with Hao and I learned much from him in that time. Brana Jelenkovic also visited our lab and guided me as I continued to work on that project after Hao's graduation. Jiazhong, despite being an experimentalist, was also an excellent theorist. Much of my cavity QED knowledge comes from reading his notes and derivations, and despite years of trying I was never able to find a single math error in any of his work. Later in my Ph.D. I had the pleasure of sharing the lab with Alban Urvoy and Josh Ramette. Both of them were excellent colleagues and we constantly had questions for each other and often ended up in lengthy and spirited debates. Our lab has also been a popular destination for undergraduates and visiting students from the high school to graduate level. In my time here I've had the pleasure of working with Nana Shumiya, Bojan Zlatković, Valentin Crépel, Milan Krstajić, Muqing Xu, Petar Bojović, Theo Lukin, Liam Parker, Haoyang Gao, Daniel Blais, Dawid Paszko, and Luke Stewart. Valentin

Crépel in particular worked with me on the new cavity apparatus for several months as we struggled to understand control theory well enough to lock our 1560 nm laser to a cavity with a 4 kHz linewidth, which he managed to do. I also look forward to the next generation of researchers coming to our lab. My time has overlapped with Alyssa Rudelis, Josiah Sinclair, and Michelle Chong, all of whom show a strong drive and enthusiasm for what comes next in our lab. I'm excited to see what they accomplish.

There are also colleagues from other labs in the CUA, and even groups outside of the CUA, to whom I owe a debt of gratitude. Whenever I wondered how something could be done better in lab, Boris Braverman was first person I'd go to. He is an experimental perfectionist who always had excellent answers to my technical questions. He and Jesse Amato-Grill also started the CUA wiki, which became an indispensable tool for our lab. I'd also like that thank Martin Zwierlein. A few years ago we asked to borrow his group's spare fiber amplifier on a temporary basis, yet it is still present in our lab. I was always worried that his group would need it back, which would cut short our laser cooling research. However he was incredibly generous and allowed us to hold on to the fiber amplifier for years. I cannot thank him enough for his generosity as my PhD would not have been possible without it.

In the last year or so our lab has undergone extensive software upgrades which have greatly increased our research capabilities. In particular the open-source Labscript and M-LOOP projects have changed the way we do research. The authors of both of these projects have been kind enough not only to share their code, but to provide support and answer questions as well, even patching bugs for us on occasion. On the Labscript side, Chris Billington, Phil Starkey, and Russ Anderson have taken the time to provide support and their reviews of my pull requests have made me a much better programmer. For M-LOOP, Harry Slatyer and Michael Hush took the time to meet with me and kindly tolerated my stream of pull requests to that project. I am deeply indebted to all of these colleagues.

Outside of the lab I was also involved in the Physics Friday socials. When I started, the Friday social crew consisted of Cody Burton, Keaton Burns, and Darius

Bunandar. As time moved on, Neel Kabadi, Sameer Abraham, Michael DeMarco, and Dimitra Pefkou joined. We've made many a beer run together and have spent plenty of time reorganizing the chairs and tables in the Pappalardo room. Many of my favorite memories from grad school came from those socials, and I'd like to thank everyone involved in them for making them so enjoyable.

I'd also like to thank the other friends I made in grad school. During my first year I could often be found searching for the local bars for the hoppiest IPAs with Sid Narayanan, Brandon Allen, Dan Abercrombie, and Cheko Cantu. As time wore on my ring of drinking buddies expanded to include Jesse, Ian Counts, Lawrence Cheuk, and Matt Nichols. Later on I could also be found losing video games with Cheko, Andrew Turner, and Dave Lingenbrink. I also am thankful for the time my pub trivia team Weakipedia, which included Sid, Cheko, Dan, Stephanie Brandt, Aaron Buikema, Nair Hid Ahmed, and my partner Cari Cesarotti. I also have fond memories of the physics softball team, run by my pal Michael Cantara. Every one of those people made my grad school experience much better in more ways than I can count.

Lastly I'd like to thank Cari. She has been a wonderful partner over the last few years, not to mention the most responsible roommate I've ever lived with. She provided constant guidance and support during the COVID-19 pandemic and showed extreme patience as I worked and slept long and unusual hours. She has made the last few years of my Ph.D. not only possible, but enjoyable.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Why Not Evaporate?

The field of atomic physics prides itself on its exquisite control over systems. Atomic clocks make the most precise measurements ever made by mankind [1, 2, 3], achieving 18 digits of precision. Additionally, trapped ion systems have demonstrated qubit gates [4, 5] with fidelities exceeding 99%. It is somewhat surprising then that despite the fact that Bose-Einstein condensates (BECs) have become common fare among atomic physics labs over the last 26 years since their first observations [6, 7], evaporation remains the standard way to produce quantum degenerate gases. Put another way, rather than exercising precise control over the motional state of atoms to cool them to condensation, labs around the world simply dump the hot atoms out of their trap to leave colder ones behind. Although this approach has opened the doors for extensive research [8, 9, 10, 11, 12, 13], it runs somewhat antithetical to the theme of pristine control in atomic physics.

The use of evaporation has some other downsides as well. In particular it can only be used for species with favorable collisional properties, it requires removal of most of the trapped cloud, and it is generally time-consuming due to the slow repopulation of the high-energy states that escape from the trap (thermalization through collisions). The first of those issues limits which species can be condensed, placing boundaries on what can be studied. The latter two of those issues have more practical conse-

quences. Typically the signal-to-noise ratio (SNR) of measurements increases with the number of atoms, so evaporation reduces the SNR by removing most of the atoms. Furthermore, experiments on degenerate gases generally use a shot-based approach in which gases are repeatedly prepared and then destructively measured. Therefore the rate at which research can be performed is limited by the rate at which clouds can be prepared, which is slowed by evaporation. So despite the widespread success of evaporation in the preparation of quantum degenerate gases, it is still worthwhile to investigate replacements for it or supplements to it.

1.2 Laser Cooling Temperature Limits

Laser cooling provides a promising alternative to evaporation as it doesn't fundamentally require the removal of atoms. However, there are several roadblocks that make laser cooling all the way to quantum degeneracy difficult, which is why evaporation has become so popular. Over time techniques for overcoming many of these roadblocks have been developed.

The earliest form of laser cooling was Doppler cooling [14, 15, 16], in which atoms are placed in two counter-propagating laser beams detuned red of the atomic transition. This configuration cools the atoms because they preferentially scatter photons from the beam propagating in the opposite direction to the atom's velocity, as that beam is blue-shifted towards the atomic resonance due to the Doppler effect. The scattering imparts an associated recoil momentum to the atom as the photon bounces off of it, slowing the atom down. This approach works well for capturing hot atoms, but it can only cool atoms down to a temperature referred to as the Doppler limit [15, 16]. This limit is simply a consequence of the fact that the atom must be moving fast enough for its Doppler shift to be approximately equal to its excited state linewidth in order for the atom to preferentially scatter from one beam or the other. The temperatures required for BEC are well below the Doppler limit, so Doppler cooling alone is not sufficient for laser cooling to BEC.

After the development of Doppler cooling came the discovery of polarization-

gradient cooling (PGC), which was initially discovered experimentally [17], then later explained theoretically [18]. PGC provided a route to sub-Doppler cooling, but had a limitation of its own: the recoil limit. This limit is set by the kinetic energy that a stationary atom would receive after emitting a single photon. It was not long before techniques to circumvent recoil limit as well were discovered. These approaches rely on dark states which do not scatter light and use tailored beam configurations to move atoms from a dark state to a bright state based on their velocity. By allowing only a narrow velocity class of atoms to remain in the dark state, atoms can be forced to repeatedly scatter photons until they happen to fall into that narrow velocity class, opening the possibility of sub-recoil cooling. The first observation of sub-recoil cooling was achieved via velocity-selective coherent population trapping (VSCPT) [19, 20]. Around the same period, Raman cooling was also proposed and demonstrated [21, 22].

At this point, the final fundamental limit to the temperatures achievable by laser cooling had been overcome. It may seem then that these approaches must inevitably be able to produce a BEC, however they still could not do so. The issue was that condensation requires not only low temperatures, but high atomic densities as well. The laser cooling schemes were developed by considering the interactions of one atom with the light field. The techniques then work for many atoms as long as they all act independently, which they do at low number densities. However, as the atomic cloud gets colder and denser, the atoms begin to see the effects of each other's presence and the cooling suffers. The simplest example of this is the issue of re-scattering. Generally these cooling schemes require a certain configuration for the light field and involve scattering light off of the atom. The scattered light's direction and polarization are uncontrolled and so it will typically heat a second atom if it is scattered again before escaping the cloud. Therefore re-scattering curtails the effectiveness of the cooling as the density increases. Along the same vein, light-assisted collisions, which will be discussed in Section 2.5.5 lead to excess atomic loss as the density increases, further limiting the efficacy of laser cooling at high densities.

1.3 Laser Cooling With High-Density Clouds

After each of the temperature limits on laser cooling had been overcome, some laser cooling research shifted focus towards overcoming the density limitations. The first major step in this direction was performed in Florian Schreck's group on strontium [23]. In that work, the authors used a dimple trap and took advantage of the narrow transition in strontium. By shining a transparency beam onto the dense region of atoms in the dimple, they could light shift those atoms enough so that they were transparent to the cooling light. Then only atoms in the low-density wings of the trap were cooled, acting as a cold bath for the atoms in the dimple. In some sense this approach is essentially evaporation with recycling; atoms with more energy than average can "evaporate" out of the dimple and the atoms remaining in the dimple rethermalize to a lower temperature. However instead of escaping the trap completely, the atoms that escape the dimple remain in the wings and are cooled and allowed to reenter the dimple. This work both pushed the boundaries of laser cooling and also showed that the difference between laser cooling and evaporation isn't as clear-cut and definitive as one might expect.

A few years after the work in Schreck's group, our group showed that 1D gases could be cooled at high densities thanks to their unusual properties which encourage atoms to avoid each other. Our group demonstrated, first in rubidium [24] then in cesium [25], that these clouds could be cooled all the way to quantum degeneracy. These demonstrations used Raman sideband cooling in an optical lattice, which is formed when two trap beams at the same wavelength interfere with each other to produce a spatially modulated potential with a periodicity on the order of the wavelength. A release-and-retrap compression sequence (pioneered in Dave Weiss' group [26]) was used to increase the occupation per lattice site by moving atoms to lattice sites nearer the center of the trap beams. The rubidium result also relied on the use of far-detuned optical pumping light to suppress light-assisted collisions [27], as will be discussed in Section 2.5.5. Although they were technically not BECs due to their 1D nature, the end result of each cooling sequence was an array of independent

1D degenerate bosonic gases with ∼10 to ∼100 atoms each and ∼100 to ∼1000 atoms in total.

Typical BEC experiments begin by creating a single 3D BEC with much more than ∼1000 atoms, so it would be easier for groups to transition to using laser cooling if it could prepare clouds more along those lines. Guided by this our lab began to investigate Raman cooling outside of the resolved-sideband limit. The research presented in this thesis is primarirly focused on this regime. We demonstrated that Raman cooling in a crossed optical dipole trap (cODT) with far-detuned optical pumping light could cool a cloud of rubidium atoms just past the threshold to BEC [28], which will be the topic of Chapter 4. The cooling was very efficient; starting with 100×10^3 atoms at PSD$\sim 10^{-3}$, 25×10^3 atoms remained as the threshold to BEC was crossed at PSD$\sim$1. This was the first instance of direct laser cooling to a true 3D BEC, where here "direct" means that the atoms in the overlap region of the cODT are cooled, in contrast to the work in Ref. [23].

Although our work in Ref. [28] showed that we could cross the threshold to BEC using only Raman cooling, the resulting cloud was very impure, with only ≈7% of the atoms occupying the BEC. In subsequent work, which is the topic of Chapter 5, we investigated what results are possible by combining these Raman cooling techniques with a brief final evaporation. It was found that we could create essentially pure BECs of $\approx 5 \times 10^3$ atoms in 575 ms, which is the fastest BEC preparation to date known to the author. Notably the apparatus required to implement this BEC preparation method is much simpler than other sub-second BEC machines [29, 30], although the resulting BECs are smaller. The construction of that sequence required multiple optimizations of the sequence's 42 parameters, which would be either very tedious or entirely impractical to do manually. However it was shown that Bayesian optimization [31], implemented by the open-source Python package M-LOOP [32, 33], could perform these optimizations automatically in only several hours. This makes implementation of such sequences much more practical, easing the difficulty in adopting this approach.

Raman sideband cooling in a 3D optical lattice could yield further improvements.

The cooling could be even faster due to the fast timescales of the system, and cooling to the 3D ground state of a lattice site could be performed without collisions. That could be particularly helpful for species which have unfavorable collisional properties. It may also be possible to release atoms from the 3D lattice to combine them into one BEC. Some preliminary efforts of Raman sideband cooling rubidium in a 3D lattice are presented in Ch. 7. Cooling to near the 3D lattice ground state in at least one direction is demonstrated, and difficulties of measuring the cloud temperature in that regime are also discussed. Preliminary attempts to cool multiple atoms in the same lattice site are also presented and it is shown that this particular goal is likely unfeasible in our system. Lastly some difficulties associated with adiabatically releasing atoms from a 3D lattice into a single combined trap are discussed.

Raman sideband cooling was also performed in 1D and 2D optical lattices, and these results are presented in Ch. 6. The work on Raman sideband cooling in a 2D lattice focused on trying to extend the results of Ref. [24] to clouds with more atoms. However it was seen that the cooling did not perform as well with more atoms, as we were not able to cool the tightly-confined direction to its ground state. Furthermore a strange and unexplained behavior was observed in which the temperature along the loose direction of the trap could be significantly larger than the temperature along the tightly-confined direction of the trap. Raman sideband cooling in a 1D lattice was shown to yield clouds with phase space densities $\sim$0.1, which is just shy of quantum degeneracy. Notably this represents more than a tenfold increase in phase space density achieved with this method compared to previous works [34, 35]. The final performance limited is in part by the slow thermalization between the indirectly-cooled loosely-confined directions of the trap with the directly-cooled tightly-confined direction of the trap. Although the insufficient thermalization in the 2D lattice is not understood, the slowed thermalization in the 1D lattice has a clear cause. In particular it is due to the minimum energy required for atoms to exchange kinetic energy along the loosely-confined direction of the trap for excitations along the tightly-confined direction of the trap, as will be discussed in Section 6.2.1.

Chapter 2

Raman Cooling Theory

This chapter explains the mechanisms behind Raman cooling and explains the different regimes in which it occurs. Its cooling abilities are discussed and equations are derived for the Raman cooling cycle rate and energy removed per cycle. Throughout this chapter I will often use the word "atom" to describe the target of the cooling, however the techniques are applicable to molecules, ions, and any other species with discrete internal structure and external motional degrees of freedom.

2.1 Raman Cooling as Two-Photon Doppler Cooling

Raman cooling has many subtleties and can operate in different regimes, which can make it seem complicated. However, at its heart Raman cooling is essentially Doppler cooling but with the single-photon transitions replaced with two-photon transitions. This section explains Raman cooling through that perspective with emphasis on conceptual understanding rather than mathematical models, which will instead be discussed later in this chapter. The analogy is straight-forward, but it is still helpful to carefully review Doppler cooling. Even readers who are very familiar with Doppler cooling are encouraged to read this description of it; it is written in a way to make the connection to Raman cooling as clear as possible.

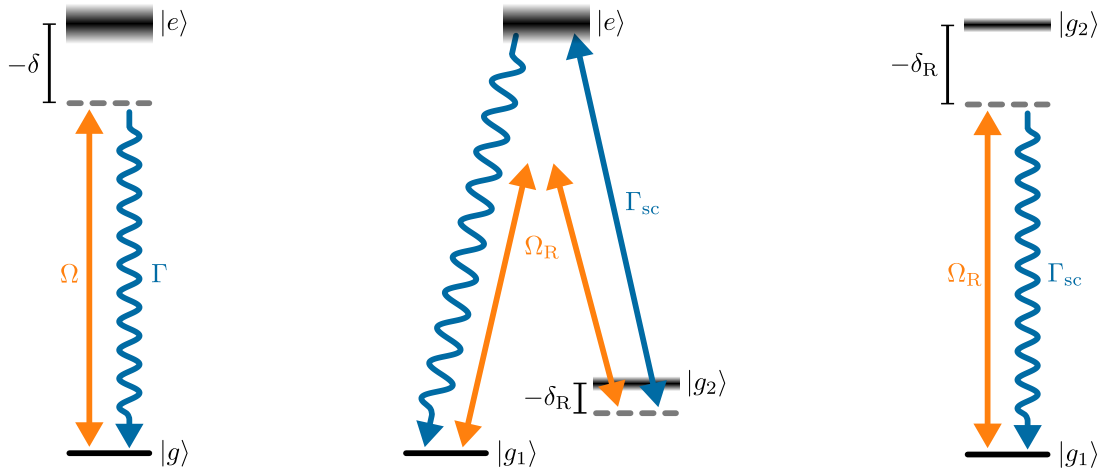
2.1.1 Doppler Cooling

Doppler cooling can be considered a two-step process and here we'll consider how it operates for a simple two-level atom. In Doppler cooling, an atom is illuminated with red-detuned light, but the detuning of the light in the atom's reference frame is shifted due to the Doppler effect. An atom moving against the beam see it blue-shifted, bringing it closer to resonance while an atom moving with the beam sees it red-shifted further from resonance. This effect means that the probability that the atom absorbs a photon from the beam and receives the associated recoil momentum kick depends on the velocity of the atom. In other words, the first step of Doppler cooling is to perform a velocity-selective transition from the atom's initial state $|g\rangle$ to another state $|e\rangle$ while simultaneously changing its external state by imparting a recoil momentum kick $\hbar k$. Here $\hbar$ is the reduced Planck constant, $\mathbf{k}$ is the wavevector of the light, and $k = |\mathbf{k}|$.

The velocity class for which this transition is resonant is the one for which the Doppler shift $\delta_{\mathrm{D}} = -\mathbf{k} \cdot \mathbf{v}$ cancels the detuning δ of the light from resonance in the lab frame such that the detuning in the atom's frame $\delta_{\mathrm{a}} = \delta + \delta_{\mathrm{D}}$ is zero. Here $\mathbf{v}$ is the velocity of the atom, the $\delta = \omega_{\mathrm{L}} - \omega_0$ is the laser's angular frequency detuning which is defined to be positive for blue detuning and negative for red detuning, ω_{L} is the angular frequency of the laser light, and ω_0 is the transition's resonant angular frequency. Because the atom's velocity is dotted with the light's wavevector, the detuning from resonance (including the Doppler shift) only depends on the component of the atom's velocity along $\mathbf{k}$, which we'll refer to as v. The excitation profile, i.e. the excitation rate as a function of v, is then a Lorentzian centered around the resonant velocity $v_0 = \delta/k$. The resonant velocity class can be tuned by adjusting the laser frequency or by adjusting the energy splitting between $|g\rangle$ and $|e\rangle$. The minimum width of the resonance is set by the linewidth of the excited state Γ, and can be larger due to broadening mechanisms such as power broadening. The rate at which an atom undergoes this transition is set by the Rabi frequency Ω of the driving beam, as well as the detuning, including the Doppler shift.

Once the atom has undergone the velocity-selective transition, it has two possible paths forward. One possibility is that the atom undergoes stimulated emission, putting the absorbed photon back into the driving beam. This is because the $|g\rangle$ and $|e\rangle$ states form a driven two-level system so the atom can Rabi flop between those two states at Rabi frequency Ω; in other words it is a reversible transition. Stimulated emission back into the driving beam provides a momentum kick $-\hbar k$ and returns the atom to $|g\rangle$, completely undoing the effects of the velocity-selective transition. Clearly this doesn't lead to cooling. Fortunately the other possible path forward can lead to cooling. The second option is that the atom spontaneously emits a photon in a random direction, transitioning back to $|g\rangle$ and again receiving a momentum kick of magnitude $|\hbar k|$ but in a random direction. This irreversibly returns the atom to $|g\rangle$ in a way that does lead to some undesirable recoil heating, but doesn't precisely undo the cooling achieved by the velocity-selective step. The return to the initial state is velocity-insensitive since the spontaneous emission occurs regardless of the atom's velocity. Once the atom is back in $|g\rangle$, a new iteration of the cooling process can begin.

The two steps of Doppler cooling have different abstract roles in the cooling process. The first (velocity-sensitive) step removes kinetic energy from the system by providing a cooling momentum kick. However, that does not reduce the entropy of the system because the atom's internal state has changed. It is the second (velocity-insensitive) step which removes entropy by returning the atom to its initial internal state. The entropy leaves the system in the form of the scattered photon which leaves in a random direction and with a random Doppler shifted frequency. The second step is also irreversible, which gives directionality to the cooling. In other words the irreversibility allows us to select cooling rather than heating. To see this, imagine that the reverse of spontaneous emission were possible, i.e. an atom in $|g\rangle$ could transition to $|e\rangle$ without absorbing a photon from the cooling beam. If that were to occur, the atom could then undergo stimulated emission into the cooling beam and receive a heating recoil kick. In other words, that would make it possible for the atom to undergo the Doppler cooling process backwards and be heated in the process. The



(a) (b) (c)

Figure 2-1: (a-b) Level diagrams for Doppler cooling and Raman cooling. The velocity-selective first step of each process is marked in orange and the velocity-insensitive second step is marked in blue. Arrows with heads on both ends represent reversible transitions while arrows with only one head represent irreversible transitions. The state $|e\rangle$ is broadened by its intrinsic linewidth Γ, and $|g_2\rangle$ has no intrinsic linewidth but is somewhat broadened by Γ_{sc} as will be discussed in Section 2.2. (c) Simplified level diagram for Raman cooling which depicts two-photon transitions as if they were single-photon transitions. This sweeps some details under the rug but more clearly shows the connection between Raman cooling and Doppler cooling. Notably (c) looks identical to (a), but with different labels.

fact that the reverse of spontaneous emission does not occur [1] makes that reverse process impossible, which allows us to only drive the forward direction and thereby cool the cloud.

2.1.2 Raman Cooling

Now that we've reviewed aspects Doppler cooling in detail, we are ready to make the connection to Raman cooling. In brief, Raman cooling follows the same two-

[1] Admittedly the counter-rotating term of the dipole Hamiltonian does allow ground state atoms in a vacuum to transition to the excited state and emit a photon in the process, but this is very energy-forbidden so we'll neglect it here. This reverse process can happen in an energy-conserving way though when a photon scattered by one atom is reabsorbed by another atom, which will be discussed in Section 2.5.4.

step process as Doppler cooling, but each one-photon transition is replaced with a two-photon transition. This requires one additional level, so in this subsection we'll consider a three-level atom with two stable ground states $|g_1\rangle$ and $|g_2\rangle$ and an excited state $|e\rangle$. Raman cooling starts with a reversible, velocity-selective, two-photon Raman transition from one state $|g_1\rangle$ to another state $|g_2\rangle$ which imparts a cooling momentum kick to the atom to remove kinetic energy from the system. Then an irreversible, velocity-insensitive, two-photon optical pumping transition is used to move atoms back from $|g_2\rangle$ to $|g_1\rangle$ which removes entropy from the system [2]. Both of these transitions occur by moving the atom through an intermediate excited state $|e\rangle$, but it is now the state $|g_2\rangle$ that plays the role that $|e\rangle$ did in Doppler cooling. This process is depicted in Fig. 2-1.

Going into a little more detail, the first step of the Raman cooling process is driven by two laser beams which have different wavevectors $\mathbf{k}_{\mathrm{R1}}$ and $\mathbf{k}_{\mathrm{R2}}$ where subscripts of 1 will be used to denote the beam driving the up-leg of the transition and subscripts of 2 will be used to denote the beam driving the down-leg. These beams are far-detuned ($\Delta_{\mathrm{R}} \gg \Gamma$) from the intermediate excited state $|e\rangle$ to minimize undesired scattering off of $|e\rangle$ in this part of the process. This constitutes a stimulated Raman transition, which is a two-photon transition in which the internal state of the atom changes and both legs of the transition are driven. The frequencies of these beams are chosen such that the difference between them is slightly different than the energy splitting between $|g_1\rangle$ and $|g_2\rangle$. In other words this two-photon transition from $|g_1\rangle$ to $|g_2\rangle$ is detuned by δ_{R} and this plays the role analogous to δ in Doppler cooling. Note that δ_{R} is not to be confused with the detuning Δ_{R} from the intermediate $|e\rangle$ state, which may seem more similar to δ at first glance but actually does *not* play the same role here.

A moving atom sees a different Doppler shift for each of the Raman coupling beams, so the Doppler effect shifts the two-photon transition by $\delta_{\mathrm{D}} = -\mathbf{v} \cdot \mathbf{k}_{\mathrm{R1}} + \mathbf{v} \cdot$

[2] Although not strictly necessary, $|g_1\rangle$ should ideally be dark to the optical pumping light. That avoids excess recoil heating which improves the cooling. Also, if $|g_1\rangle$ is not dark and can scatter into $|g_2\rangle$, then it is possible for the atom to undergo the Raman "cooling" process in reverse, leading to heating. A dark $|g_1\rangle$ state provides the best performance.

$\mathbf{k}_{R2} = -\mathbf{v} \cdot \Delta\mathbf{k}$ where $\Delta\mathbf{k} = \mathbf{k}_{R1} - \mathbf{k}_{R2}$. Note that the sign on shift for the down-leg is reversed because the photon is emitted into that beam rather than absorbed from it. To see that this is the case, note that blue-shifting the down-leg would push the final state dashed line further downwards in Fig. 2-1(b). The Doppler shift here is essentially identical to the Doppler shift in Doppler cooling, with the slight change that $\mathbf{k}$ is replaced by $\Delta\mathbf{k}$. Similarly, the cooling momentum kick imparted onto the atom during the Raman transition as it absorbs a photon from one beam and emits a photon into the other is $\hbar\mathbf{k}_{R1} - \hbar\mathbf{k}_{R2} = \hbar\Delta\mathbf{k}$, whereas the momentum imparted was $\hbar\mathbf{k}$ for Doppler cooling. The above implies that $\Delta\mathbf{k}$ plays the same role for Raman cooling that $\mathbf{k}$ does for Doppler cooling. As before, only the projection of $\mathbf{v}$ along the $\Delta\mathbf{k}$-direction affects the resonance condition and we'll refer to that component as v. Notably $|g_2\rangle$ is a stable state so it doesn't have an intrinsic linewidth, and so the linewidth of this transition is *not* limited by the excited state linewidth Γ. This opens the possibility for selective excitation of very narrow velocity classes, circumventing the Doppler limit.

As with Doppler cooling, the velocity-sensitive step is coherent and reversible so it removes kinetic energy but not entropy. After that first step, the atom can transition back to $|g_1\rangle$ from $|g_2\rangle$ using the Raman beams, but that would undo the cooling momentum kick in the process. A second irreversible path back to $|g_1\rangle$ from $|g_2\rangle$ is needed. This is achieved via optical pumping, which is a spontaneous two-photon Raman transition where the up-leg is driven but the down-leg occurs via spontaneous emission. For clarity, I'll generally refer to this step as "optical pumping" and the former step as "Raman coupling" even though both are Raman processes. The optical pumping step plays the same role for Raman cooling that the spontaneous emission does for Doppler cooling. The spontaneous emission during optical pumping makes it irreversible, and the entropy leaves the system in the form of the scattered photon. This returns the atom to the initial state $|g_1\rangle$ at the cost of some recoil heating, at which point the atom is ready for another cycle of cooling. The rate at which atoms are returned to $|g_2\rangle$ is proportional to Γ_{sc}, the scattering rate of the optical pumping beam. As will be discussed further in Section 2.2, this leads to some broadening of

Doppler Cooling	Raman Cooling
Absorption	Raman Coupling
Spontaneous Emission	Optical Pumping
$\|g\rangle$	$\|g_1\rangle$
$\|e\rangle$	$\|g_2\rangle$
Ω	Ω_{R}
δ	δ_{R}
$\mathbf{k}$	$\Delta\mathbf{k}$
Γ	Γ_{sc}

Table 2.1: The correspondence between aspects of Doppler cooling and Raman cooling. Raman cooling can be though of as Doppler cooling in which the one-photon transitions are replaced with two-photon transitions. When making that analogy, the aspects of Doppler cooling in the left column are replaced by the corresponding aspects of Raman cooling in the right column.

the state $|g_2\rangle$ by Γ_{sc}, which limits the velocity resolution of the Raman transition in the much same way that Γ limits the velocity resolution in Doppler cooling.

At this point the analogy between Doppler cooling and Raman cooling is complete. Both are two-step processes in which a reversible velocity-sensitive cooling transition is followed by an irreversible velocity-insensitive transition which returns the atom to its initial internal state where it can start another cooling cycle. For convenient reference, Table 2.1 summarizes the analogous components of the two processes.

2.2 Key Differences Between Doppler and Raman Cooling

Although Doppler cooling and Raman cooling are very similar in principle as discussed in Section 2.1, there are some key differences between the two which have significant effects on their results. This section points out those differences and discusses their consequences.

Recoil Heating — Due to the extra transitions involved in Raman cooling, it generally has worse recoil heating per cooling cycle than Doppler cooling. Firstly, in Doppler cooling the irreversible step simply involves one spontaneous decay, so

only one recoil energy is added on average in that step (in addition to a recoil energy of heating added during the absorption). In Raman cooling the irreversible step requires absorbing an optical pumping photon then reemitting it, which imparts two momentum kicks: one from absorption and one from emission. Furthermore, in Raman cooling the excited state $|e\rangle$ can decay to either $|g_1\rangle$ or $|g_2\rangle$ which implies that the atom isn't always successfully optically pumped to $|g_1\rangle$ whenever it scatters an optical pumping photon. Sometimes it will scatter a photon but decay back to $|g_2\rangle$. That imparts extra recoil heating without completing the Raman cooling cycle, which only finishes when the atom returns to $|g_1\rangle$. This aspect of Raman cooling is reflected in the effective recoil temperature, which is the subject of Section 2.5.1. That effective recoil temperature plays an analogous role to that of the normal recoil temperature, but it accounts for the branching ratio and other parameters involved in Raman cooling. Although recoil heating is worse for Raman cooling than for Doppler cooling, there are other differences in which Raman cooling is superior to Doppler cooling.

Tunable Scattering Rate and Velocity Resolution — A major difference between Doppler cooling and Raman cooling is that in Doppler cooling $|e\rangle$ has an intrinsic linewidth Γ due to its finite lifetime, while in Raman cooling $|g_2\rangle$ is stable and therefore has zero intrinsic linewidth. This was mentioned briefly in Section 2.1 but warrants further comment. The fact that $|e\rangle$ has width Γ is the reason that the Doppler limit exists; an atom must be moving fast enough for its Doppler shift to be $\sim\Gamma$ in order for it to be spectrally resolvable from a stationary atom. However, $|g_2\rangle$ has no intrinsic linewidth so there is no fundamental limit to how precisely the Raman transition can resolve an atom's velocity. This implies that the normal Doppler limit doesn't apply to Raman cooling, enabling it to reach lower temperatures.

In practice the velocity resolution of the Raman transition is broadened by the Raman coupling Rabi frequency Ω and the optical pumping scattering rate Γ_{sc}. This is analogous to how the velocity resolution in Doppler cooling is limited by the cooling laser intensity (which causes power broadening) and the excited state linewidth. Importantly, in Raman cooling Γ_{sc} can be varied at will by controlling the intensity

of the optical pumping beam, which would be akin to Doppler cooling on a transition with a tunable linewidth. That tunability is a powerful tool as it makes it possible to trade between fast cooling with large Γ_{sc} and high-velocity-resolution cooling with small Γ_{sc}.

This aspect of Raman cooling makes it very similar to "quenched" cooling, in which an atom is cooled via velocity-selective excitation on a narrow optical transition, then returned to the ground state by optically pumping out of the long-lived excited state [36, 37, 38]. In fact the primary difference between Raman cooling and quenched cooling is that Raman cooling's velocity-selective excitation is a two-photon transition between stable ground states while quenched cooling's velocity-selective excitation is a single-photon transition to an optically excited stable state. From this perspective, quenched cooling is halfway between Doppler cooling and Raman cooling, where only one of the single-photon transitions has been replaced with a multi-photon transition. However, quenched cooling is only available for very select atoms with narrow optical transitions.

The ability to turn off Γ_{sc} completely also becomes very helpful when performing pulsed Raman cooling as that makes it easier to create precise velocity-selection profiles. Pulsed cooling was not used for the work presented in this thesis, but for completeness it is briefly discussed in Sections 2.3.2 and 2.4.

Freedom of Detuning — In order to reach the lowest temperatures achievable by Doppler cooling, the frequency of the cooling beam must be set $\sim\Gamma$ to the red of the resonance. Adjusting the frequency away from the optimal value by $\sim\Gamma$ or more leads to higher temperatures, so there is very little room for adjustment. In Raman cooling there is much more freedom in choosing the frequencies of the beams involved. The beams used for the Raman coupling need to be detuned from resonance far enough to suppress scattering of Raman beam photons. However it is the difference in frequency between these beams, not the absolute frequency of the beams, that matters for selecting velocity classes. That means that the absolute frequency can be adjusted over a wide range and so detunings $\sim$10 GHz [39, 28, 40] to $\sim$100 THz [34, 24] have been used. The beam used for optical pumping can be set to any frequency.

In practice it is convenient to set the frequency on or near resonance to reduce the amount of power required for a given scattering rate. However the viable detunings are only limited by available laser power so the allowed frequency range is much larger than that for Doppler cooling.

The value of being able to choose laser frequencies at will may not be immediately apparent, but it is absolutely critical for the results presented in this thesis. This freedom enabled us to select laser frequencies in order to mitigate light-assisted collisions, which can prohibit laser cooling at high densities. This topic will be discussed further in Section 2.5.5.

2.3 Types of Raman Cooling

Section 2.1 explained Raman cooling as an extension to Doppler cooling which simply replaces each one-photon transition with a two-photon transition. Despite being so simple in this sense, Raman cooling can still often seem confusing. Some of this confusion occurs because Raman cooling can be performed in a few different ways. This section aims to explain some different forms of Raman cooling in an attempt to avoid that confusion.

2.3.1 Free-Space vs. Sideband-Resolved

When Raman cooling is performed on atoms in a harmonic trap, the trap frequency ω_t constitutes a timescale that can affect the system's dynamics. In the limit that the trap frequency ω_t is small compared to the optical pumping scattering rate or the Raman coupling Rabi frequency ($\omega_t \ll \Gamma_{sc}$ or $\omega_t \ll \Omega_R$), the trap levels are unresolved. In this limit the cooling works qualitatively the same way as it does without a trap, i.e. in free space. For this reason, it is referred to as the free-space limit. In the other extreme ($\omega_t \gg \Gamma_{sc}, \Omega_R$), the motional sidebands of the trap are well-resolved. In this limit Raman cooling is generally called Raman sideband cooling. Typically the phrase "Raman cooling" is used to refer specifically to Raman cooling in the free-space limit with unresolved-sideband while the phrase "Raman sideband

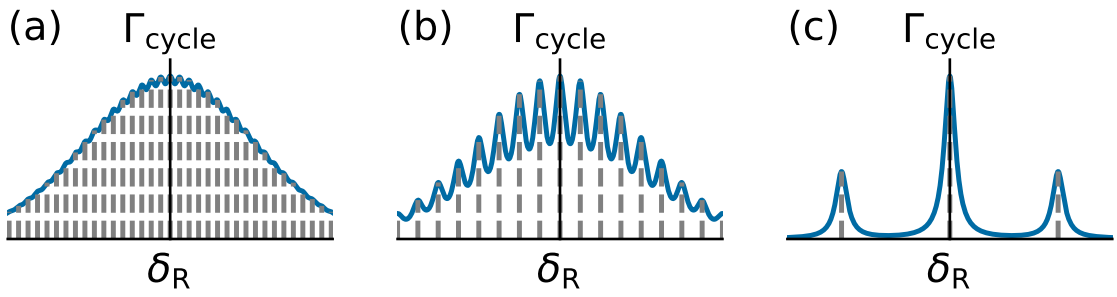


Figure 2-2: Raman cycle rate Γ_{cycle} (blue curves) as a function of Raman transition detuning δ_R in (a) the free-space limit, (b) the intermediate regime, and (c) the resolved-sideband limit for a thermal cloud. Vertical gray dashed lines mark sideband positions which are spaced by the trap frequency and each sideband has a Lorentzian excitation spectrum. The spectrum is nearly indistinguishable from a Gaussian in (a) where the trap frequency is small compared to the width of the Lorentzian. In (b) the trap frequency is larger and the sidebands are partially resolved. In (c) the trap frequency is larger still and the sidebands are fully resolved. This figure should only be interpreted qualitatively to show the transition between the regimes. The actual matrix elements of transitions and other effects are not accounted for.

cooling" is used when cooling in the resolved-sideband limit. Notably these limits are not unique to Raman cooling, they exist in Doppler cooling as well. The nomenclature there is similar: "Doppler cooling" typically refers to Doppler cooling in the free-space limit without resolved motional sidebands, while "Sideband cooling" refers to Doppler cooling in the resolved-sideband limit.

In either limit the cooling mechanism is the same, but there are some qualitative differences. For example, in the free-space limit the Raman cooling cycle rate for a thermal cloud monotonically falls off as the Raman transition detuning δ_R is increased due to the lower population at larger velocities. By contrast, in the sideband limit the cycle rate shows discrete peaks corresponding to when δ_R is approximately a multiple of ω_t. The lower population of high-energy states, as well as some other effects, causes the height of these peaks to fall off as δ_R is increase. This is depicted pictorially in Fig. 2-2. A last point worth noting is that, as will be discussed in Section 2.4, Raman sideband cooling is capable of reaching sub-recoil temperatures.

2.3.2 Pulsed vs. Continuous

The work presented in this thesis was performed with "continuous" Raman cooling. This means that the Raman coupling and optical pumping beams are left on simultaneously and continuously throughout the cooling. This is akin to how beams are typically left on continuously when performing Doppler cooling. However, previous Raman cooling work in the free-space limit [22, 41, 42, 43, 44] and some work in the resolved-sideband limit [45, 46] used pulsed Raman cooling, so that approach is briefly discussed in this section to provide additional context for the reader. Each pulsed cooling cycle involves turning the Raman coupling beams on then off following some carefully-chosen intensity-vs-time pulse profile while the optical pumping beam is turned off. These pulses can involve chirping the beam frequencies as well [43]. After the Raman coupling pulse has finished, the atoms are optically pumped back to $|g_1\rangle$.

As will be discussed in Section 2.4, pulsed Raman cooling is one method to achieve sub-recoil temperatures. In short, pulsing the Raman coupling makes it possible to have arbitrary velocity excitation profiles, i.e. curves of probability of an atom undergoing the Raman transition as a function of velocity. This makes it possible to create a very narrow (even compared to the recoil velocity) dark velocity class where atoms can accumulate over many cycles. The velocity-dependence of the excitation profile comes from the Doppler shift of the Raman beams as seen by moving atoms. Therefore the Doppler effect still plays a critical role in pulsed Raman cooling.

The pulsed cooling approach does not fit the analogy to Doppler cooling as well as continuous Raman cooling does, but it is possible and instructive to extend the analogy to imagine what the analogous approach for pulsed Raman cooling would look like. The analogous approach would be pulsed Doppler cooling in a two level system, which would involve pulsing on the Doppler cooling beam with a carefully-chosen intensity-vs-time pulse profile to selectively transfer atoms from $|g\rangle$ to $|e\rangle$ based on their velocity. Atoms would then decay back to $|g\rangle$ at which point a new cooling cycle could begin. This approach isn't very helpful though because unlike

the optical pumping out of $|g_2\rangle$ during Raman cooling, the spontaneous emission out of $|e\rangle$ in Doppler cooling cannot be turned off. This means that the duration of the pulse would have to be short compared to $1/\Gamma$ in order to avoid pulse errors due to spontaneous decay. Fourier broadening implies that the excitation profile would then be wide compared to Γ, meaning that velocity classes with Doppler shifts differing by $\lesssim \Gamma$ could not be resolved. Therefore pulsed Doppler cooling would not provide better velocity resolution than continuous Doppler cooling and so would not be able to reach lower temperatures. Hence why pulsed Doppler cooling isn't done. In contrast, during Raman cooling the optical pumping beam can be turned off, which allows an arbitrary amount of time for the Raman coupling pulse and enables it to have arbitrary velocity resolution. This is why pulsed cooling schemes have been used in Raman cooling but not for Doppler cooling.

2.4 Sub-Recoil Raman Cooling

There are some subtleties associated with how cold Raman cooling can cool an atom. Raman cooling is able to reach sub-recoil temperatures, but there are some requirements for that which are not necessarily obvious. This section discusses the temperatures achievable with continuous Raman cooling in the free-space limit, and then mentions methods for achieving sub-recoil cooling.

First it is worth noting that although $|g_2\rangle$ has zero intrinsic linewidth, the Raman transition in Raman cooling is generally broadened. There are two fundamental sources for this broadening: the optical pumping and the Raman coupling. These effects will be discussed further in Section 2.5.2, but are briefly summarized here. The optical pumping beam mixes $|g_2\rangle$ with $|e\rangle$, so the lower dressed state (which is mostly $|g_2\rangle$ but has a bit of $|e\rangle$ mixed in) inherits some of the linewidth from $|e\rangle$. This broadens the state by an amount set by Γ_{sc}. Along the same lines, the Raman coupling is driven at a Rabi frequency Ω_{R}. Atoms moving at velocities for which $|\delta_{\mathrm{a}}| \lesssim \Omega_{\mathrm{R}}$ can undergo the Raman transition even in the limit $\Gamma_{\mathrm{sc}} \to 0$. In other words, the velocity excitation profile is power broadened by Ω_{R}. Other broadening mechanisms are also

possible and can limit the velocity resolution of the Raman transition as well. For example, if the splitting between $|g_1\rangle$ and $|g_2\rangle$ depends on the magnetic field, then a time-dependent or spatially-varying magnetic field can lead to broadening.

However, Ω_R, Γ_{sc}, and technical broadening mechanisms can all be made arbitrarily small in principle and the velocity resolution of the Raman transition can be made arbitrarily narrow. It may seem that this implies that continuous Raman cooling in the free-space limit can cool to arbitrarily low temperatures. Unfortunately this is not quite the case, as can be seen from the analog with Doppler cooling. Although the Doppler limit is generally quoted as $\hbar\Gamma/(2k_B)$, that expression is only valid in the limit that $\Gamma \gg E_r/\hbar$. For narrow transitions where $\Gamma \ll E_r/\hbar$, Doppler cooling is instead limited to $E_r/2$ [47]. Similarly continuous Raman cooling can only cool a single atom to a temperature on the order of the recoil energy. This is true even if the parameters are set to make the transition narrow compared to $E_r/\hbar$. There is an intuitive explanation for this limit. If an atom is moving slowly compared to a recoil velocity, then it will likely be moving faster due to recoil heating after scattering one or more photons. That means that the atom would likely be heated overall by a "cooling" cycle. Therefore Doppler cooling cycles and Raman cooling cycles stop cooling on average once the temperature is below some value on the order of the recoil energy, no matter how fine the velocity resolution is.

Thus, continuous Raman cooling in the free-space limit can cool a single atom to about the recoil limit. A more quantitative expression for the minimum temperatures achievable this way, called the effective recoil limit, will be derived in Section 2.5.1. That said, Raman cooling can provide sub-recoil temperatures; it just needs some small additions beyond what has been described so far in order to do so. The first Raman cooling demonstrations achieved sub-recoil cooling by pulsing the Raman coupling light [22, 41, 42, 48, 43, 44], which was discussed in Section 2.3.2. Pulsing the light makes it possible to have tailored velocity excitation profiles to excite faster atoms while leaving atoms near zero-velocity unperturbed. When cooling to sub-recoil temperatures these "faster" atoms are often moving slower than the recoil velocity. After faster atoms undergo a Raman transition, they can be optically pumped back to

the initial state, randomizing their momentum due to the associated recoils. Although atoms moving slower than about one recoil velocity will be still heated on average by the Raman cooling cycle, this process can be repeated until population has built up in the states with nearly zero-velocity, thus achieving sub-recoil cooling. This process works well for individual atoms. However, at higher atomic densities, cold atoms may be knocked out of their velocity-selective dark state by other atoms still undergoing the cooling process. When that occurs, the cooling process for the previously-cold atom likely will need to be restarted. This is one of the difficulties of pulsed cooling at high atomic densities. The pulsed Raman cooling approach was not used in the work in this thesis, and is only mentioned here for the sake of completeness.

Another approach for sub-recoil Raman cooling is Raman sideband cooling, which is the Raman cooling analog of normal Doppler sideband cooling and was discussed in Section 2.3.1. Normal sideband cooling (typically just called "sideband cooling") is simply Doppler cooling in the resolved-sideband limit, and it offers a way to cool two-level species to below the standard Doppler limit. Resolving sidebands requires trapping the target species with a trap frequency large compared to the transition linewidth. That approach has been shown to work well in ion traps where large trap frequencies are readily attainable [49], but achieving such tight traps for charge-neutral species is more difficult. A workaround for neutral atoms, which works for ions as well, is to replace the Doppler cooling with Raman cooling. This is called Raman sideband cooling [50, 51, 34, 39], which is simply Raman cooling in the resolved-sideband limit as was discussed in Section 2.3.1. Raman transitions can be made arbitrarily narrow in principle, which reduces the trap frequencies required to resolve sidebands. This makes it much more practical to perform cooling with resolved sidebands on neutral species, particularly for species like alkalis which lack narrow optical transitions. By driving the red sideband, atoms can be transitioned to lower energy states until they accumulate in the ground state of the trap. Doing so can lead to sub-recoil temperatures [51].

The Raman cooling work we performed in the free-space limit, published in Refs [28, 40] and presented in Chapters 4 and 5, achieves sub-recoil temperatures

in a different way without pulsed cooling or resolved sidebands. Our cooling is performed in a trap at higher density than other similar work, which means that our atoms collide and thermalize. We take advantage of this by selectively cooling atoms with kinetic energies above the effective recoil limit. The atomic cloud then rethermalizes and repopulates the targeted velocity states so that the cloud can be cooled further. This approach allows sub-recoil cooling of an atomic ensemble, though the cooling becomes exponentially slower for temperatures well below the recoil energy. Sub-recoil cooling via thermalization will be discussed further in Section 2.5.3.

In summary, Raman cooling outperforms Doppler cooling in regards to the achievable temperatures. Simple continuous Raman cooling in the free-space limit (i.e. outside the resolved-sideband limit) enables sub-Doppler cooling. Furthermore, sub-recoil cooling is also possible using Raman cooling with the addition of resolved sidebands, pulsed cooling, or thermalization.

2.5 Raman Cooling Dynamics

This section discusses the dynamics of Raman cooling in the free-space limit. Formulae for the effective recoil temperature, Raman cooling cycle rate, and cooling rate are derived. Considerations due to loss and heating mechanisms are also presented.

2.5.1 Effective Recoil Temperature

In Doppler cooling, cycling transitions are often used so that the excited state $|e\rangle$ essentially always decays back to the initial state $|g\rangle$. However, Raman cooling requires at least two ground states coupled to $|e\rangle$. During the optical pumping step, an atom will sometimes decay to the desired initial state $|g_1\rangle$, but sometimes it will decay back to $|g_2\rangle$ instead. This is undesirable as it implies that Raman cooling generally requires more scattering events and more recoil heating than Doppler cooling in order to close the cooling cycle. A consequence of this is that it introduces a new characteristic temperature scale for the Raman cooling dynamics called the effective recoil temperature, which is derived in this subsection.

The effective recoil temperature is derived by calculating the expected value for the energy removed during a Raman cooling cycle, then finding the critical velocity where that value hits zero. Atoms moving at faster velocities will be cooled on average by a Raman cooling cycle, while atoms at lower velocities will be heated on average. The effective recoil temperature is then defined as the temperature at which the RMS speed of atoms is equal to that critical velocity.

Suppose an atom with mass m moving with at velocity $\mathbf{v}$ undergoes a Raman transition. Its change in kinetic energy is then:

$$\Delta K_{\text{Raman}} = \frac{(m\mathbf{v} + \hbar\Delta\mathbf{k})^2}{2m} - \frac{(m\mathbf{v})^2}{2m} \tag{2.1}$$

After expanding and using $\delta_{\text{D}} = -\mathbf{v} \cdot \Delta\mathbf{k}$, this can be rewritten as

$$\Delta K_{\text{Raman}} = -\hbar\delta_{\text{D}} + E_{\text{r}\Delta\mathbf{k}} \tag{2.2}$$

where $E_{\text{r}\Delta\mathbf{k}} = (\hbar\Delta\mathbf{k})^2/(2m)$ is the recoil energy associated with a momentum kick $\hbar\Delta\mathbf{k}$. In words, the Raman transition removes a Doppler shift of kinetic energy, but adds a $\Delta\mathbf{k}$-recoil of kinetic energy.

After the Raman transition the atom must be optically pumped to complete the cooling cycle. The recoil heating during optical pumping has some subtleties which will require some discussion. The number of optical pumping photon scatterings required on average to optically pump the atom depends on the branching ratio of $|e\rangle$ to $|g_1\rangle$, and we'll denote the probability that $|e\rangle$ decays to $|g_1\rangle$ as b. The probability that an atom is successfully optically pumped on the nth scattering after decaying to the wrong state during the first $(n-1)$ scatterings is $b(1-b)^{n-1}$. Therefore the average number of scatterings to successfully optically pump is

$$\langle n_{\text{scat}} \rangle = \sum_{n=1}^{\infty} nb(1-b)^{n-1} = \frac{1}{b} \tag{2.3}$$

Each scattering has two recoil kicks associated with it: one from absorbing a photon from the pump beam and one from spontaneously emitting a photon. The

spontaneously emitted photons leave in random uncorrelated directions, so these momentum kicks add in quadrature. This leads to E_r/b of heating on average, where $E_\mathrm{r} = (\hbar\mathbf{k}_\mathrm{pump})^2/(2m)$ is the recoil energy of a single optical pumping photon and $\mathbf{k}_\mathrm{pump}$ is the wavevector of the optical pumping beam. Assuming that the momentum kicks from the absorption add in quadrature (as will be discussed later on in this section, this isn't always true), then the absorptions contribute an additional E_r/b of heating on average, for a total of $2E_\mathrm{r}/b$.

The average total change in kinetic energy during a Raman cooling cycle is then

$$\Delta K_\mathrm{cycle} = -\hbar\delta_\mathrm{D} + E_{\mathrm{r}\Delta\mathbf{k}} + \frac{2E_\mathrm{r}}{b} \tag{2.4}$$

The critical velocity along the $\Delta\mathbf{k}$-direction at which the kinetic energy remains unchanged on average in a Raman cooling cycle can be calculated by setting $K_\mathrm{cycle} = 0$, yielding

$$v_{\Delta K=0} = \frac{E_{\mathrm{r}\Delta\mathbf{k}} + 2b^{-1}E_\mathrm{r}}{\hbar\,|\Delta\mathbf{k}|} \tag{2.5}$$

That velocity corresponds to a kinetic energy in the $\Delta\mathbf{k}$-direction of

$$K_{\Delta K=0} = \frac{1}{4}E_{\mathrm{r}\Delta\mathbf{k}}\left(1 + 2b^{-1}\frac{E_\mathrm{r}}{E_{\mathrm{r}\Delta\mathbf{k}}}\right)^2 \tag{2.6}$$

Finally, by defining $(1/2)k_\mathrm{B}T_\mathrm{r}^\mathrm{eff} = K_{\Delta K=0}$ we arrive at the formula for the effective recoil temperature:

$$\boxed{k_\mathrm{B}T_\mathrm{r}^\mathrm{eff} = \frac{1}{2}E_{\mathrm{r}\Delta\mathbf{k}}\left(1 + 2b^{-1}\frac{E_\mathrm{r}}{E_{\mathrm{r}\Delta\mathbf{k}}}\right)^2} \tag{2.7}$$

It is worth mentioning again here that Eqn. 2.7 is valid when the momentum kicks due to absorption from the optical pumping beam add in quadrature so that their recoil energies add, rather than adding linearly in which case their recoil momenta would add.

The physical interpretation of the effective recoil temperature is straightforward. At $T_\mathrm{r}^\mathrm{eff}$, an atom moving slower than the RMS thermal velocity wouldn't be cooled on average if it were to undergo a Raman cooling cycle; the recoil heating during the cycle

would outweigh the cooling on average. An atom can only be cooled on average if its velocity component along the $\Delta\mathbf{k}$-direction is larger than the RMS velocity of a cloud at the effective recoil temperature. This makes $T_{\mathrm{r}}^{\mathrm{eff}}$ a characteristic temperature at which Raman cooling dynamics change. As discussed in Section 2.4, simple continuous Raman cooling of a single atom can only cool to about this temperature but not beyond, akin to how Doppler cooling on a line narrower than the recoil energy can only cool to approximately the recoil energy. Some additions to that simple approach, such as those discussed in Section 2.4, are required for Raman cooling to cool atoms to below this temperature.

Notably, the effective recoil temperature is somewhat less fundamental than the recoil temperature and is not fixed for a given atom. For instance, $T_{\mathrm{r}}^{\mathrm{eff}}$ depends $\Delta\mathbf{k}$ through $E_{\mathrm{r}\Delta\mathbf{k}}$, Changing the wavelength of the beams driving the Raman transition, or even just changing the angle between those beams, will change the effective recoil temperature. Furthermore, real atoms have many different states, and different choices of which states to use for $|g_1\rangle$ and $|e\rangle$ will result in different branching ratios, and therefore different effective recoil temperatures.

At this point it is worth revisiting the question of how the recoil kicks from the absorption of optical pumping beam photons affect the atom. The heart of the complication is that kinetic energy scales as the square of momentum, and colinear momentum kicks add linearly while random momentum kicks add in quadrature. For example, suppose an atom starts at rest and receives two arbitrary momentum kicks $\hbar\mathbf{k}_a$ and $\hbar\mathbf{k}_b$ of the same magnitude with associated single-recoil energy E_{r}. The atom's increase in kinetic energy is then

$$\begin{aligned}\Delta K &= \frac{\left(\hbar\mathbf{k}_a + \hbar\mathbf{k}_b\right)^2}{2m} \\ &= 2E_{\mathrm{r}}\left(1 + \cos\theta_{ab}\right)\end{aligned} \tag{2.8}$$

Where θ_{ab} is the angle between $\mathbf{k}_a$ and $\mathbf{k}_b$. If $\mathbf{k}_a$ and $\mathbf{k}_b$ occur in random and uncorrelated directions, then the $\cos\theta_{ab}$ term averages out, leading to $\Delta K = 2E_{\mathrm{r}}$, i.e. one E_{r} per kick. On the other hand, if $\mathbf{k}_a$ and $\mathbf{k}_b$ are always colinear, then the cosine term

does not average out, leading to $\Delta K = 4E_\mathrm{r}$. The generalization to more momentum kicks is straightforward; after N kicks, the final kinetic energy is NE_r if the kick directions are properly randomized or N^2E_r if they are always the same direction. If the momentum kicks are partially, but not fully, correlated then the final kinetic energy will be somewhere in between.

For an initially non-stationary atom, correlations between a momentum kick's direction and the atom's initial momentum have a similar effect. If an atom has initial momentum $\mathbf{p}$ and receives a momentum kick $\hbar\mathbf{k}$, then its change in kinetic energy is

$$\Delta K = 2\frac{\mathbf{p}\cdot\hbar\mathbf{k}}{2m} + E_\mathrm{r} \tag{2.9}$$

The average increase in kinetic energy is just E_r for a thermal distribution because in that case $\mathbf{p}\cdot\hbar\mathbf{k}$ averages to zero. The Raman coupling generally selects a certain momentum class though, so the initial momentum of an atom undergoing optical pumping isn't generally fully randomized in Raman cooling. Therefore a correlation between the atom's initial momentum and the optical pumping beam's wavevector can exist, which can affect the expected change in kinetic energy during optical pumping.

From the above it is clear that multiple correlated momentum kicks lead to excess heating of atoms. Given that the momentum kicks due to absorption of photons from the optical pumping beam always occur in the same direction, the assumption that they add in quadrature used when deriving T_r may seem to be an odd choice. However, that assumption is valid if there is a process that randomizes the atom's momentum between scatterings. That situation can be described by repeatedly applying Eqn. 2.9, randomizing $\mathbf{p}$ each time, which causes its term to average out. In that case the momentum kicks again add in quadrature rather than linearly, so the recoil heating energy scales linearly with the number of scatterings.

There are at least two ways that the momentum can be randomized between scatterings. One possibility is that if the atom is part of an ensemble then thermalization can randomize the atom's momentum. Therefore $\Gamma_\mathrm{sc} \ll 1/\tau_\mathrm{c}$ is a sufficient condition for ensuring that optical pumping absorption momentum kicks add in quadrature,

where τ_c is the typical timescale for thermalization. Another possiblity which works even for single atoms is that the atom's momentum can be randomized by its trapping potential. For simplicity let us assume that the trap is harmonic with trap frequency ω_t, though this is not necessary. If the scattering rate is slow compared to the trap frequency $\Gamma_{sc} \ll \omega_t$, then the atom will oscillate in the trap between scatterings, randomizing its momentum. Notably $\Gamma_{sc} \ll \omega_t$ is the same criteria for reducing heating using festina lente effects as will be discussed in Section 2.5.4, however these are distinct effects [3] and that same criteria has appeared here for a different reason.

If neither of the criteria $\Gamma_{sc} \ll 2\pi/\tau_c$ or $\Gamma_{sc} \ll \omega_t$ are met, and no other momentum-randomizing mechanism is used, then Eqn. 2.7 is not valid and the actual effective recoil temperature will be larger due to the increased recoil heating during optical pumping. If the excess recoil heating can be calculated, then generalizing Eqn. 2.7 is straightforward. For example, in the limit that all of the absorption momentum kicks add lineraly but are not correlated with the atom's initial momentum (e.g. if the optical pumping beam is perpendicular to $\Delta\mathbf{k}$), then the average recoil heating due to absorption is

$$\langle \Delta K_{abs} \rangle = \sum_{n=1}^{\infty} n^2 E_r b(1-b)^{n-1} = \left(\frac{b^3 - 3b + 2}{b^2(1-b)} \right) E_r \tag{2.10}$$

That heating term can be inserted into Eqn. 2.4, and then the same steps can be taken to calculate the corrected T_r^{eff}. The same procedure can also be done to account for any other heating effects that occur at a rate proportional the the Raman cooling cycle rate.

Generally in the final cooling stages of the Raman cooling work presented in this thesis, where the performance of the Raman cooling is most critical, our experimental parameters satisfy $\Gamma_{sc} \ll \omega_t$. Satisfying that criteria is likely necessary for efficient Raman cooling at high PSD, both to stay in the festina lente regime and to avoid excess recoil heating. This is why Eqn. 2.7 was derived under the assumption of

[3] The festina lente regime is a multi-atom effect associated with rescattering light that was scattered off of another atom, while here we are dealing with a single atom effect associated with scattering multiple photons directly from the optical pumping beam.

randomized optical pumping beam absorption recoil kicks.

One final interesting point relevant to this section is that the branching ratio b for a given $|e\rangle \rightarrow |g_1\rangle$ transitions is not exactly a fixed parameter as one might expect. At first glance it would seem obvious that this would be set purely by internal matrix elements of the atom, but this is not the case when multiple atoms are present. As an extreme example, consider a fermionic atom is in $|e\rangle$ in a harmonic trap that already has an atom in $|g_1\rangle$ in each of its trap levels. Then the excited atom cannot decay to $|g_1\rangle$ due to the Pauli exclusion principle; every possible $|g_1\rangle$ state is already occupied. Thus the branching ratio to $|g_1\rangle$ is zero in that situation, even if the atom would have a nonzero branching ratio to $|g_1\rangle$ if no other atoms were around. Bosons have a similar effect, but their branching ratio into $|g_1\rangle$ is actually increased due to bosonic enhancement if there are already other atoms in that state. Imagining the same extreme scenario as above but with bosonic atoms, the matrix element for decay into $|g_1\rangle$ any trap state is enhanced by $\sqrt{2}$ by the presence of atom there. This comes from the fundamental symmetries of bosons and is reflected by the fact that for a bosonic creation operator $a^\dagger |1\rangle = \sqrt{2}\,|2\rangle$. Here the number in the ket here denotes the number of atoms in a given state, which must include both the internal and external degrees of freedom. Because the matrix element to any $|g_1\rangle$ final state is increased by $\sqrt{2}$, the decay rate into those states is doubled and so the branching ratio into $|g_1\rangle$ is increased. The increased branching ratio is beneficial because it decreases the average number of scatterings required to optically pump the atoms, which decreases the amount recoil heating per Raman cycle, which in turn reduces the effective recoil temperature. Along similar lines, if lower-energy trap states are more occupied with atoms in $|g_2\rangle$ than higher-energy trap states, then bosonic atoms are more likely to decay to those lower-energy states. That reduces the expected recoil heating from the last spontaneous emission in a cooling cycle.

These quantum statistics effects generally start to play a role when the expected occupation of quantum states approaches unity, which occurs when the phase space density approaches unity. Therefore these effects are most prominent for the cold, dense clouds at or beyond the threshold to degeneracy achieved near the end of

cooling. In Chapter 4 we will mention some results that seem to demonstrate this effect. Towards the end of the cooling sequence discussed there, the phase space density approaches one. Assuming the branching ratios of an isolated atom, the cooling parameters at that point in the sequence should lead to heating because the expected recoil heating would be larger than the amount of kinetic energy removed by the Raman transition on average. Despite that, cooling is still observed, which suggests that bosonic enhancement may be reducing the recoil heating per cycle. However, more careful calibrations and cross checks would need to be made before we could definitively claim to have observed these effects.

2.5.2 Cycle Rate

The cooling rate achieved via Raman cooling depends on how much energy is removed per cycle and how frequently an atom goes through a cooling cycle. This subsection derives the latter quantity for a simple three-level atom. The cycle rate is first derived in the limit $\Omega_R \gg \Gamma_{sc}$ as the math is simple in that limit, which makes the derivation intuitive and easy to understand. A more general form is subsequently provided based on the derivation for the cycle rate of quenched cooling [38].

In the limit $\Omega_R \gg \Gamma_{sc}$, $|g_1\rangle$ and $|g_2\rangle$ form a simple driven two-level system [4]. If an atom starts in $|g_1\rangle$ then after a time τ the probability that an atom has transitioned to $|g_2\rangle$ is [5]:

$$P(g_2) = \frac{\Omega_R^2}{\Omega_{Rabi}^2} \sin^2\left(\Omega_{Rabi}\tau/2\right) \tag{2.11}$$

where $\Omega_{Rabi} = \sqrt{\Omega_R^2 + \delta_a^2}$ is the generalized Rabi frequency of the Raman transition. If the Raman coupling is left on continuously, the atom Rabi flops back and forth many times before being optically pumped in the $\Omega \gg \Gamma_{sc}$ limit. Therefore it is

[4] This treatment of the problem is somewhat simplified for clarity and neglects some effects such as the light shifts of the levels due to the Raman coupling beams. See Ref [52] for a more complicated derivation which includes more effects at the cost of leading to some algebraically complicated results which are less intuitive. Note that Ref [52] does not include effects of broadening from the optical pumping beam and that their Rabi frequencies are defined to be half as large as the definition used in this thesis.

[5] See e.g. Ref. [53] Complement A_I Eqn.(35) with the substitutions $|V_{fi}| = \hbar\Omega_R/2$ and $(E_f - E_i) = -\delta_a$.

prudent to take the time average $\langle P_{g_2} \rangle = \Omega_{\mathrm{R}}^2/(2\Omega_{\mathrm{Rabi}}^2)$. By inserting some definitions and rearranging some terms this can be written as

$$\langle P_{g_2} \rangle = \left(\frac{1}{2}\right) \frac{1}{1 + \left(\frac{\delta_{\mathrm{R}} + |\Delta\mathbf{k}| v}{\Omega_{\mathrm{R}}}\right)^2} \tag{2.12}$$

This shows that the excitation profile in this limit is a Lorentzian in v (the atom's velocity along $\Delta\mathbf{k}$) centered at $v_0 = -\delta_{\mathrm{R}}/|\Delta\mathbf{k}|$ with full-width-half-max (FWHM) $2\Omega_{\mathrm{R}}/|\Delta\mathbf{k}|$.

While an atom is in $|g_2\rangle$, it scatters optical pumping photons at the rate Γ_{sc}. When an optical pumping photon is scattered, the atom has a chance to decay to the desired $|g_1\rangle$ state, but unfortunately it can also decay back to the $|g_2\rangle$ state. Therefore the optical pumping rate is actually smaller than Γ_{sc} by a factor of b, the branching ratio of $|e\rangle$ to $|g_1\rangle$. The cycle rate Γ_{cycle} can then be written as Γ_{sc} multiplied by the probability that the atom is in $|g_2\rangle$ where it can be optically pumped, times the probability that the scattering successfully pumps the atom into $|g_1\rangle$:

$$\Gamma_{\mathrm{cycle}} = b\left(\frac{\Gamma_{\mathrm{sc}}}{2}\right) \frac{\Omega_{\mathrm{R}}^2}{\Omega_{\mathrm{R}}^2 + (\delta_{\mathrm{R}} + |\Delta\mathbf{k}|\, v)^2} \tag{2.13}$$

This result is valid in the limit $\Omega_{\mathrm{R}} \gg \Gamma_{\mathrm{sc}}$.

The Raman cooling cycle rate can also be calculated in a more general case by following the derivation for the quenched cooling cycle rate in Ref. [38] with a few minor changes. The main differences are that the target state of the velocity-selective transition $|g_2\rangle$ has no intrinsic linewidth for Raman cooling and the Raman beams are not chirped to artificially broaden their excitation spectrum. The derivation is based on the master equation for the internal state of the atom:

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\frac{i}{\hbar}[\mathcal{H}, \rho] + \mathcal{C}(\rho) \tag{2.14}$$

Where $\mathcal{C}$ is the Lindblad superoperator

$$\mathcal{C}(\rho) = \sum_{j=1,2} C_j \rho C_j^\dagger - \frac{1}{2} C_j^\dagger C_j \rho - \frac{1}{2} \rho C_j^\dagger C_j \tag{2.15}$$

and C_j are the collapse operators representing the decays from $|e\rangle$. More specifically the collapse operators are $C_1 = \sqrt{b\Gamma}\,|g_1\rangle\langle e|$ and $C_2 = \sqrt{(1-b)\Gamma}\,|g_2\rangle\langle e|$. The Hamiltonian $\mathcal{H}$ simply consists of the internal state energies and the Rabi couplings between the states due to the driving lasers.

We'll now make the rotating wave approximation and the approximation that the excited state population is small, which has two requirements. First, the intermediate state detuning of the Raman coupling beams must be large compared to the Raman beam Rabi frequencies. This is generally already a requirement for efficient Raman cooling, since this is necessary to avoid scattering of the Raman beam photons. Secondly, this assumes that the optical pumping beam does not saturate the $|g_2\rangle \rightarrow |e\rangle$ transition. This is also already a requirement for good Raman cooling, otherwise $|g_2\rangle$ would be broadened to $\sim\Gamma$ which would reduce the Raman transition's velocity resolution to be comparable to that of normal Doppler cooling. Doing that would defeat the purpose of Raman cooling rather than Doppler cooling. In this approximation, the excited state population can be adiabatically eliminated, giving a two-level system $|g_1\rangle$ and $|g_2\rangle$ with an effective linewidth equal to Γ_{sc} (see Eqn. 8 of Ref. [38]). The final result can be written down based on Eqn. 12 of Ref. [38], inserting the effects of

the branching ratio b and the Doppler shift [6]:

$$\boxed{\Gamma_{\text{cycle}} = b\left(\frac{\Gamma_{\text{sc}}}{2}\right)\frac{\frac{\Omega_{\text{R}}^2}{2}}{\left(\delta_{\text{R}} + |\Delta\mathbf{k}|\, v\right)^2 + \frac{\Omega_{\text{R}}^2}{2} + \left(\frac{\Gamma_{\text{sc}}}{2}\right)^2}} \tag{2.16}$$

The above result is a Lorentzian centered at δ_{R} with full-width-half-max (FWHM, in angular frequency)

$$\text{FWHM} = 2\sqrt{\frac{\Omega_{\text{R}}^2}{2} + \left(\frac{\Gamma_{\text{sc}}}{2}\right)^2} \tag{2.17}$$

which shows the broadening due to the nonzero Ω_{R} and Γ_{sc}. In the limit of $\Gamma_{\text{sc}} \gg \Omega_{\text{R}}$ the FWHM approaches Γ_{sc}, and in the limit $\Omega_{\text{R}} \gg \Gamma_{\text{sc}}$ the FWHM approaches $\sqrt{2}\Omega_{\text{R}}$. The width in velocity can be calculated by dividing the width in angular frequency by $|\Delta\mathbf{k}|$.

2.5.3 Sub-Recoil Cooling Via Thermalization

As mentioned in Section 2.4, the sub-recoil Raman cooling work presented in this thesis achieved sub-recoil temperatures using thermalization. This subsection conceptually explains that approach then treats it mathematically.

Section 2.5.1 showed that atoms moving below a critical velocity would be heated on average if they were to undergo a Raman cooling cycle. This is because the amount of energy removed during the Raman transition is outweighed by the recoil heating induced by the optical pumping step. This aspect of the cooling is the mechanism behind the effective recoil temperature $T_{\text{r}}^{\text{eff}}$. However, in a thermal cloud the atoms

[6] Eagle-eyed readers may notice that Eqn. 2.16 does not quite approach Eqn. 2.13 in the limit $\Gamma_{\text{sc}} \to 0$, differing by a coefficient of 1/2 on Ω_{R} in the numerator. This difference comes from the fact that the time-averaged population in the second state of a driven two-level atom without scattering differs from the steady-state solution of the optical Bloch equation. In the full time-dependent solution to the optical Bloch equation, the atom oscillates between its states with an amplitude that decays with a timescale $\sim 1/\Gamma_{\text{sc}}$ from the scattering-free solution to the optical Bloch equation steady-state value. Because the atom is also optically pumped at this same timescale (at which point it enters a different velocity class, disrupting the two-level system evolution), the time-averaged population in the second state is actually somewhere between those values. In other words, the coefficient on Ω_{R} should actually be somewhere between 1/2 and 1 to account for the finite time that the atom evolves under the optical Bloch equation in a given velocity class. A derivation to determine exactly what that factor should be is beyond the scope of this text and would only lead to small changes in the dynamics.

do not all move at the RMS thermal velocity; their velocities follow a Maxwell-Boltzmann distribution. Therefore, even at low temperatures, there are some atoms in the high-energy tail of the distribution with velocities larger than the critical velocity. Those high-energy atoms can be cooled by putting them through a Raman cooling cycle, removing energy from the system. After those atoms are cooled, the system rethermalizes to a lower temperature and repopulates the high-energy states.

This thermalization approach is similar to evaporative cooling in a sense because it relies on repopulating the high-energy tails of a thermal distribution via thermalizing collisions, but there are a few key differences. First of all, ideally the high-energy atoms are cooled rather than released from the trap, so the atom loss can be greatly reduced compared to evaporation. This makes it somewhat akin to the strontium work from Ref. [23] mentioned in Section 1.3 in which atoms with large potential energy were cooled. However in our work we cool atoms with large kinetic energy rather than atoms with large potential energy. This means that there is no separate low-density bath of atoms; we cool the high-density cloud directly. The fact that we cool atoms with more kinetic energy than average is also qualitatively different than evaporative cooling which addresses atoms with much more potential energy than average.

There are other important consequences of the fact that atoms with more energy than average are cooled rather than removed from the trap. Although evaporation leads to a gain in PSD overall, it reduces the number of atoms in the trap. In order to be efficient and retain an acceptable number of atoms after a full evaporation sequence, the trap depth U must typically be set somewhere in the range of $\approx 5 k_{\mathrm{B}} T$ or more [7]. However the high-energy tails are repopulated slowly, becoming exponentially slower as $\eta = U/(k_{\mathrm{B}} T)$ is increased. This makes evaporation somewhat slow, and it becomes even slower if η is increased to make it more efficient. By contrast, Raman cooling doesn't explicitly require removing atoms from the cloud. That makes it possible to efficiently cool atoms with kinetic energies that are smaller multiples of $k_{\mathrm{B}} T$, which exponentially speeds up the cooling. This is particularly the case when the temperature is well above $T_{\mathrm{r}}^{\mathrm{eff}}$, but remains true until the temperature has dropped

to a factor of a few below $T_{\mathrm{r}}^{\mathrm{eff}}$. Thus the reduced loss of Raman cooling compared to evaporation cooling enables it to cool significantly faster. This speed is demonstrated in Ch. 4 where we present the fastest-prepared BEC to date. A related potential benefit of Raman cooling is associated with the fact that repopulating states with energy $E \gg k_{\mathrm{B}}T$ requires an average number of collisions that is exponential in $E/(k_{\mathrm{B}}T)$. Therefore it can require significantly fewer collisions to populate velocity classes with a few $k_{\mathrm{B}}T$ of kinetic energy than to populate states with several $k_{\mathrm{B}}T$ of potential energy. This implies that Raman cooling may be more favorable for cooling species with less-favorable collisional properties. It may even open the door to condensing new species that have yet to be condensed due to their poor collisional properties.

Given that waiting for collisions slows down the cooling and requires more favorable collisional properties, it may seem obvious that pulsed Raman cooling would be a more favorable way to achieve sub-recoil temperatures and create high-PSD clouds. However there are some subtleties that make thermalization-based sub-recoil cooling competitive. Pulsed cooling becomes less efficient below $T_{\mathrm{r}}^{\mathrm{eff}}$ because Raman cooling cycles heat the atoms on average. Generally many cycles have to be completed until an atom happens to randomly end up in a near-zero-velocity state, which requires scattering many photons. Unfortunately scattering photons can have negative consequences which become worse at high densities. As will be discussed in Section 2.5.5, each scattering leads to the possibility of a light-assisted collision which typically removes two atoms from the trap. Furthermore, heating due to rescattering (which will be discussed in Section 2.5.4) becomes more of an issue at higher densities, so scattering many photons off of one atom can lead to excess heating of other atoms. Lastly, it is also possible that when an atom ends up in a low-velocity state it may collide with a warmer atom, leaving the low-velocity state and starting the sub-recoil cooling process all over again. Therefore the many scatterings required for pulsed cooling can cause excess loss and heating when applied to dense clouds. On the other hand, sub-recoil cooling by thermalization requires many fewer scatterings because it targets atoms that have enough kinetic energy to be cooled on average, which reduces the negative effects associated with scattering photons. Furthermore,

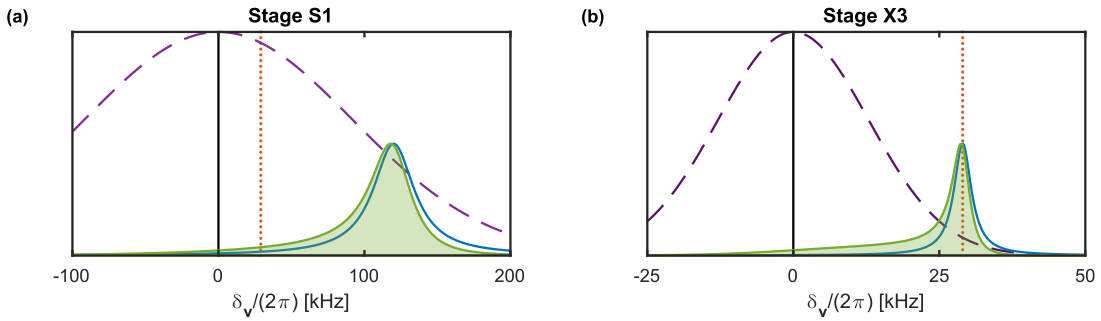


Figure 2-3: From Ref. [28]. Thermal velocity distribution (dashed purple line), time averaged Lorentzian Raman excitation profile (solid blue line), rescaled product of the thermal distribution and the excitation profile (shaded green region), and Doppler shift at the critical velocity $v_{\Delta K=0}$ (29 kHz, dotted red) for cooling sequence parameters in (a) stage S1 and (b) stage X3 of the cooling sequence presented in Chapter 4. On average, a cooling cycle will cool atoms with velocity greater than $v_{\Delta K=0}$, and heat atoms with velocity below $v_{\Delta K=0}$.

thermalizing collisions with other atoms are no longer detrimental; in fact they are necessary.

In the limit that the thermalization rate Γ_{therm} is large compared to the Raman cooling cycle rate Γ_{sc} at the resonant velocity class, much of the quantitative analysis of sub-recoil cooling via thermalization is straightforward. This is because the Raman cooling then doesn't affect the population of velocity classes. In the opposite limit, the Raman cooling would deplete velocity classes, making the results harder to calculate. For reference, it has been found that $\Gamma_{\text{therm}} \approx \Gamma_{\text{col}}/2.7$ [54, 55, 56, 57] where Γ_{col} is the collision rate (equal to the inverse of the collision time). We'll now derive the average energy removed per Raman cooling cycle $\langle \Delta K_{\text{cycle}} \rangle$ in the limit $\Gamma_{\text{therm}} \gg \Gamma_{\text{sc}}$ accounting for the distribution in atomic velocities for a cloud of atoms. As derived in Sections 2.5.1 and 2.5.2, the Raman cycle rate and the average energy removed in a Raman cycle are functions of the atom's velocity v along the $\Delta\mathbf{k}$-direction, or equivalently they are functions of the Raman transition Doppler shift δ_{D}. An atom can only be optically pumped to complete a Raman cycle while it is in $|g_2\rangle$, and the fraction of time that it is in $|g_2\rangle$ is given by $\Gamma_{\text{cycle}}/(b\Gamma_{\text{sc}})$. The average energy removed per cooling cycle $\langle \Delta K_{\text{cycle}} \rangle$ is found by multiplying the average energy removed in a Raman transition for a given velocity by the fraction of the time that an atom at that

velocity spends in $|g_2\rangle$, then averaging over all possible velocities while weighting by the velocity distribution:

$$\langle \Delta K_{\text{cycle}} \rangle = \int_v \Delta K_{\text{cycle}}(v) \frac{\Gamma_{\text{cycle}}(v)}{b\Gamma_{\text{sc}}} p(v) \mathrm{d}v \tag{2.18}$$

This approach is depicted in Fig. 2-3, which uses the actual values for $\Delta\mathbf{k}$, δ_{R}, etc. used in Ref. [28] and presented in Chapter 4. There the velocity distribution is assumed to be a Maxwell-Boltzmann distribution, which is accurate for thermal clouds, given by

$$p(v) = \sqrt{\frac{m}{2\pi k_{\text{B}} T}} \exp\left(-\frac{mv^2}{2k_{\text{B}}T}\right) \tag{2.19}$$

The excitation profile is a Lorentzian with width limited by the Raman coupling Rabi frequency Ω_{R}, so Eqn. 2.13 is used for $\Gamma_{\text{cycle}}(v)$. The optical pumping rate is slow compared to the trap frequencies in that work, so Eqn. 2.4 is used for $\Delta K_{\text{cycle}}(v)$.

Fig. 2-3(a) shows the parameters when the cloud's temperature is large compared to $T_{\text{r}}^{\text{eff}}$. The parameters are tuned to put the resonant velocity class for the Raman transition at approximately the RMS thermal velocity, which makes the cooling much faster than evaporation at that point in the sequence. Fig. 2-3(b) shows the parameters when the cloud's temperature is small compared to $T_{\text{r}}^{\text{eff}}$. The Raman coupling rate Ω_{R} is reduced to provide narrower velocity resolution and the resonant velocity is put very near the critical velocity. Interestingly, the expected average energy removed in the Raman transition for Fig. 2-3(b) is $h * 15$ kHz while the expected recoil heating is $h * 22$ kHz assuming the bare atomic branching ratio. This could indicate the effects of bosonic enhancement mentioned in Section 2.5.1, in which the branching ratio to $|g_1\rangle$ is increased by bosonic enhancement due to the presence of other atoms. Such bosonic enhancement would increase the probability that an atom decays to $|g_1\rangle$ after each optical pumping scattering, thereby reducing the average recoil heating in each Raman cooling cycle and making it possible to cool atoms with velocities below $v_{\Delta K=0}$. It is also worth pointing out that the typical velocity of an atom undergoing a Raman transition in Fig. 2-3(b), i.e. the mean of the shaded green curve, is still less than twice the typical thermal velocity. This corresponds to a kinetic energy less

than four times the thermal average, which is still smaller than the typical η used in evaporative cooling [7]. That indicates that the cooling rate is still competitive with efficient evaporation despite the fact that the temperature has dropped below $T_{\mathrm{r}}^{\mathrm{eff}}$. The integral over velocity does not have a nice closed form solution, but it is straightforward to integrate numerically.

It is also instructive to view the value of $\langle \Delta K_{\mathrm{cycle}} \rangle$ as a function of δ_{R} and Ω_{R} at different temperatures, which is shown in Fig. 2-4. This shows how the mean energy removed per cycle is affected by the Raman cooling control parameters at a given temperature. Notably many parameter values lead to heating rather than cooling, indicated by the white regions in the plots. At large temperatures, large detunings $\delta_{\mathrm{R}} \sim (2\pi)100$ kHz and large $\Omega_{\mathrm{R}} \sim (2\pi)75$ kHz are optimal. As the temperature is lowered, optimal cooling happens are smaller δ_{R}, due to the decreased population at high-velocities. The optimal Ω_{R} also decreases due to the need for narrower velocity resolution, and the average energy removed per cycle decreases as well. Going from 30 μK to 2.8 μK ($T_{\mathrm{r}}^{\mathrm{eff}} = 2.8$ μK here) corresponds to roughly a factor of $\approx$10 decrease in temperature and a factor of $\approx$20 decrease in predicted $\langle \Delta K_{\mathrm{cycle}} \rangle$ for optimal parameters. By comparison, going from 2.8 μK to 0.61 μK corresponds to a factor of $\approx$5 in temperature but $\approx$150 in predicted $\langle \Delta K_{\mathrm{cycle}} \rangle$ for optimal parameters. In fact, the predicted optimal $\langle \Delta K_{\mathrm{cycle}} \rangle$ is negligible at 0.61 μK. This is part of the reason why the Raman cooling slows down below $T_{\mathrm{r}}^{\mathrm{eff}}$. Note that the predicted results here do *not* include the effects of bosonic enhancement.

Part of the reason for the dramatic fall off of $\langle \Delta K_{\mathrm{cycle}} \rangle$ is associated with how Gaussians and Lorentzians scale in their tails. A Lorentzian falls off polynomially as $1/x^2$ while a Gaussian falls off exponentially as $\exp(-x^2)$, so Gaussians fall off much faster than Lorentzians. In order to cool on average, the cooling must preferentially target atoms moving faster than $v_{\Delta K=0}$. Naïvely it may seem that this simply requires setting $\delta_{\mathrm{R}} > |\Delta \mathbf{k}|\, v_{\Delta K=0}$, but this alone is not sufficient. When $T \ll T_{\mathrm{r}}^{\mathrm{eff}}$, the population of atoms with Doppler shifts $\approx \delta_{\mathrm{R}}$ is exponentially small while the rate at which atoms near zero-velocity are off-resonantly excited by the Raman transition is only polynomially small. This means that the typical velocity of an atom that undergoes a

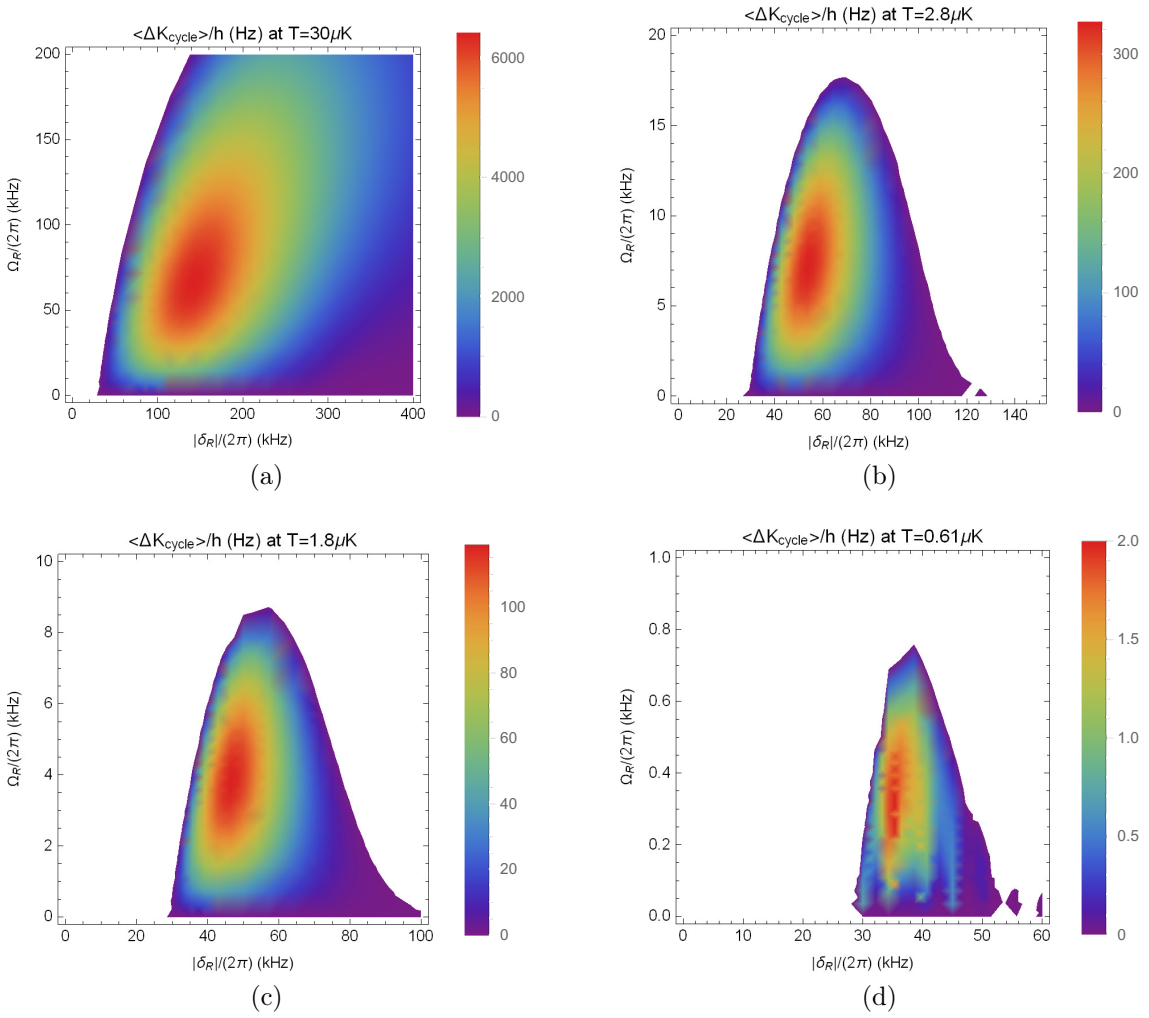


Figure 2-4: Mean energy removed per Raman cycle $\langle \Delta K_{\text{cycle}} \rangle$ as a function of δ_{R} and Ω_{R} at various temperatures corresponding to points in the cooling sequence presented in Chapter 4: (a) 30 μK, the initial temperature in that sequence, (b) 2.8 μK, the effective recoil temperature $T_{\text{r}}^{\text{eff}}$, (c) 1.8 μK, the temperature at the start of stage X3, and (d) 0.61 μK, the temperature at the threshold to BEC in that sequence. The values of $\Delta\mathbf{k}$, b, etc. are the same as used in the experimental results presented in Chapters 4 and 5. The Doppler shift at the critical velocity is $(2\pi)29$ kHz and the calculations do *not* include the effects of bosonic enhancement. Note that the axes and color scale ranges change between plots. White regions correspond to parameter values which produce heating on average. The strange patterns in (d) are due to issues with the numerics.

Raman transition is actually much smaller than δ_R because the effect of the Gaussian Maxwell-Boltzmann distribution outweighs the effects of the Lorentzian excitation profile when they are multiplied together. That can mean that most "cooling" cycles actually address low-velocity atoms and lead to heating. This effect leads to the long tail towards low velocities of the green curve showing the product of the Gaussian and Lorentzian in Fig. 2-3(b), and that tail causes some counterintuitive behavior.

Increasing $|\delta_R|$ moves the resonant velocity class to higher velocities which would seem to imply that the cooling would primarily address higher-velocity atoms and would remove more energy per cycle on average. However due to the weighting from the Maxwell-Boltzmann distribution, eventually increasing the detuning leads to exciting lower velocities on average and removing less energy per cycle on average. One way to think of this is that when $|\delta_R|$ is very large, then the Doppler shift of the vast majority of atoms is negligible compared to δ_R and those atoms see essentially the same detuning regardless of velocity. In that case nearly all atoms are essentially equally likely to undergo a Raman transition, so the low-velocity atoms are addressed about as much as most of the high-velocity atoms. This is why the mean energy removed per cycle decreases and even drops below zero as $|\delta_R|$ is increase at fixed Ω_R in Fig. 2-4. In order to overcome this, the width of the Lorentzian must be reduced so that its effect dominates the product with the Gaussian despite its weaker scaling. This means that Ω_R and Γ_{sc} must be set to low values not just to achieve sufficient velocity-resolution, but low enough so that the narrowness of the Lorentzian outweighs the exponential scaling of the Gaussian. For any finite width of the Gaussian, a sufficiently narrow Lorentzian exists which can selectively excite atoms moving faster than $v_{\Delta K=0}$ on average. This is why all of the plots in Fig. 2-4 have a region that gives some amount of cooling on average, and it is why that region has a tail towards large $|\delta_R|$ at small Ω_R. Although parameters that produce some amount of cooling on average in principle always exist, achieving a sufficiently narrow Lorentzian may require reducing Ω_R and Γ_{sc} to the point that the cooling becomes impractical.

The cooling rate P_{cool} is also straightforward to calculate in the limit $\Gamma_{therm} \gg \Gamma_{cycle}$. The form is fairly similar to that for $\langle \Delta K_{cycle} \rangle$. It can be calculated by

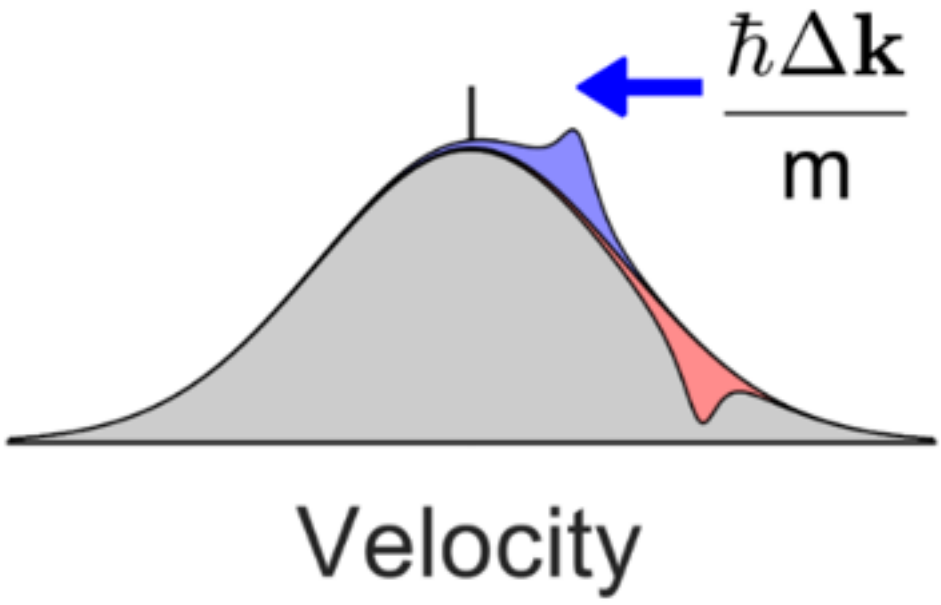


Figure 2-5: The quasi-steady-state velocity distribution of a cloud undergoing continuous Raman cooling in the limit of slow thermalization $\Gamma_{\text{therm}} \ll \Gamma_{\text{cycle}}$, taken from Ref. [28]. The red region shows where the velocity class population is depleted due to Raman cooling of atoms at that velocity. The blue region shows the excess of population accumulated where atoms end up after a Raman cooling cycle before the rethermalize with the rest of the cloud.

multiplying the average energy removed in a Raman transition for a given velocity by Γ_{cycle} for that velocity, then averaging over all possible velocities weighted by the velocity distribution:

$$\begin{aligned} P_{\text{cool}} &= \int_v \Delta K_{\text{cycle}}(v)\Gamma_{\text{cycle}}(v)p(v)\mathrm{d}v \\ &= b\Gamma_{\text{sc}} \left\langle \Delta K_{\text{cycle}} \right\rangle \end{aligned} \tag{2.20}$$

Again this integral has no nice closed form solution but can easily be integrated numerically.

The dynamics outside of the limit $\Gamma_{\text{therm}} \gg \Gamma_{\text{cycle}}$ are generally more complicated and are likely best solved numerically. The simulation must take into account the depletion of population in velocity classes [56] due to the Raman cooling and properly account for the distribution of atom velocities after they have undergone a Raman cooling cycle, which is generally different than the thermal velocity distribution. It

also must include the repopulation of the depleted states via thermalization. The quasi-steady-state velocity distribution is qualitatively pictured in Fig. 2-5, showing a depletion of the velocity states near the resonant velocity class and accumulation of population at other velocities. Josh Ramette with guidance from Alban Urvoy performed some simulations along these lines, but the results are currently unpublished. Some qualitative results can still be intuited in the limit $\Gamma_{\text{cycle}} \gg \Gamma_{\text{therm}}$. In that limit the cooling rate is no longer strongly affected by Γ_{cycle}; it is instead limited by the rate at which the depleted velocity states are repopulated. Therefore the cooling rate becomes proportional to Γ_{therm} for the depleted velocity classes. Away from the resonant velocity Γ_{cycle} is low and so the velocity classes are not depleted, so the behavior there is effectively in the $\Gamma_{\text{therm}} \gg \Gamma_{\text{cycle}}$ limit. At intermediate velocities the dynamics are more complicated and depend on both Γ_{cycle} and Γ_{therm}.

2.5.4 Rescattering and Festina Lente

Photons must be scattered during the optical pumping stage of Raman cooling in order to remove entropy from the system. Ideally the scattered photons would then simply leave the system without further interactions. However, atomic clouds can be large and dense enough such that a photon scattered off of one atom may then scatter off of another. These rescatterings are generally detrimental to the cooling because the rescattered photons have uncontrolled polarization and wavevectors. That means that they tend to lead to extra recoil heating and can "depump" atoms to undesired states such that they need to be optically pumped again, introducing further recoil heating.

At first it may seem that a simple solution to this issue would be to detune the optical pumping light from the atomic resonance by an amount Δ, increasing the power as needed to maintain the desired optical pumping rate. In that case it may seem that the atomic cloud would be optically thin because the scattered photons would be off-resonant from the atomic transition. However, this emphatically is *not* the case [7]. The problem is that the scattered photon can drive an undesired stimulated

[7] Section 2.5.5 will discuss light-assisted collisions, which can be reduced by detuning the optical

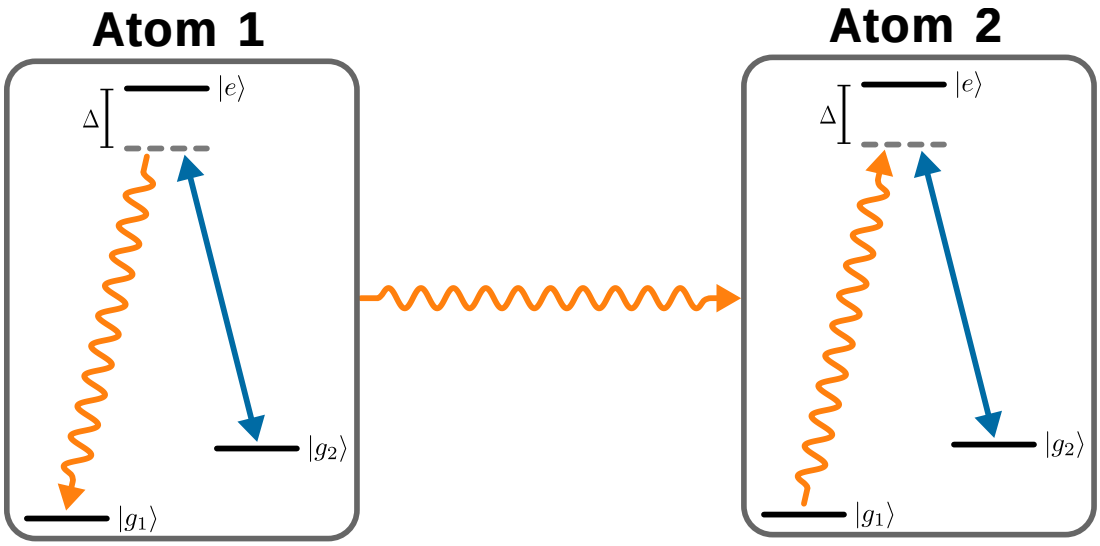


Figure 2-6: Depiction of a Raman rescattering process. Atom 1 is optically pumped from $|g_2\rangle$ back to $|g_1\rangle$ by the optical pumping beam (blue arrows) to complete a Raman cooling cycle, emitting a scattered photon (orange arrows) in the process. The scattered photon travels to atom 2, which is initially in $|g_1\rangle$. The scattered photon and optical pumping beam then drive atom 2 through an unwanted Raman transition from $|g_1\rangle$ to $|g_2\rangle$. Although the transition is detuned from the intermediate state $|e\rangle$ by Δ, the unwanted two-photon Raman transition from $|g_1\rangle$ to $|g_2\rangle$ is resonant. Atom 2 must then be optically pumped back to $|g_1\rangle$ before it can undergo another Raman cooling cycle, and that additional optical pumping step leads to more recoil heating and the potential for additional rescatterings or light-assisted collisions.

Raman transition in tandem with the light remaining in the optical pumping beam, as depicted in Fig. 2-6.

The process behind the unwanted Raman transition is as follows. Consider two atoms, and neglect their motion for now. Atom 1 undergoes a Raman cooling cycle and at the end is optically pumped back from $|g_2\rangle$ to $|g_1\rangle$. The optical pumping produces a spontaneously emitted photon, which leaves from atom 1. That scattered photon then propagates to atom 2, which is in state $|g_1\rangle$. At this point Atom 2 is now driven by two light sources: the scattered photon and the optical pumping beam. Together those beams can drive an undesired Raman transition, moving atom 2 from $|g_1\rangle$ to $|g_2\rangle$. Note that this unwanted stimulated Raman transition driven by the scattered photon is distinct from the velocity-selective stimulated Raman transition and the velocity-insensitive optical pumping transition used for Raman cooling. The unwanted two-photon transition is detuned from the intermediate state $|e\rangle$ by the optical pumping laser's detuning Δ, but the two-photon transition from $|g_1\rangle$ to $|g_2\rangle$ is resonant. Therefore that transition sees the full resonant absorption cross section. Atom 2 then needs to be optically pumped back to $|g_1\rangle$ again before it can start another Raman cooling cycle, and doing that introduces more recoil heating and more chances for additional rescatterings or light-assisted collisions. This form of rescattering can be extremely detrimental to the cooling, particularly for cold and dense clouds.

Fortunately, the probability that a scattered photon is later rescattered before escaping the cloud can be reduced by operating in the festina lente regime [58, 59, 60, 61]. For context, the phrase "festina lente" comes from Latin and means "make haste slowly." It references the balance required between doing things quickly to complete a time-sensitive task as soon as possible, while still performing the task slowly enough to avoid making mistakes. In the festina lente regime, reducing the optical pumping rate reduces the probability that a scattered photon is rescattered. Therefore a balance must be struck between increasing the optical pumping rate to

pumping light. So detuning the optical pumping light does help for other effects, but doesn't resolve the issue of rescattering photons.

cool quickly and decreasing it to keep the rescattering probability sufficiently low; hence the use of the phrase.

In the above description of rescattering, the atomic motion was neglected, however it has important consequences for the rescattering probability. The festina lente regime is a bit different in the free-space and resolved-sideband limits. I will first discuss it in the free-space limit as that is the relevant regime for the work presented in Chapters 4 and 5. Festina lente in the resolved-sideband limit will be discussed later in this subsection.

When an atom is optically pumped, the frequency of the scattered photon is constrained to ensure conservation of energy. Its energy depends not only on the energy of the optical pumping photons and the energy splitting between $|g_1\rangle$ and $|g_2\rangle$, but on the atom's motion as well. The scattered photon will be shifted towards the red if the atom is moving along the optical pumping beam propagation direction, or shifted towards the blue if it is moving in the opposite direction. Similarly, the scattered photon will be shifted towards the blue if it is emitted in the atom's direction of motion, or shifted towards the red if it is emitted opposite the atom's direction of motion. More precisely, the frequency of the scattered photon depends on $\mathbf{v} \cdot \Delta\mathbf{k}$, where here $\Delta\mathbf{k}$ is the difference in wavevectors between the optical pumping beam and the scattered photon. This means that atom 1's velocity affects the energy of the scattered photon. Similarly, atom 2's velocity affects what photon energy is required to make its undesired Raman transition resonant. Therefore the unwanted Raman rescattering process is only resonant if atom 1 and atom 2 have matching velocities. In particular, they must have the same velocity v in the $\Delta\mathbf{k}$-direction. Again, here $\Delta\mathbf{k}$ is the difference in wavevectors between the optical pumping beam and the scattered photon.

This velocity-sensitivity aspect of the rescattering is exactly the same as for the desired velocity-selective stimulated Raman transition that constitutes the first step of a Raman cooling cycle. The only difference is that here the up-leg of the Raman transition is driven by a scattered photon which has uncontrolled direction polarization, and so it typically does not drive a helpful cooling transition even though it

is velocity-selective. Instead it effectively just undoes a successful optical pumping scattering by moving atom 2 to the internal state and v that atom 1 had before its last optical pumping scattering. It is worth emphasizing that the rescattering not only undoes the change in internal state of optical pumping (it moves atom 2 from $|g_1\rangle$ to $|g_2\rangle$ which effectively undoes the optical pumping of atom 1 from $|g_2\rangle$ to $|g_1\rangle$ because one atom still needs to be repumped), it also undoes the recoil heating of the last optical pumping scattering due to the velocity-selectivity of the rescattering. The net effect is that atom 1 and atom 2 just swap internal states and swap v, so there is no excess heating thus far.

The internal and external state swapping that occurs during rescattering would not be particularly problematic on its own; it would just slow down the optical pumping rate for a given scattering rate Γ_{sc} because successful optical pumping scatterings could effectively be undone. However, the problem is worsened by the finite branching ratio b of $|e\rangle$ to $|g_1\rangle$. When atom 1 is optically pumped there are generally multiple scatterings before it is successfully optically pumped, i.e. before $|e\rangle$ happens to decay to $|g_1\rangle$. Each of those unsuccessful optical pumping scatterings lead to some recoil heating. If the successful optical pumping scattering is followed by a rescattering event, then atom 2 is moved to $|g_2\rangle$ and the recoil heating from that last scattering is undone, but the recoil heating from the previous failed optical pumping attempts is *not* undone. Furthermore, atom 2 then needs to be optically pumped which leads to additional recoil heating from each of its optical pumping attempts. This leads to a greater number of total failed optical pumping attempts before all atoms end up in $|g_1\rangle$, which increases the average recoil heating per cooling cycle and thereby makes the cooling less efficient. That aspect of the rescattering is much more problematic as it essentially increases the effective recoil temperature $T_{\mathrm{r}}^{\mathrm{eff}}$. It is effectively the same as reducing b because more optical pumping attempts are required before the excitation to $|g_2\rangle$ is removed from the system.

As with the stimulated Raman transition step of Raman cooling, the undesired rescattering Raman transition does not have to be exactly resonant to occur. This is because the $|g_2\rangle$ state has some width due to the optical pumping beam scattering rate

Γ_{sc} and Raman coupling Rabi frequency Ω_R (See Section 2.2). Therefore the fraction of atoms that are close enough to the two-photon resonance to rescatter depends on those parameters, and so they effect the rescattering probability. Reducing those parameters narrows the resonance width of the unwanted rescattering transitions and thereby reduces their likelihood.

The general expression for rescattering probability is quite complicated, but it has simple forms in various limits. The derivation is beyond the scope of this thesis, but is performed in Ref [59]. The authors calculate the average change in kinetic energy when two atoms, one in $|g\rangle$ and one in $|e\rangle$, are placed in a harmonic trap and then allowed to evolve until the excitation has left the system. Rescattering effects are included, so the effective cross section for reabsorption can be calculated by taking the ratio of the expected recoil heating to the recoil energy and accounting for the spatial distribution of the atoms in the trap. In the free-space limit, the discrete levels of the trapping potential are unresolved and the atomic velocities follow a Maxwell-Boltzmann distribution. There the effective cross section has the form [59]

$$\sigma_{\text{eff}} = \sigma_0 \,\text{Re}\left[\frac{i\sqrt{\pi}\Gamma'\zeta^*}{2\omega_r}\exp\left(\zeta^2\right)\text{erfc}\left(\zeta\right)\right] \tag{2.21}$$

where $\sigma_0 = 4\pi/k^2$ is the resonant cross section, Γ' is the width of $|g_2\rangle$[8], ω_r is the one-photon recoil frequency, $\text{erfc}(x)$ is the complementary error function, ζ is $(\Gamma' + i\omega_r)/(\sqrt{2}\omega_D)$, and ω_D is the Doppler width in angular frequency. The form of Eqn. 2.21 is too complicated to be particularly enlightening as is, but its forms in the limits of large or small Γ'/ω_D are very clear and interesting.

In the limit $\Gamma' \gg \omega_D$, Eqn. 2.21 reduces to $\sigma_{\text{eff}} \approx \sigma_0$ [59]. The physical interpretation of that is clear: The state $|g_2\rangle$ is broadened so much that essentially no atoms are moving fast enough to be off-resonant for the rescattering process. The effective cross section then reduces to the full resonant cross section. In that case the rescattering probability is as if the optical pumping beam were not detuned at all.

[8]Ref. [47] was written with pulsed Raman cooling in mind, so it assumes that the Raman coupling beams are turned off during optical pumping and so the state is only broadened by the optical pumping rate. Here we have replaced that width with Γ' which can include the broadening due to the Raman coupling.

In the limit $\Gamma' \ll \omega_D$, Eqn. 2.21 reduces to [59]

$$\sigma_{\text{eff}} \approx \sigma_0 \frac{\sqrt{\pi}\Gamma'}{2\sqrt{2}\omega_D} \exp\left(-\frac{\omega_r^2}{2\omega_D^2}\right) \tag{2.22}$$

which has a few interesting properties. First of all there is exponential suppression of σ_{eff} when $\omega_D \ll \omega_r$. This is because when the recoil velocity is large compared to the thermal spread in velocities, the first atom will be moving much faster than the typical thermal velocities after decaying and receiving a recoil kick. Due to the Maxwell-Boltzmann distribution, it is then exponentially unlikely that the second atom is moving with that velocity, so it is quite likely off-resonance for rescattering. The other interesting feature of Eqn. 2.22 is that the cross section is suppressed by the ratio Γ'/ω_D. This is a reflection of what was explained qualitatively earlier: the width of the rescattering resonance is Γ'. When the thermal spread in Doppler shifts is large compared to the width of the resonance, the fraction of atoms with a Doppler shift near enough to resonance is proportional to Γ'/ω_D. Therefore, σ_{eff} can be reduced by reducing Γ', which reduces the likelihood of rescattering.

The probability of rescattering depends on the effective cross section and the cloud's spatial distribution. In the limit that rescattering probability is small, which is necessary for efficient cooling, it can be expressed as $p \approx \sigma_{\text{eff}} \langle nl \rangle$ where $\langle nl \rangle$ is the mean column density of the cloud weighted by the emission radiation pattern. The evaluation of $\langle nl \rangle$ can be tedious and will depend on the trap geometry. For spherically uniform radiation from an axially-symmetric harmonic trap it is [62]:

$$\langle nl \rangle = \sqrt{\frac{\pi}{8}} n(0) \sigma_s \frac{\operatorname{artanh}\left(\sqrt{1 - 1/\varepsilon^2}\right)}{\sqrt{\varepsilon^2 - 1}} \tag{2.23}$$

where $n(0)$ is the peak number density, σ_s is the cloud's RMS radius along the tightly-confined direction of the trap, $\operatorname{artanh}(x)$ is the inverse hyperbolic tangent function, and $\varepsilon < 1$ is the ratio of the loose-direction to tight-direction trap frequencies. For

the case of a spherically symmetric trap $\varepsilon = 1$, this reduces to

$$\langle nl \rangle = \sqrt{\frac{\pi}{8}} n(0) \sigma_{\mathrm{s}} \tag{2.24}$$

It is worth explicitly stating that reducing the optical pumping rate doesn't just reduce the rescattering rate simply because there are fewer scatterings; it actually makes it less likely that any given scattered photon will be rescattered. In other words, if the the optical pumping rate is reduced by a factor of two, then the rescattering rate will drop by a factor of four in the free-space festina lente regime. One factor of two comes from the reduced incidence rate of photons, but the other factor of two comes from the reduced probability that any given scattered photon is subsequently rescattered. It is also worth explicitly stating that the benefits of reducing the scattering rate in the festina lente regime are distinct from the benefits for reducing the scattering rate to reduce recoil heating discussed in Section 2.5.1. The festina lente effects discussed here are multi-atom effects associated with rescattering a photon that has already been scattered from the optical pumping beam by another atom; the effect discussed in Section 2.5.1 is a single-atom effect associated with how multiple momentum kicks of different photons directly from the optical pumping beam add up.

The festina lente regime has a qualitatively similar effect in the resolved-sideband limit. In the resolved-sideband limit, the atoms undergo trap oscillations faster than they scatter photons, which complicates the analysis. Quantitative results are presented in Ref. [59] and are summarized qualitatively here. As in the free-space limit, the resonance condition of the unwanted Raman transition determines how the rescattering can occur. In the resolved-sideband limit, the good eigenbasis to work in is the harmonic trap eigenstates rather than the momentum eigenstates. Considering just one dimension for now, suppose that atom 1 is optically pumped from $|g_2\rangle$ to $|g_1\rangle$ and remains in the same harmonic oscillator state $|n\rangle$ that it started in before the optical pumping. If atom 2 undergoes an unwanted rescattering starting in harmonic trap state $|m\rangle$, then it must remain in state $|m\rangle$ as it transitions from $|g_1\rangle$

to $|g_2\rangle$ if it is to conserve energy. Similarly if atom 1 had transitioned from $|n\rangle$ to $|n+1\rangle$ during the optical pumping, then atom 2 would have to transition from $|m\rangle$ to $|m-1\rangle$ in order to conserve energy. The energy added to atom 1 by the recoil is removed from atom 2 if it undergoes the unwanted rescattering resonantly. Just as in the free-space limit, the rescattering removes the energy from atom 2 that was added to atom 1, essentially undoing the optical pumping. However, more recoil heating occurs as atom 2 is optically pumped back to $|g_1\rangle$ due to the finite branching ratio, as was the case in the free-space limit. One difference is that although the atoms had to have matching momentum in the free-space limit, in the resolved-sideband limit they can be in different trap states [9] $|m\rangle$ and $|n\rangle$.

As in the free-space limit, the resonance condition of the unwanted Raman transition does not have to be fulfilled exactly. The $|g_2\rangle$ states are broadened by $\Gamma_{\rm sc}$. Therefore off-resonant excitation can occur, particularly when those broadening parameters are large. Off resonant excitations make it possible for atom 2 to transition from $|m\rangle$ to $|m+1\rangle$ even if atom 1 starts and ends in $|n\rangle$ for example. These off-resonant excitations lead to heating on average and thus are undesirable. As in the free-space limit, reducing $\Gamma_{\rm sc}$ reduces the rate of these off-resonant excitations and thus reduces some of the effects of rescattering.

Another new feature of the festina lente regime which appears in the resolved-sideband limit is the effect of the commensurability of the trap frequencies in the three spatial directions. Suppose that the trap is spherically symmetric and let $|n_x, n_y, n_z\rangle$ denote the trap levels in each direction. If atom 1 transitions from $|n_x, n_y, n_z\rangle$ to $|n_x+1, n_y, n_z\rangle$, then atom 2 can resonantly rescatter the emitted photon by transitioning from $|m_x, m_y, m_z\rangle$ to any of the states $|m_x-1, m_y, m_z\rangle$, $|m_x, m_y-1, m_z\rangle$, or $|m_x, m_y, m_z-1\rangle$. However, if the trap frequencies are properly mismatched, then the only the final state $|m_x-1, m_y, m_z\rangle$ is resonant and the others are not. Along similar lines, suppose that the trap frequency in the x-direction is twice that of the

[9] This is due to the equal spacing between trap levels in a harmonic trap. For an anharmonic trap with different spacings between levels, the rescattering can only happen resonantly when atom 1 ends in the state that atom 2 started in. Note that off-resonant scattering is still possible for other states though.

y-direction. In that case both the final states $|m_x - 1, m_y, m_z\rangle$ and $|m_x, m_y - 2, m_z\rangle$ would be resonant for atom 2. Generally, when one trap frequency is an integer multiple of another (to within the energy resolution set by the broadening), more than one final state for atom 2 is possible. That increases the probability of rescattering. Thus the ratios of the trap frequencies in the three directions can affect the rescattering dynamics. In particular, the likelihood of rescattering is increased when the trap frequencies are commensurate [59].

2.5.5 Light-Assisted Collisions

Laser cooling schemes are generally devised for individual atoms, and so they work for clouds of atoms as long as the atoms act independently. However, when the density becomes large enough, the atoms begin to interact with each other, usually to the detriment of the cooling[10]. One such effect is rescattering of scattered photons, which was discussed in Section 2.5.4. Another such effect is the process of light-assisted collisions, in which two atoms and a photon interact in a way that imparts significant kinetic energy to the atoms. The typical energy can be large compared to a trap depth, resulting in the ejection of both atoms from the trap. That process is the subject of this subsection.

Light-assisted collisions are photo-molecular processes which transfer energy from the photon to the atoms. The energy transferred is typically much more than a normal recoil energy. The process is depicted in Fig. 2-7. The black curves in Fig. 2-7 show the energy of different states for a pair of atoms separated by a distance R. The lower curve shows the potential when both atoms are in their electronic ground state, which is an S-state for alkalis. The molecular interaction potential is for that pair of states is small except at small R due to the $1/R^6$ scaling of the interaction. For the purposes of conceptually understanding light-assisted collisions, it can be approximated that this interaction potential is completely flat. When one atom is moved into an excited electronic state (typically a P-state for alkalis) as it absorbs a photon, the interaction

[10] One notable exception to this is that atoms can collide and thermalize. That can be bad for cooling schemes, but it is required for sub-recoil cooling by thermalization.

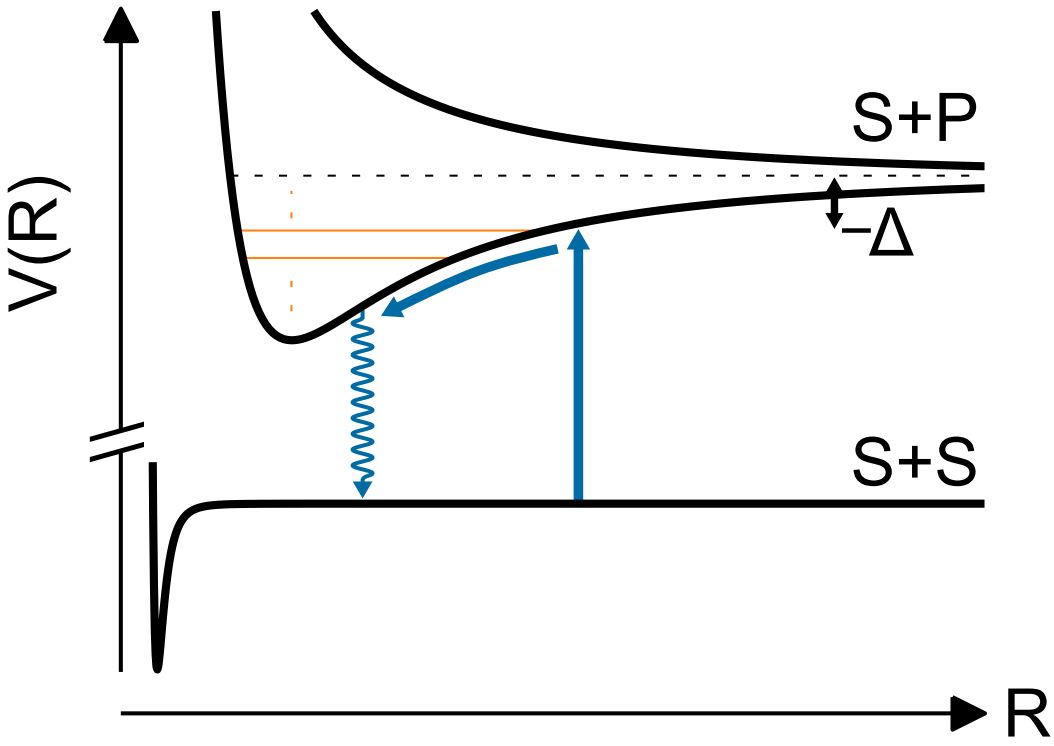


Figure 2-7: The light-assisted collision process. The lower black curve represents the energy of two atoms in their ground state as a function of the distance between them. They do not interact strongly so the curve is flat except at very short distances where the $1/R^6$ potential is appreciable. The upper black curves represent the energy of two atoms when one is in its excited state, which leads to longer range $1/R^3$ interactions, which can be attractive or repulsive depending on the state symmetry. A light-assisted collision occurs when one atom is excited to one of the upper molecular branches by a driving beam (upward blue arrow). Excitation typically occurs near the Condon point, which is where the transition is resonant despite the laser detuning Δ due to the molecular interaction level shift. After one atom is excited, the atoms roll down their relative potential. The excited atom soon decays (wiggly downward blue arrow) due to the finite excited state lifetime, but before doing so the atoms can pick up enough kinetic energy to eject both of them from the trap. The likelihood of this process occurring during a given scattering can be reduced by adjusting Δ to avoid the discrete molecular states, represented as thin orange horizontal lines in the well of the attractive S+P interaction curve. Ellipses indicate the presence of other discrete molecular states which are not shown.

potential changes dramatically. The interaction then scales as $1/R^3$ and so has a much longer effective range. That interaction can also be attractive (lower branch) or repulsive (upper branch) depending on the symmetrization of the wavefunction.

Laser cooling generally involves scattering light from some drive beam, represented by the upward blue arrow in Fig. 2-7, which is detuned from the bare atomic resonance by Δ. When two atoms are separated by a certain distance, called the Condon radius, the molecular interaction shifts the energy levels such that the nominally detuned light can resonantly excite an atom to one of the excited interaction potential branches. The interaction potential when one atom is excited is very steep, so the atoms roll down their relative potential and pick up kinetic energy in the process. At some point the excited atom decays due to its finite lifetime, but the atoms can pick up a significant amount of kinetic energy before then. After the excited atom decays, the interaction potential returns to the lowest curve in Fig. 2-7, at which point they essentially stop interacting because the curve is relatively flat. The kinetic energies gained while one atom was excited can be large compared to typical trap depths, so the atoms can fly out of the trap. The end result is that an attempted scattering off of a single atom during laser cooling can result in the ejection of two atoms from the trap. Because this process requires two atoms, it generally happens more frequently at higher densities where finding two atoms close together becomes more likely. Light-assisted collisions have shown been shown to lead to significant loss when laser cooling at high densities [26]. The rate of light-assisted collisions can be reduced by cooling at lower densities, but this has other consequences for the cooling, many of which will be discussed in Section 2.5.6.

Fortunately it is possible to suppress the rate of light-assisted collisions through careful choice of Δ. Note that the attractive excited state branch potential has a well, which leads to a discrete set of molecular bound states. These are depicted as horizontal orange lines in that well in Fig. 2-7. As pointed out in Ref. [27], choosing Δ to avoid those molecular bound states can suppress the rate of light-assisted collisions. This is because tuning the light partway between molecular resonances makes the excitation to the molecular potential off-resonant. Furthermore, changing Δ moves

the Condon point, and so choosing a value of Δ that puts the Condon point nearer to a node in the ground state molecular wavefunction can also help reduce the rate of light-assisted collisions. Notably, increasing Δ can improve the ratio of "good" to "bad" collisions; i.e. the ratio of normal scatterings (which are generally required for laser cooling) to light-assisted collisions. So although increasing Δ requires increasing the power in order to maintain a fixed scattering rate, doing so can lead to fewer light-assisted collisions for a given number of normal scatterings. This is where the freedom of detuning during Raman cooling mentioned in Section 2.2 plays a significant role.

Calculating the expected light-assisted collision rate as a function of Δ is non-trivial. The matrix elements for coupling to the excited state depend on the Franck-Condon factors. Calculating those factors requires knowledge of the interaction parameters for both ground-ground and ground-excited interaction potentials. Additionally, the energies of the molecular bound states must be calculated, which requires determining the eigenvalues of the excited state potential. Properly accounting for hyperfine structure adds further complications as well. The process for performing these calculation is outlined in Ref. [27].

Although predicting good values of Δ is fairly involved, experimentally finding good values of Δ is relatively straightforward and will be discussed in Chapter 4. The approach there is to prepare a cloud of atoms then scatter an approximately constant number of photons and measure the remaining atom number. By comparing the fraction of atoms remaining at different Δ, good detunings that lead to less loss can easily be identified.

2.5.6 Other Loss and Heating Mechanics

While a cloud is being Raman cooled, there are other processes going on simultaneously that affect the dynamics and the results of the cooling. Those processes provide important considerations that must be accounted for when choosing parameters to create an efficient cooling sequence. This section discuss several such process and how they affect both the clouds dynamics and the final results.

Trap Heating — When studying cooling dynamics it is typically assumed that the

trapping potential is fairly ideal. However real traps generally have some technical issues that can lead to heating. Noise that leads to modulation of the trap potential can lead to parametric heating. For harmonic traps heating occurs when the trap potential is modulated at twice the trap frequency, in which case the cloud is heated exponentially [63]. For optical traps generated by far-detuned light, this process can occur due to intensity noise because the trap depth is proportional to the light intensity. In that case, heating can be reduced by introducing feedback to stabilize the trap beam intensity at the frequency band of interest. Similarly, pointing noise, i.e. modulation of the position of the trap minimum, also leads to heating. For pointing noise, the heating occurs at a constant rate when the trap minimum's position is modulated at the trap frequency [63]. Optical traps can also lead to recoil heating due to off-resonant scattering of trap light [64] and may be able to drive light-assisted collisions. Generally the heating from all of these processes occurs over time, so their effects can be reduced by cooling the cloud quickly to reduce the amount of time spent in the trap.

Vacuum Loss — Ultracold gas experiments are performed in vacuum chambers to avoid collisions between the ultracold gas and molecules in the air. However, although the vacuum chamber pressure is generally very small (ours is typically $\sim 10^{-10}$ mbar $= 10^{-13}$ atm), it is nonzero. That finite pressure means that there is a background gas of atoms and molecules in the chamber which occasionally collide with atoms in the ultracold gas. The background gas has a much higher temperature than the ultracold cloud, so these collisions generally impart significant kinetic energy onto the ultracold atom and eject it from the trap. This process happens at a fixed rate for each trapped atom, leading to an exponential decay in the number of trapped atoms with a timescale called the "vacuum lifetime." Since it happens independently for each atom in the ultracold cloud, it is considered a one-body loss process. In order for the cooling to be efficient, it must happen on a timescale comparable to, or much faster than, the vacuum lifetime. If the cooling is too slow, then the gains in PSD due to the decreasing temperature will be outpaced by the loss in PSD due to the decreasing atom number. This (along with other loss mechanisms) provides

constraints on the achievable temperatures with sub-recoil cooling via thermalization because that cooling becomes exponentially slow when the temperature becomes small compared to $T_{\mathrm{r}}^{\mathrm{eff}}$. In principle the vacuum lifetime can be increased arbitrarily, so it is a practical issue rather than a fundamental one. Notably the vacuum lifetime is independent of the ultracold cloud's density, which makes it different from multi-body loss processes.

Spin-Changing and Hyperfine-Changing Collisions — Raman cooling involves at least two nominally stable states, $|g_1\rangle$ and $|g_2\rangle$, which generally have different energies. For concreteness suppose that $|g_2\rangle$ is at a higher energy than $|g_1\rangle$. Although those states are typically stable for free atoms, collisions between atoms have the potential to move atoms between those states. This implies that two atoms in $|g_2\rangle$ could collide and change states to $|g_1\rangle$, releasing energy in the process. Depending on the states involved, this process is called a spin-changing or hyperfine-changing collision [57]. If the energy released per atom is larger than the trap depth and the atomic cloud is collisionally thin, then the atoms are typically ejected from the trap. If the cloud is collisionally thick, then those atoms can collide with other atoms and impart some or all of their energy to the cloud. Furthermore, if the energy released per atom is less than the trap depth then the atoms typically remain trapped and slowly rethermalize with the cloud. As discussed in the paragraph on Oort cloud heating below, this deposits all of the energy from the collision into the cloud. In order for an atom to undergo one of these collisions, it must find another atom nearby. Therefore the rate (per atom) of these collisions scales linearly with the number density of the cloud n. Unfortunately this is the same scaling as for thermalizing collisions as well, which are 2-body collisions in which the internal atomic states do not change. This means that the ratio of good/thermalizing collisions to bad/spin-changing/hyperfine-changing collisions cannot be tuned by adjusting the cloud density. In real atoms, other ground states besides $|g_1\rangle$ and $|g_2\rangle$ may exist and may become populated or provide other states for collisions products. In alkalis, there are generally two hyperfine manifolds in the ground state. Atoms in the upper manifold can undergo hyperfine-changing collisions, decaying to the lower hyperfine manifold and in the

process release an amount of energy equal to the hyperfine splitting, which is $\sim$ GHz. Atoms can also undergo spin-changing collisions in which they change spin states within a hyperfine manifold. The states within each manifold are non-degenerate if an external magnetic field is applied, in which case an atom can release an amount of energy equal to the differential Zeeman shift between the initial and final state. For atoms with many stable ground states, the choice of which states to use for $|g_1\rangle$ and $|g_2\rangle$ in the Raman cooling process can affect the prevalence of spin-changing and hyperfine-changing collisions. In the work presented in this thesis, we use the $|F = 2, m_F = -2\rangle$ spin-stretched state of ^{87}Rb for $|g_1\rangle$, primarily because it remains a good dark state even for large optical pumping detuning when pumping with light on the D_2 line. However, it is also a good choice because it is the lowest energy state within the upper hyperfine manifold so it two $|2, -2\rangle$ atoms colliding are immune to spin-changing collisions. Furthermore, the rate for hyperfine-changing collisions between two $|2, -2\rangle$ is very low, too low to be measured in Ref. [65] even though they were able to detect hyperfine-changing collisions for other spin states in the upper hyperfine manifold. This is because the $|2, -2\rangle$ state is fully spin-polarized so in order to conserve momentum when decaying to $F = 1$, angular momentum must be moved to the external degrees of freedom. That process is called dipolar relaxation and is generally slow compared to spin-changing collisions in most atoms [57].

3-Body Loss — Typically when talking about ultracold clouds we tend to think that the ground state of the system is a pure BEC or Fermi gas, but strictly speaking this is not true. The actual ground state of ultracold species would be a solid[11]. The only reason that we can talk about ultracold gases as thermal systems is because the timescale for thermalizing to a solid can be much slower than the timescale of thermalization of atoms in the gas state. Although the thermalization towards a solid can be slow, it does still happen and leads to loss from the ultracold cloud. Generally two atoms alone cannot collide and stick together to form the beginnings of a solid because doing so would be impossible while conserving energy and momentum. This

[11] With the exception that helium's ground state (i.e. state at absolute zero) would be liquid helium.

is different than for spin-changing and hyperfine-changing collisions because the fact that the atoms are stuck together in the final state adds a constraint to the possible momenta in the final state (the two atoms must have the same final momentum if they are stuck together) and that over-constrains the problem. A third object is needed in order to make it possible for two atoms to stick together while conserving energy and momentum. Typically the third object is another atom in the cloud, in which case the process is called 3-body loss. In a 3-body loss collision, three atoms collide simultaneously. Two of the atoms stick together to form a molecule, and the third atom remains free but provides a way for the collision to conserve energy and momentum. Because an atom must find two other nearby atoms in order to undergo a 3-body loss collision, the 3-body loss rate per atom scales as n^2. That is a stronger scaling with density than the n scaling of the 2-body thermalizing collisions which are necessary for sub-recoil cooling via thermalization. Therefore the degree of 3-body loss can be reduced by lowering the density of the cloud, which can be done at fixed total atom number by reducing the confining strength of the trap. Doing so does come at a cost though as it reduces the critical temperature and slows down thermalization.

Oort Cloud Heating — In the early days of BEC, it was noticed that the lifetime of BECs in magnetic traps could be increased by keeping on the RF knife used during evaporation even after the evaporation has finished [66, 67]. In Ref. [66] doing so increased the lifetime of the BEC from 1 second to 20 seconds. The explanation for that behavior is that there are processes that produce highly energetic atoms. If the energy of those atoms is lower than the trap depth, they remain in the trap and form an "Oort cloud" of unthermalized atoms. Those atoms then slowly thermalize with the cold cloud, departing up to a trap depth of energy per Oort cloud atom in the process. Keeping the RF knife on reduces the trap depth, which makes it easier for the highly energetic atoms to escape. That both reduces the number of atoms in the Oort cloud and reduces the average energy imparted per Oort cloud atom. This is somewhat different from evaporative cooling because the hot atoms are not just the atoms from the high-energy tail of the thermal distribution. The

high-energy state populations are larger than expected from a thermal distribution because they include non-thermal atoms which gained high energy through processes other than thermalizing collisions. Therefore the effects of varying the trap depth can be significant even if the trap depth is set too large for evaporation to play a significant role. This same process occurs in non-magnetic traps; the BEC lifetime is shortened by large trap depths in general, so the trap depth must be kept low. There are various possible sources for the high-energy atoms in the Oort cloud. Generally grazing collisions with atoms that have much more than a trap depth of energy can impart slightly less than a trap depth of energy onto previously cold atoms. For example, a room temperature atom in the background gas of the vacuum chamber could undergo a grazing collision with an ultracold atom, giving it slightly less than a trap depth of energy in the process and thereby putting it into the Oort cloud. Similarly, products of loss processes such as light-assisted collisions and spin-changing or hyperfine-changing collisions can also undergo grazing collisions with other cold atoms. For rubidium, the products of 3-body loss in particular may be very prone to causing this due to an unfortunate coincidence in which a common energy for 3-body products (set by a molecular state's binding energy) is close to the d-wave scattering resonance [62].

Raman Heating in Other Hyperfine Manifold — Alkali atoms have two hyperfine manifolds in their ground state which have opposite g_F. This is problematic for degenerate Raman cooling, i.e. Raman sideband cooling where the two beams driving the Raman transition have the same frequency. This is frequently the case when performing Raman sideband cooling, in which case the Raman transition detuning is typically controlled by adjusting the magnetic field to shift the energy difference between Zeeman states. When the magnetic field is set to drive cooling Raman transitions for atoms in one hyperfine manifold, it unfortunately is then tuned to drive heating Raman transitions for atoms in the other hyperfine manifold due to its opposite g_F. Thus care must be taken to avoid heating in one manifold while cooling in the other. A simple way to avoid that is to quickly optically pump the atoms out of the undesired manifold. By pumping them fast compared to Ω_{R}, one

can take them out of the undesired hyperfine manifold before they undergo a heating cycle. Another way to avoid this is to introduce a beam which light shifts the various Zeeman states differently. In particular if the energy splitting between two states in one manifold is sufficiently different than the energy splitting between two states in the other manifold then only one is resonant at a time. Thus the heating transitions in one manifold can be off-resonant while the cooling transition in the other is on-resonant. This occurred somewhat naturally in much of our Raman sideband cooling work due to the far-detuned circularly-polarized optical pumping beam which can induce significant light shifts. More discussion on heating in the undesired hyperfine manifold is provided in Ref. [68].

2.6 Raman Coupling Internal Matrix Elements

Thus far we have taken the velocity-selective two-photon Raman transition Rabi frequency Ω_R as a given value. In practice it can be calculated given the beam powers, detunings, polarizations, and trap configuration. The calculation is conceptually straightforward, but the results can be somewhat surprising. Cancellations, or near-cancellations, between different terms can occur when summing over all possible intermediate states and paths from $|g_1\rangle$ to $|g_2\rangle$. That can lead to much smaller Ω_R than one may have otherwise expected. This section shows how to calculate Ω_R in general, and then shows example calculations for ^{87}Rb, showing effects of cancellations and near-cancellations in the process. Notably this section will only calculate the contribution to Ω_R from the matrix elements corresponding to the internal state of the atom. This is equal to the actual Raman transition Rabi frequency for an atom in the free-space limit, but when the atom is tightly confined, then Ω_R will depend on the external state matrix elements as well. The effects of tight confinement will be discussed in Section 2.7.

The approach taken here uses a direct application of second-order perturbation theory. Another approach is presented in Ref. [68] which uses the formalism of effective magnetic fields to derive Ω_R with an emphasis on application to degenerate

(i.e. the Raman beams are at the same frequency) Raman sideband cooling. That approach, while more elegant and often simpler to evaluate, is more complicated to derive and can therefore be less intuitive. For this reason, second-order perturbation theory is used directly here as a complementary approach.

2.6.1 Background and Conventions

This subsection contains basic formulas from quantum physics which will be used in the rest of the chapter. Its intent is not to teach introductory quantum mechanics, but to include certain results for convenient reference and to make clear to the reader which conventions are being used. The conventions and notation used here mainly follows that used in Ref. [69] and Ref. [70].

The light field created by a beam will are assumed to be of the form

$$\mathcal{E}(\mathbf{r}, t) = E_0 \hat{\epsilon} \cos\left(\mathbf{k} \cdot \mathbf{r} - \omega_{\mathrm{L}} t\right) \tag{2.25}$$

where E_0 is the electric field amplitude, $\hat{\epsilon}$ is the electric field polarization unit vector, $\mathbf{k}$ is the beam's wavevector, and ω_{L} is the frequency of the driving laser. In the lab, typically the beam intensity or its power and waist are measured, so it is helpful to have formulas for calculating E_0 from those values. A beam with electric field amplitude E_0 has intensity (averaged over cycles) of $I = E_0^2 c\epsilon_0/2$ in SI units where c is the speed of light and ϵ_0 is the vacuum permittivity. This can be inverted to give

$$E_0 = \sqrt{\frac{2I}{c\epsilon_0}} \tag{2.26}$$

Often in the lab beams have a Gaussian profile, in which case the intensity isn't uniform across the beam's cross section. For Gaussian beams, the peak intensity at the center of the beam I_{p} is related to the power P and beam width w via

$$I_{\mathrm{p}} = \frac{2P}{\pi w^2} \text{ (For a Guassian Beam)} \tag{2.27}$$

Here the beam width is the $1/e$ radius of the beam's electric field amplitude, which

is the same as the $1/e^2$ radius of the beam intensity, which in turn is $\sqrt{2}$ times the $1/e$ power radius or the $1/e$ power diameter divided by $\sqrt{2}$. Combining the last two equations yields the peak electric field amplitude as a function of Gaussian beam power and beam width

$$E_{0,\mathrm{p}} = \sqrt{\frac{4P}{c\epsilon_0 \pi w^2}} \tag{2.28}$$

And the electric field amplitude falls off as a Gaussian from its peak value when moving radially away from the beam axis.

$$E_0(r) = E_{0,\mathrm{p}} \exp\left(-\frac{r^2}{w^2}\right) \tag{2.29}$$

Lastly the beam width can be calculated from the beam waist w_0 (its width at its narrowest point) as $w(z) = \sqrt{1 + (z/z_\mathrm{R})^2}$ where z is the position along the beam axis and $z_\mathrm{R} = \pi w_0^2/\lambda$ is the Rayleigh range.

The Hamiltonian for the interaction between the atom and the field is $\mathcal{H}(t) = -\mathbf{d} \cdot \mathcal{E}(\mathbf{r}, t)$. For atoms with a single valence electron $\mathbf{d} = -e\mathbf{r}$ where e is the elementary charge and $\mathbf{r}$ is the electron's position vector. After making the dipole approximation, that can be rewritten as $\mathcal{H}(t) = -\hbar\Omega \cos(\mathbf{k} \cdot \mathbf{r} - \omega_\mathrm{L} t)$ with $\hbar\Omega = d_{eg} E_0$ where $d_{eg} = e \langle e|\hat{\epsilon} \cdot \mathbf{r}|g\rangle$. After making the rotating wave approximation, the matrix element between the ground and excited state is $\mathcal{H}_{eg} = \langle e|\mathcal{H}|g\rangle = \hbar\Omega/2$. The factor of $1/2$ comes when dropping the counter-rotating term. In brief, $\mathcal{H} = -\mathbf{d} \cdot \mathcal{E}(t)$ and $\mathcal{E}(t) = E_0 \hat{\epsilon} \cos(\omega_\mathrm{L} t) = (E_0 \hat{\epsilon}/2)(\exp(i\omega_\mathrm{L} t) + \exp(-i\omega_\mathrm{L} t))$. When the counter-rotating term is dropped, the factor of $1/2$ from converting the cosine to a sum of complex exponentials remains. Essentially, half of the electric field's amplitude goes to driving the co-rotating term, while the other half goes into driving the counter-rotating term.

The spherical unit vectors are $\hat{u}_{-1} = (\hat{x} - i\hat{y})/\sqrt{2}$, $\hat{u}_0 = \hat{z}$, and $\hat{u}_{+1} = -(\hat{x} + i\hat{y})/\sqrt{2}$. An electric field with polarization $\hat{\epsilon} = \hat{u}_{-1}$ drives only σ^- transitions. Similarly an electric field with polarization $\hat{\epsilon} = \hat{u}_0$ only drives π transitions and an electric field with polarization $\hat{\epsilon} = \hat{u}_{+1}$ only drives σ^+ transitions. More generally if the electric field polarization in Cartesian coordinates is $\hat{\epsilon} = A_x\hat{x} + A_y\hat{y} + A_z\hat{z}$ then the polarization

in terms of spherical unit vectors is

$$\hat{\epsilon} = \frac{A_x - iA_y}{\sqrt{2}}\hat{u}_{-1} + A_z\hat{u}_0 + \frac{-A_x - iA_y}{\sqrt{2}}\hat{u}_{+1} \tag{2.30}$$

2.6.2 Calculating Single-Photon Transition Rabi Frequencies

Calculating Raman transition Rabi frequencies requires first calculating the Rabi frequencies of single-photon transitions from more convenient lab units, like intensity or beam power and waist. This is particularly easy to do using the popular Steck D-line data notes, at least for species for which those notes exist. For other species, the process is the same but the reader will have to find the values of the corresponding constants from other sources and carefully ensure that the same conventions, e.g. normalization conditions, are used. In this case we'll use the notes for ^{87}Rb [71].

Each of the polarizations π, σ^+, and σ^- couples the initial state to a different final state. Therefore it is easiest to decompose the beam polarization into a sum of those components, then treat each component separately. The component in a given polarization can be written as $\hat{d}_i \cdot \hat{\epsilon}$ where $\hat{d} = \mathbf{d}/|\mathbf{d}|$ is the unit vector of the dipole matrix element and its components are $\hat{d}_i$ where i can be any of π, σ^+, or σ^-. Each d_i is then equal to its corresponding spherical unit vector, so $\hat{d}_{\sigma^-} = \hat{u}_{-1}$, $\hat{d}_{\pi} = \hat{u}_0$, and $\hat{d}_{\sigma^+} = \hat{u}_{+1}$. The Clebsch-Gordan coefficient C_i for the transition for a given component can then be read off from the tables of values for the various possible initial and final hyperfine states for each D-line. The dipole matrix element component d_i is then calculated by multiplying the Clebsch-Gordan coefficient by the reduced matrix element for the D-line of the transition, either $\langle J = 1/2 \| er \| J' = 1/2 \rangle$ for the D_1 transition or $\langle J = 1/2 \| er \| J' = 3/2 \rangle$ for the D_2 transition. The values of those reduced matrix elements are recorded in Table 7 of Ref. [71]. Notably they differ due to the normalization condition used so $\langle J = 1/2 \| er \| J' = 1/2 \rangle = \sqrt{2}\,\langle J = 1/2 \| er \| J' = 3/2 \rangle$. In words, the reduced matrix element for the D_2 line is a factor of $\sqrt{2}$ larger than that of the D_1 line. That will be helpful to know when comparing the contributions to the Raman coupling from the D_1 and D_2 lines.

The Rabi frequency can then be calculated using the relation $\hbar\Omega_i = \mathbf{d}_i \cdot \mathcal{E} =$

$d_i E_0(\hat{d}_i \cdot \hat{\epsilon})$. Using Eqn. 2.26 this can be written in a more convenient form

$$\boxed{\hbar\Omega_i = \begin{cases} \mathrm{C}_i \langle J = 1/2 \| er \| J' = 1/2 \rangle (\hat{d}_i \cdot \hat{\epsilon}) \sqrt{\frac{2I}{c\epsilon_0}} & {}^{87}\mathrm{Rb}\ D_1\ \mathrm{Line} \\ \mathrm{C}_i \langle J = 1/2 \| er \| J' = 3/2 \rangle (\hat{d}_i \cdot \hat{\epsilon}) \sqrt{\frac{2I}{c\epsilon_0}} & {}^{87}\mathrm{Rb}\ D_2\ \mathrm{Line} \end{cases}} \tag{2.31}$$

Or using Eqn. 2.28 this can be used to calculate the peak value of Ω_i at the center of the cross section of a Gaussian beam at any point long its path

$$\hbar\Omega_{i,\mathrm{p}} = \begin{cases} \mathrm{C}_i \langle J = 1/2 \| er \| J' = 1/2 \rangle (\hat{d}_i \cdot \hat{\epsilon}) \sqrt{\frac{4P}{c\epsilon_0 \pi w^2}} & {}^{87}\mathrm{Rb}\ D_1\ \mathrm{Line} \\ \mathrm{C}_i \langle J = 1/2 \| er \| J' = 3/2 \rangle (\hat{d}_i \cdot \hat{\epsilon}) \sqrt{\frac{4P}{c\epsilon_0 \pi w^2}} & {}^{87}\mathrm{Rb}\ D_2\ \mathrm{Line} \end{cases} \tag{2.32}$$

And the value away from the beam axis falls off as $\exp(-r^2/w^2)$ as in Eqn. 2.29 because the Rabi frequency is proportional to the electric field.

These equations can also be written in another convenient form by substituting in simple values for parameters then including the scalings (note that $\hbar$ has now been moved to the other side)

$$\Omega_i = \begin{cases} (2\pi * 1.0513\ \mathrm{MHz})\, \mathrm{C}_i (\hat{d}_i \cdot \hat{\epsilon}) \sqrt{\frac{I}{1\ \mathrm{W/m^2}}} & {}^{87}\mathrm{Rb}\ D_1\ \mathrm{Line} \\ (2\pi * 1.4847\ \mathrm{MHz})\, \mathrm{C}_i (\hat{d}_i \cdot \hat{\epsilon}) \sqrt{\frac{I}{1\ \mathrm{W/m^2}}} & {}^{87}\mathrm{Rb}\ D_2\ \mathrm{Line} \end{cases} \tag{2.33}$$

Note that the intensity is given in SI units ($\mathrm{W/m^2}$) in the above formula rather than $\mathrm{mW/cm^2}$ which is a bit more typical in AMO physics ($1\ \mathrm{mW/cm^2} = 10\ \mathrm{W/m^2}$). The formulas for $\Omega_{i,\mathrm{p}}$ in terms of Gaussian beam parameters can be written in the same way

$$\Omega_{i,\mathrm{p}} = \begin{cases} (2\pi * 0.83879\ \mathrm{MHz})\, \mathrm{C}_i (\hat{d}_i \cdot \hat{\epsilon}) \left(\frac{1\ \mathrm{m}}{w}\right) \sqrt{\frac{P}{1\ \mathrm{W}}} & {}^{87}\mathrm{Rb}\ D_1\ \mathrm{Line} \\ (2\pi * 1.1846\ \mathrm{MHz})\, \mathrm{C}_i (\hat{d}_i \cdot \hat{\epsilon}) \left(\frac{1\ \mathrm{m}}{w}\right) \sqrt{\frac{P}{1\ \mathrm{W}}} & {}^{87}\mathrm{Rb}\ D_2\ \mathrm{Line} \end{cases} \tag{2.34}$$

And again the value away from the beam axis falls off as $\exp(-r^2/w^2)$.

2.6.3 General Calculation Outline

To avoid unwanted scattering off of the intermediate excited state, Raman coupling is typically done in the limit of large intermediate state detuning Δ_{R}. In that case the excited state can be adiabatically eliminated and Ω_{R} can be calculated via second-order perturbation theory. For a simple three-level atom, the result after making the rotating wave approximation is the well-known formula $\Omega_{\mathrm{R}} = \Omega_1\Omega_2/(2\Delta_{\mathrm{R}})$ where Ω_1 is the Rabi frequency of the beam driving the up-leg and Ω_2 is the Rabi frequency of the beam driving the down-leg.

More generally, real atoms have multiple intermediate excited states that can contribute to the coupling. In that case second-order perturbation theory[12] gives the effective coupling between $|g_1\rangle$ and $|g_2\rangle$

$$\langle g_2|\mathcal{H}_{\mathrm{eff}}|g_1\rangle = \frac{\hbar\Omega_{\mathrm{R}}}{2} = \frac{1}{2}\sum_i \langle g_2|\mathcal{H}|e_i\rangle \langle e_i|\mathcal{H}|g_1\rangle \left(\frac{1}{E_2 - E_i} + \frac{1}{E_1 - E_i}\right) \tag{2.35}$$

where E_1 and E_2 are the energies of the initial and final states respectively (including the energy of the photons) and E_i is the energy of the intermediate state $|e_i\rangle$. To proceed further, I'll introduce some more new notation. For now let us assume that beam 1 drives the up-leg of the Raman transition and beam 2 drives the down-leg[13]. Let $\omega_{\mathrm{L},1}$ and $\omega_{\mathrm{L},2}$ be the angular frequencies of the beams driving the up-leg and down-leg respectively. Next, let $\omega_{0,g1,ei}$ be the resonant angular frequency for the $|g_1\rangle \to |e_i\rangle$ transition, and let $\omega_{0,g2,ei}$ be the resonant angular frequency for the $|g_2\rangle \to |e_i\rangle$ transition. This implies $\delta_{\mathrm{R}} = \hbar(\omega_{\mathrm{L},1} - \omega_{\mathrm{L},2}) - \hbar(\omega_{0,g2,ei} - \omega_{0,g1,ei})$. In words, δ_{R} is the angular frequency detuning of the Raman transition from the two-photon $|g_1\rangle \to |g_2\rangle$ resonance when the up-leg is driven by beam 1 and the down-leg is driven by beam 2. Let $\Delta_{\mathrm{R},i} = \omega_{\mathrm{L},1} - \omega_{0,g1,ei}$ be the angular frequency detuning of beam 1 (the up-leg driving beam) from resonance for the $|g_1\rangle \to |e_i\rangle$ transition. Additionally, let $\Omega_{1,i}$ be the single-photon transition angular Rabi frequency for beam 1 coupling

[12] See Eqn. 26 on pg. 45 of Ref. [53]. Here $|g_1\rangle$ and $|g_2\rangle$ are different states, so the zeroth-order term in that equation drops out. I'll also assume that the drives do not couple $|g_1\rangle$ and $|g_2\rangle$ directly so that the first order term drops out as well, leaving only the indirect coupling through the excited states.

[13] The case that either beam can drive either leg will be discussed later in this section.

$|g_1\rangle$ to $|e_i\rangle$, and similarly let $\Omega_{2,i}$ be the corresponding value for beam 2 coupling $|g_2\rangle$ to $|e_i\rangle$.

Eqn. 2.35 can be significantly simplified in the rotating wave approximation. The matrix elements to the intermediate excited states are then $\hbar\Omega_{1,i}/2$ and $\hbar\Omega_{2,i}/2$, which gives

$$\Omega_{\mathrm{R}} = \frac{1}{\hbar}\sum_i \left(\frac{\hbar\Omega_{2,i}^*}{2}\right)\left(\frac{\hbar\Omega_{1,i}}{2}\right)\left(\frac{1}{E_2 - E_i} + \frac{1}{E_1 - E_i}\right) \tag{2.36}$$

In the rotating wave approximation, the value of E_i is then $E_1 - \hbar\omega_{\mathrm{L},1} + \hbar\omega_{0,g1,ei}$ because it differs from the initial state by the absorption of one photon from beam 1 and excitation of the atom to state $|e_i\rangle$ from $|g_1\rangle$. That can rewritten as $E_i = E_1 - \hbar\Delta_{\mathrm{R},i}$. Similarly, $E_2 = E_i - \hbar\omega_{0,g2,ei} + \hbar\omega_{\mathrm{L},2}$ because it differs from the intermediate state by the emission of one photon into beam 2 and the de-excitation of the atom to state $|g_2\rangle$ from state $|e_i\rangle$. The above expression for E_i can then be inserted into the expression for E_2, which gives $E_2 = E_1 - \hbar\delta_{\mathrm{R}}$. These results can be inserted into Eqn. 2.36 to give

$$\Omega_{\mathrm{R}} = \sum_i \frac{\Omega_{2,i}^*\Omega_{1,i}}{4}\left(\frac{1}{\Delta_{\mathrm{R},i} - \delta_{\mathrm{R}}} + \frac{1}{\Delta_{\mathrm{R},i}}\right) \tag{2.37}$$

Generally $\Delta_{\mathrm{R},i}$ is large compared to δ_{R} in order to have a good ratio of Raman coupling to scattering off of the intermediate state $|e_i\rangle$. In that case, the contribution of δ_{R} to the denominator can be neglected, giving the final general expression of Ω_{R} in the rotating wave approximation.

$$\boxed{\Omega_{\mathrm{R}} = \sum_i \frac{\Omega_{2,i}^*\Omega_{1,i}}{2\Delta_{\mathrm{R},i}}} \tag{2.38}$$

This very similar to the expression $\Omega_{\mathrm{R}} = \Omega_1\Omega_2/(2\Delta_{\mathrm{R}})$ mentioned at the beginning of the chapter for a three-level atom. The only difference is that all possible paths contribute, which requires summing over all possible intermediate states. Notably each term in the sum can be positive or negative (or have any complex phase) so terms for different intermediate states can interfere constructively or destructively. Keep in mind that the actual Rabi frequency of the Raman coupling can be affected by matrix elements for the atom's external state, which will be discussed in Section 2.7.

It will turn out that the actual Raman coupling Rabi frequency is equal to the value calculated from the internal matrix elements alone when the cooling is in the free-space limit.

At this point it is worth discussing the effects of the counter-rotating component for each term in Eqn. 2.35 which was dropped in the construction of Eqn. 2.38. This discussion is easiest to have in second quantization, so let $a_1^\dagger$ and a_1 be the creation and annihilation operators for the beam 1 photons. Also let σ_{1ei}^+ and σ_{1ei}^- be the raising and lowering operators for the atoms state on the $|g_1\rangle$ to $|e_i\rangle$ transition, i.e. $\sigma_{1ei}^+ = |e_i\rangle\langle g_1|$ and $\sigma_{1ei}^- = |g_1\rangle\langle e_i|$. Let the corresponding operators for beam 2 and $|g_2\rangle$ be the same, but with the 1 subscripts changed to 2. Consider any individual term of the sum in Eqn. 2.35. The $\mathbf{d}\cdot\mathcal{E}$ Hamiltonian for beam 1 in second quantization has the form $(\hbar\Omega_{1,i}/2)(\sigma_{1ei}^+ + \sigma_{1ei}^-)(a_1^\dagger + a_1)$, which can be expanded to $(\hbar\Omega_{1,i}/2)(\sigma_{1ei}^+ a_1^\dagger + \sigma_{1ei}^+ a_1 + \sigma_{1ei}^- a_1^\dagger + \sigma_{1ei}^- a_1)$. When this operator acts on $|g_1\rangle$, as it does in the right inner product in each term of Eqn. 2.35, the components with σ_{1ei}^- give zero, so the only part that can give a nonzero contribution is $(\hbar\Omega_{1,i}/2)(\sigma_{1ei}^+ a_1^\dagger + \sigma_{1ei}^+ a_1)$. The $\sigma_{1ei}^+ a_1$ contribution is the co-rotating component corresponding to absorbing a photon and exciting the atom, while the $\sigma_{1ei}^+ a_1^\dagger$ contribution is the counter-rotating component corresponding to emitting a photon and exciting the atom. A similar argument applies to the down-leg, except only the σ_{2ei}^- terms survive when they act on the intermediate $|e_i\rangle$ state in the left inner product in the numerator of Eqn. 2.35, leaving $(\hbar\Omega_{2,i}/2)(\sigma_{2ei}^- a_2^\dagger + \sigma_{2ei}^- a_2)$. The two inner products then give a result of the form

$$\left(\frac{\hbar\Omega_{1,i}}{2}\right)\left(\frac{\hbar\Omega_{2,i}}{2}\right)\left\langle g_2\right|\left(\sigma_{2ei}^- a_2^\dagger \sigma_{1ei}^+ a_1^\dagger + \sigma_{2ei}^- a_2^\dagger \sigma_{1ei}^+ a_1 + \sigma_{2ei}^- a_2 \sigma_{1ei}^+ a_1^\dagger + \sigma_{2ei}^- a_2 \sigma_{1ei}^+ a_1\right)\left|g_1\right\rangle \tag{2.39}$$

Two of those operator products strongly violate energy conservation between the initial and final state (not just the intermediate state) because they either remove a photon from each beam or add a photon to each beam, and so those contributions can be neglected. This leaves

$$\left(\frac{\hbar\Omega_{1,i}}{2}\right)\left(\frac{\hbar\Omega_{2,i}}{2}\right)\left\langle g_2\right|\left(\sigma_{2ei}^- a_2^\dagger \sigma_{1ei}^+ a_1 + \sigma_{2ei}^- a_2 \sigma_{1ei}^+ a_1^\dagger\right)\left|g_1\right\rangle \tag{2.40}$$

Notably both of those terms have the same prefactor.

It is convenient to keep track of the different paths from $|g_1\rangle$ to $|g_2\rangle$ by adding to the ket the number of photons in each beam so $|N_1, N_2, g_1\rangle$ corresponds to the initial state with N_1 photons in beam 1 and N_2 photons in beam two. The first of those operator products corresponds to the co-rotating path $|N_1, N_2, g_1\rangle \rightarrow |N_1 - 1, N_2, e_i\rangle \rightarrow |N_1 - 1, N_2 + 1, g_2\rangle$, while the second of those operator products corresponds to the counter-rotating path $|N_1, N_2, g_1\rangle \rightarrow |N_1 + 1, N_2, e_i\rangle \rightarrow |N_1 + 1, N_2 - 1, g_2\rangle$. The intermediate state of the counter-rotating path has a very different energy than the initial state. That means that this path involves a large denominator in the second order coupling and so it is generally smaller than the co-rotating term.

In addition to the different energies in the intermediate state for the two paths, note that the number of photons in each beam in the final state is different for each path. In the co-rotating path, beam 1 loses a photon and beam 2 gains a photon, but the opposite occurs in the counter-rotating path. This has two noteworthy consequences. First, the counter-rotating path imparts the opposite momentum kick $-\hbar\Delta\mathbf{k}$ to the atom. That implies that it affects the atom's external states differently and so the counter-rotating terms should be summed separately from the co-rotating terms as they will couple the atom to different momentum states. The two terms should only be added together after including the matrix elements for the external state of the atom, as will be discussed in Section 2.7. Second, the final state energy for the co-rotating path is δ_{R}, but for the counter-rotating path it is $\delta_{\mathrm{R}} + 2(\omega_{\mathrm{L},1} - \omega_{\mathrm{L},2})$. If the beams are far from degenerate, then the counter-rotating path is far from the two-photon Raman resonance and will have little effect on the dynamics. That gives further justification for making the rotating wave approximation in that situation; not only is the intermediate state off-resonant, so is the final state. However, if the beams are degenerate, then the co-rotating and counter-rotating paths have the same two-photon Raman detuning. In that case they can both contribute and only differ in the momentum imparted on the atom and the amplitude of their contribution due to the different energies of the intermediate states. The Raman coupling Rabi frequency of the counter-rotating term with momentum impulse $-\hbar\Delta\mathbf{k}$ can be written from

Eqn. 2.36. The intermediate state now has energy $E_i = E_1 + \hbar\omega_{\mathrm{L},1} + \hbar\omega_{0,g1,ei}$ and the final state has energy $E_2 = E_i - \hbar\omega_{\mathrm{L},2} - \hbar\omega_{0,g2,ei}$. This gives

$$\Omega_{\mathrm{R},-\Delta\mathbf{k}} = -\sum_i \frac{\Omega^*_{2,i}\Omega_{1,i}}{4}\left(\frac{1}{\omega_{\mathrm{L},2} + \omega_{0,g2,ei}} + \frac{1}{\omega_{\mathrm{L},1} + \omega_{0,g1,ei}}\right) \tag{2.41}$$

Again note that this contribution should *not* simply be added with Eqn. 2.38 to account for the counter-rotating path because they have different corresponding external state matrix elements due to the opposite associated momentum kicks. Hence the addition of the $-\Delta\mathbf{k}$ subscript in Eqn. 2.41.

It is now worth mentioning the effects that can occur when the two Raman beams are degenerate. In this situation, either beam can drive the up-leg of the transition. The two co-rotating paths are now $|N_1, N_2, g_1\rangle \rightarrow |N_1 - 1, N_2, e_i\rangle \rightarrow |N_1 - 1, N_2 + 1, g_2\rangle$ which corresponds to the case that beam 1 drives the up-leg, and $|N_1, N_2, g_1\rangle \rightarrow |N_1, N_2 - 1, e_i\rangle \rightarrow |N_1 + 1, N_2 - 1, g_2\rangle$ which corresponds to the case that beam 2 drives the up-leg. The two possibilities give opposite momentum kicks because they swap the absorption and emission between the beams, which means that they have different corresponding external state matrix elements. Therefore it is necessary to sum up each possibility separately, performing the sum in Eqn. 2.38 once as written for the $+\hbar\Delta\mathbf{k}$ coupling, then again with the roles of beam 1 and 2 swapped to get the $-\hbar\Delta\mathbf{k}$ coupling. The roles of beams 1 and 2 can be swapped by swapping the indices referring to the beams in the subscripts (be sure not to swap the indices referring $|g_1\rangle$ and $|g_2\rangle$ though).

With degenerate beams there are also two counter-rotating paths. These are $|N_1, N_2, g_1\rangle \rightarrow |N_1 + 1, N_2, e_i\rangle \rightarrow |N_1 + 1, N_2 - 1, g_2\rangle$ which corresponds to the case that beam 1 drives the up-leg, and $|N_1, N_2, g_1\rangle \rightarrow |N_1, N_2 + 1, e_i\rangle \rightarrow |N_1 - 1, N_2 + 1, g_2\rangle$ which corresponds to the case that beam 2 drives the up-leg. Interestingly, because the counter-rotating path contributes the opposite momentum kick relative to the corresponding co-rotating term (i.e. the co-rotating term which involves the same beam driving the up-leg), it can be included in the sum for the other configuration. In other words, the counter-rotating term of the path where beam 2 drives the up-leg provides

the same $+\hbar\Delta\mathbf{k}$ momentum kick as the co-rotating term of the path where beam 1 drives the up-leg. Therefore both of those contributions can be summed together since they act the same way on the atom's external state. Generally the single-photon Rabi frequencies will be different for each beam driving each transition and will need to be calculated separately to account for the differing polarizations and so on. To specify this, the subscripts on the single-photon Rabi frequencies below will include not only which beam they correspond to, but which states are connected as well. For example, $\Omega_{2,g1,ei}$ is the single-photon Rabi frequency of beam 2 driving the $|g_1\rangle \rightarrow |e_i\rangle$ transition. The sum for the Raman transition Rabi frequency $\Omega_{\mathrm{R},+\Delta\mathbf{k}}$ corresponding to a $+\hbar\Delta\mathbf{k}$ momentum kick, including the counter-rotating contribution, is

$$\Omega_{\mathrm{R},+\Delta\mathbf{k}} = \sum_i \left[\frac{\Omega^*_{2,g2,ei}\Omega_{1,g1,ei}}{2\Delta_{\mathrm{R},i}} - \frac{\Omega^*_{1,g2,ei}\Omega_{2,g1,ei}}{4} \left(\frac{1}{\omega_{\mathrm{L},1} + \omega_{0,g2,ei}} + \frac{1}{\omega_{\mathrm{L},2} + \omega_{0,g1,ei}} \right) \right] \tag{2.42}$$

The expression for the Raman transition Rabi frequency corresponding to a $-\hbar\Delta\mathbf{k}$ kick is similar, but with the indices referring to the beams switched. Generally the counter-rotating term contributes an amount comparable to the co-rotating term only when the intermediate state detuning is on the order of the transition resonant frequencies.

To gain insight into the results derived in this subsection, it is helpful to evaluate Ω_{R} for a given atom with a given beam configuration, which will be done for ^{87}Rb in the following subsections.

2.6.4 Simple Non-Degenerate Rubidium Example

As an example, the Raman coupling internal state matrix element will be calculated for the transition used in the experimental Raman cooling work presented in this thesis. In particular, the Raman coupling will be between the $|F = 2, m_F = -2\rangle$ and $|2, -1\rangle$ states of the ^{87}Rb $5S_{1/2}$ state, which play the roles of $|g_1\rangle$ and $|g_2\rangle$ respectively. The beams driving the Raman transition are non-degenerate which implies that only

one beam can drive the up-leg and the other can only drive the down-leg[14]. The quantization axis will be taken to be along the vertical z-direction as the magnetic field points along that direction and one of the beams propagates along that direction. The up-leg beam is a circularly polarized beam propagating in the horizontal plane with a 500 μm beam width at the atoms. The down-leg beam is purely σ^--polarized and propagates in the z-direction. Its cross section is elliptical with a 30 μm waist in one direction and a 1 mm waist in the other direction. The Raman transition Rabi frequency will be calculated at the center of the intersection of the beams for the case that both of those beams are set to their full power settings, which are 51.6 μW and 10.2 mW for the up-leg and down-leg beams respectively.

Alkalis have large oscillator strengths for their D_1 and D_2 lines and small oscillator strengths for other transitions. For this reason we can generally make the approximation that the other transitions do not contribute and only sum over intermediate states on the D_1 and D_2 lines. In the experiment of Chapters 4 and 5 in particular, the Raman transition light was 4.33 GHz red of the $|5S_{1/2};2,-2\rangle \rightarrow |5P_{1/2};2',-2\rangle$ transition which is very small compared the transition frequency ∼380 THz. Therefore the rotating wave approximation is very accurate. The use of the rotating wave approximation is further justified by the fact that the beams are non-degenerate so the counter-rotating path is far from the two-photon resonance as discussed in Section 2.6.3. The detuning is also small compared to the ∼7 THz splitting between the D_1 and D_2 lines. That allows a further approximation that only the intermediate states in the D_1 line contribute. Given the above approximations and selection rules, only one term is left in the sum of Eqn. 2.38, namely through the intermediate state $|5P_{1/2};2',-2\rangle$. The full path is $|5S_{1/2};2,-2\rangle \xrightarrow{\pi} |5P_{1/2};2',-2\rangle \xrightarrow{\sigma^-} |5S_{1/2};2,-1\rangle$ where the characters over the arrows indicate the type of transition. This gives

$$\Omega_{\mathrm{R}} = \frac{\Omega_\sigma^* \Omega_\pi}{2\Delta_{\mathrm{R}}} \tag{2.43}$$

[14]The other way around is technically also possible, but that two-photon transition is far off-resonant due to the energy difference between the beams and the atomic states, so it is neglected here.

Where Ω_π and Ω_σ are the Rabi frequencies for the up-leg and down-leg respectively.

To proceed, the single-photon Rabi frequencies for each beam must be calculated as demonstrated in Section 2.6.2. We'll start with the up-leg beam. That beam drives the $|5S_{1/2}; 2, -2\rangle \xrightarrow{\pi} |5P_{1/2}; 2', -2\rangle$ transition, which has Clebsch-Gordan coefficient $C_i = -\sqrt{1/3}$ according to Table 16 of Ref. [71]. Its polarization can be written as [15] $(\hat{z} + i\hat{y})/\sqrt{2}$, which can be re-expressed as $\hat{u}_0/\sqrt{2} + (\hat{u}_{-1} + \hat{u}_{+1})/2$. Only the $\hat{u}_0$ component, which drives the π transition, contributes. The other components only drive Raman transitions to other states which are far off-resonant due to the frequency difference between the beams and the energy splittings between the atomic states. For the π-transition the dot product $(\hat{d}_\pi \cdot \hat{\epsilon})$ then comes out to $1/\sqrt{2}$. Those results can be inserted with the beam parameters into Eqn. 2.34 (using the expression for the D_1 line) to give $-(2\pi)4.92$ MHz for the Rabi frequency at the center of the beam. Note that it is usually important to keep track of the negative sign as it indicates which terms in the sum over intermediate states will interfere. Here though there is only one term in the sum, so it will not have an effect.

The single-photon Rabi frequency for the down-leg is calculated in a similar manner. It drives the $|5P_{1/2}; 2', -2\rangle \xrightarrow{\sigma^-} |5S_{1/2}; 2, -1\rangle$ leg, which has Clebsch-Gordan coefficient $C_i = -\sqrt{1/6}$ according to Table 17 of Ref. [71]. Its polarization is pure σ^-, so $(\hat{d}_{\sigma^-} \cdot \hat{\epsilon}) = 1$. The Rabi frequency at the center of the beam can then be calculated by calculating the the peak intensity of its elliptic beam profile and using Eqn. 2.31. As a quick algebraic trick, the peak intensity is the same as it would be for a round beam profile with a waist equal to the geometric mean of the elliptic beam profile waists, which is $\sqrt{(30\ \mu\text{m})(1\ \text{mm})} = 173\ \mu$m, and that can be used with Eqn. 2.34. Doing so gives $-(2\pi)200$ MHz for the Rabi frequency at the center of the beam.

Now that the single-photon transition Rabi frequencies have been calculated, they can be plugged into Eqn. 2.43. This gives the final result $\Omega_R = -(2\pi)113$ kHz at the center of the beams. That value can be reduced by turning down the power in either

[15] The beam actually propagates slightly off of the x-axis in the $x-y$-plane and so the circular polarization isn't quite in the $y-z$-plane. Therefore there are actually small phases that should be included between the components in that superposition, but those phases do not affect the calculated Rabi frequency and are neglected here for simplicity. This is effectively equivalent to solving the problem in a rotated coordinate system which has its x-axis along the beam direction.

of the beams. More specifically, the value of Ω_R scales with the square root of the power in each beam, so reducing the power in either beam by a factor of four will cut Ω_R by a factor of two. That scaling can also be used to calculate Ω_R away from the centers of the beams using the formulas for a Gaussian beam profile.

2.6.5 Non-Degenerate Rubidium Example, Large Detuning

As a second, slightly more complicated example, this section contains a calculation of the Raman coupling internal state matrix element for beams with a detuning comparable to the splitting between the D_1 and D_2 lines Δ_{FS}. The purpose of this section is to show the near cancellation between the contributions of those two lines. This section will calculate Ω_R assuming that the beams are non-degenerate, as was the case in Section 2.6.4. It will also use the $|2,-2\rangle$ and $|2,-1\rangle$ states as $|g_1\rangle$ and $|g_2\rangle$ respectively, as was done in Section 2.6.4. For simplicity it will be assumed that the up-leg beam is purely π-polarized and the down-leg beam is purely σ^--polarized so that the $(\hat{d}_i \cdot \hat{\epsilon})$ dot products are one. Because the emphasis will be on observing the near cancellation between the D_1 and D_2 lines, rather than explicitly calculating Ω_R for a specific experimental setup, the beam intensities I_{up} and I_{down} will be left as arbitrary symbolic values. It will turn out that they only enter the results through their geometric mean $\overline{I} = \sqrt{I_{up} I_{down}}$.

As in the previous section, we will only include excided states from the D_1 and D_2 lines in the sum in Eqn. 2.35 as the oscillator strengths for other transitions are much weaker. We will make the rotating wave approximation as well, which is justified given that the beams are not degenerate (see Section 2.6.3), and use Eqn. 2.38. The possible paths given these approximations are listed in Table 2.2.

The sum in Eqn. 2.38 can be split into two parts, one part for the states of the D_1 line and one for the states in the D_2 line, then common terms can be factored out.

Path	$C_{i,\text{up}}$	$C_{i,\text{down}}$
$\lvert 5S_{1/2};2,-2\rangle \xrightarrow{\pi} \lvert 5P_{1/2};2',-2\rangle \xrightarrow{\sigma^-} \lvert 5S_{1/2};2,-1\rangle$	$-\sqrt{1/3}$	$-\sqrt{1/6}$
$\lvert 5S_{1/2};2,-2\rangle \xrightarrow{\pi} \lvert 5P_{3/2};2',-2\rangle \xrightarrow{\sigma^-} \lvert 5S_{1/2};2,-1\rangle$	$-\sqrt{1/6}$	$-\sqrt{1/12}$
$\lvert 5S_{1/2};2,-2\rangle \xrightarrow{\pi} \lvert 5P_{3/2};3',-2\rangle \xrightarrow{\sigma^-} \lvert 5S_{1/2};2,-1\rangle$	$-\sqrt{1/6}$	$+\sqrt{1/3}$

Table 2.2: Possible paths from $|5S_{1/2};2,-2\rangle$ to $|5S_{1/2};2,-1\rangle$ that must be accounted for when calculating Ω_R for the hypothetical beam configuration described in Section 2.6.5. Each path corresponds to a term in Eqn. 2.38. The Clebsch-Gordan coefficients for the up-leg $C_{i,\text{up}}$ and down-leg $C_{i,\text{down}}$ are included for each path, taken from Ref. [71]. The reduced matrix element for paths through $|5P_{3/2}\rangle$ and $|5P_{1/2}\rangle$ differ by a factor of $\sqrt{2}$, which must be accounted for when summing over paths.

Doing so and inserting Eqn. 2.31 yields

$$\begin{aligned}\Omega_\text{R} = &\left|\langle J=1/2\|er\|J'=1/2\rangle\right|^2 \frac{\overline{I}}{c\epsilon_0\hbar^2}\sum_{D_1}\frac{C^*_{i,\text{down}}C_{i,\text{up}}}{\Delta_{\text{R},i}} \\ &+\left|\langle J=1/2\|er\|J'=3/2\rangle\right|^2 \frac{\overline{I}}{c\epsilon_0\hbar^2}\sum_{D_2}\frac{C^*_{i,\text{down}}C_{i,\text{up}}}{\Delta_{\text{R},i}}\end{aligned} \tag{2.44}$$

That can be further factored using the fact that, as noted in Section 2.6.2, the reduced matrix element for the D_2 line is a factor of $\sqrt{2}$ larger than that for the D_1 line for the normalization convention used in Ref. [71].

$$\Omega_\text{R} = \frac{\overline{I}\left|\langle J=1/2\|er\|J'=1/2\rangle\right|^2}{c\epsilon_0\hbar^2}\left(\sum_{D_1}\frac{C^*_{i,\text{down}}C_{i,\text{up}}}{\Delta_{\text{R},i}} + 2\sum_{D_2}\frac{C^*_{i,\text{down}}C_{i,\text{up}}}{\Delta_{\text{R},i}}\right) \tag{2.45}$$

Lastly, we can make the approximation that the detuning is the same for all of the states within one of the D lines, namely Δ_{D1} for the D_1 line and Δ_{D2} for the D_2 line. This is a reasonable approximation when the detuning is large compared to the hyperfine splitting in the excited states [16], which should be the case because we are calculating Ω_R in the limit that the detuning is large compared to Δ_FS in this

[16] For the particular calculation in this subsection, that approximation is exact for the D_1 line sum since it only has one term. For other beam configurations though multiple D_1 transition hyperfine states can contribute, in which case this approximation is not exact.

subsection. Making this approximation yields

$$\Omega_{\mathrm{R}} = \frac{\overline{I}\left|\langle J=1/2\|er\|J'=1/2\rangle\right|^2}{c\epsilon_0\hbar^2}\left(\frac{1}{\Delta_{D1}}\sum_{D_1} C^*_{i,\mathrm{down}}C_{i,\mathrm{up}} + \frac{2}{\Delta_{D2}}\sum_{D_2} C^*_{i,\mathrm{down}}C_{i,\mathrm{up}}\right) \tag{2.46}$$

At this point the Clebsch-Gordan coefficients from Table 2.2 can be inserted into the sums, giving

$$\begin{aligned}\Omega_{\mathrm{R}} &= \frac{\overline{I}\left|\langle J=1/2\|er\|J'=1/2\rangle\right|^2}{c\epsilon_0\hbar^2}\left(\frac{1}{\Delta_{D1}}\left(\frac{1}{3\sqrt{2}}\right) + \frac{2}{\Delta_{D2}}\left(-\frac{1}{6\sqrt{2}}\right)\right) \\ &= \frac{\overline{I}\left|\langle J=1/2\|er\|J'=1/2\rangle\right|^2}{3\sqrt{2}c\epsilon_0\hbar^2}\left(\frac{1}{\Delta_{D1}} - \frac{1}{\Delta_{D2}}\right)\end{aligned} \tag{2.47}$$

The scaling of Ω_{R} with the detuning in Eqn. 2.47 is worth discussion. At detunings small compared to Δ_{FS}, either the Δ_{D1} or Δ_{D2} term dominates and the other is negligible. Then Ω_{R} scales as $\overline{I}/\Delta_{D1}$ or $\overline{I}/\Delta_{D2}$ depending on which is closer to resonance. In this case, the scaling with the inverse of detuning is the same scaling as for the trap depth of a far off-resonant dipole trap. This isn't so surprising because both effects come from second order perturbation theory with the same Hamiltonian. The main difference is that the initial and final state are the same when calculating the trap depth due to the AC Stark shift, while the initial and final state are different when calculating the Raman coupling Rabi frequency. The other difference is that the legs of the Raman transition are driven by different beams while for a dipole trap both legs are driven by one beam (hence why $\overline{I}$ appears instead of just the I of one beam).

However, at detunings large compared to Δ_{FS}, the scaling of Ω_{R} changes. First of all, as can be seen from Eqn. 2.47, Ω_{R} would be zero if Δ_{FS} were zero. Therefore a splitting between the D_1 and D_2 lines is necessary for the Raman coupling. Using

the relation $\Delta_{D2} = \Delta_{D1} + \Delta_{\mathrm{FS}}$, the scaling can be written as

$$\begin{aligned}\Omega_{\mathrm{R}} &\propto \overline{I}\left(\frac{1}{\Delta_{D1}} - \frac{1}{\Delta_{D1}+\Delta_{\mathrm{FS}}}\right) \\ &= \frac{\overline{I}}{\Delta_{D1}}\left(1 - \frac{1}{1+\Delta_{\mathrm{FS}}/\Delta_{D1}}\right) \\ &\approx \frac{\overline{I}}{\Delta_{D1}}\left(\frac{\Delta_{\mathrm{FS}}}{\Delta_{D1}}\right)\end{aligned} \tag{2.48}$$

In this limit the Raman coupling rate is suppressed by a factor of $\Delta_{\mathrm{FS}}/\Delta_{D1}$ relative to the $\overline{I}/\Delta_{D1}$ scaling. Put another way, it scales as $\overline{I}/\Delta_{D1}^2$ for $|\Delta_{D1}| \gg \Delta_{\mathrm{FS}}$, which is the same scaling as for the scattering rate. This has a few consequences. First of all, a larger intensity is needed to achieve a given Rabi frequency when $|\Delta_{D1}| \gtrsim \Delta_{\mathrm{FS}}$ than would be expected from the $\overline{I}/\Delta_{D1}$ scaling nearer to resonance. Additionally, the ratio of Ω_{R} to the rate of undesired scattering of light from the Raman beams saturates as $|\Delta_{D1}|$ is increased past Δ_{FS}. Furthermore, if one were to increase $|\Delta_{D1}|$ while increasing $\overline{I}$ to keep the Rabi frequency constant, then the light shift due to the Raman coupling beams would start to increase once $|\Delta_{D1}|$ becomes comparable to Δ_{FS}. Before performing the work presented in Ch. 4, we attempted to perform Raman cooling using the same 1064 nm light used to generate the dipole traps to drive the Raman transitions as well, but we were unable to produce a BEC without evaporation. However, after changing the system to use the 795 nm light to perform the Raman coupling, we were able to produce a BEC. One possible explanation for this is that the trap depth created by the 1064 nm Raman coupling light was too large and led to excess Oort cloud heating as discussed in Section 2.5.6, and then that was not an issue for the 795 nm light. This could be the case because the detuning of the 1064 nm light is $\Delta_{D1} \approx -(2\pi)95$ THz which is large compared to $\Delta_{\mathrm{FS}} \approx (2\pi)7$ THz [71] while the $\Delta_{D1} \approx -(2\pi)4.33$ GHz detuning of the 795 nm light is much smaller than Δ_{FS}.

The suppression of the Raman coupling rate for detunings large than $\sim\Delta_{\mathrm{FS}}$ isn't specific to this beam configuration or to rubidium; it is a more general property of these transitions. The hand-waving explanation for this is as follows. The electron

has spin 1/2, so flipping it changes its contribution to m_F from $-1/2$ to $+1/2$ (or vice versa), thereby changing m_F by one. Therefore, just flipping the electron's spin is sufficient to drive the transition between two states that differ in m_F by 1. However in order to flip the electron's spin, one must be able to resolve different electron states (e.g. $J = 1/2$ or $J = 3/2$) spectroscopically, i.e. one must be able to preferentially drive the D_1 or D_2 line which are split by Δ_{FS}. For detunings large compared to Δ_{FS}, the electron's state isn't resolved and can't be changed. That is why the ability to drive the Raman transitions with $\Delta m_F = \pm 1$ falls off faster for detunings large compared to Δ_{FS} than for detunings small compared to Δ_{FS}. A similar argument applies to Raman transitions which attempt to drive $\Delta m_F = \pm 2$ transitions using a σ^+ component and a σ^- component. However in that case flipping the electron spin isn't sufficient to cause the targeted change in m_F; it is necessary to change the nuclear spin as well. Therefore the drive must be able to resolve the hyperfine structure, and the Raman transition Rabi frequency falls off faster once the detuning is comparable to the hyperfine splitting in the excited state. The proofs for these statements are beyond the scope of this thesis, but can be done using identities associated with Clebsch-Gordan coefficients and Wigner $3j$ and $6j$ symbols. Notably this aspect of Raman coupling is important for species which have small Δ_{FS}.

2.7 Raman Coupling External Matrix Elements

The Raman transition used in Raman cooling involves not only a change of the atom's internal state, but of its external state as well. In addition to actually providing the cooling, the external matrix elements can also affect the Raman coupling Rabi frequency. This was somewhat touched upon in the Section 2.6 where it was noted that certain Raman coupling paths provided different momentum kicks than other ones and therefore needed to be summed separately. This section delves into that in more detail, justifying that those sums needed to be treated separately and showing how the external state matrix elements can affect the Raman transition Rabi frequency.

2.7.1 Free-Space Coupling

The $\mathbf{d} \cdot \mathcal{E}$ Hamiltonian for a single beam can be written out to explicitly show the spatial variation of the electric field from Eqn. 2.25. Converting the cosine to a sum of complex exponentials puts the Hamiltonian into a particularly enlightening form

$$\begin{aligned} \mathcal{H} &= d_0 E_0 \left(\hat{d} \cdot \hat{\epsilon}\right) \frac{1}{2} \left(e^{+i(\mathbf{k}\cdot\hat{\mathbf{r}} - \omega_\mathrm{L} t)} + e^{-i(\mathbf{k}\cdot\hat{\mathbf{r}} - \omega_\mathrm{L} t)}\right) \\ &= d_0 E_0 \left(\hat{d} \cdot \hat{\epsilon}\right) \frac{1}{2} \left(e^{-i\omega_\mathrm{L} t} e^{+i\mathbf{k}\cdot\hat{\mathbf{r}}} + e^{+i\omega_\mathrm{L} t} e^{-i\mathbf{k}\cdot\hat{\mathbf{r}}}\right) \end{aligned} \tag{2.49}$$

Hats are put on $\hat{\mathbf{r}}$ to indicate that it is an operator which acts on a state when $\mathcal{H}$ is applied to it. By contrast, $\mathbf{k}$ is *not* an operator; it is just a vector of real numbers describing the propagation direction of the light. This means that the term $\exp(i\mathbf{k}\cdot\hat{\mathbf{r}})$ is the momentum-space translation operator[17], which maps a momentum eigenstate $|\hbar\mathbf{k}_0\rangle$ to another momentum eigenstate $|\hbar(\mathbf{k}_0 + \mathbf{k})\rangle$, i.e. $\exp(i\mathbf{k}\cdot\hat{\mathbf{r}}) = |\hbar(\mathbf{k}_0 + \mathbf{k})\rangle \langle\hbar\mathbf{k}_0|$. This property is easy to show by writing a momentum eigenstate in the position basis, which is $(|\mathbf{r}\rangle \langle\mathbf{r}|) |\hbar\mathbf{k}_0\rangle = \exp(i\mathbf{k}_0 \cdot \mathbf{r}) |\mathbf{r}\rangle$ (up to normalization and with an implicit integral over $\mathbf{r}$). Applying the momentum translation operator to that state gives

$$e^{i\mathbf{k}\cdot\hat{\mathbf{r}}} |\mathbf{r}\rangle \langle\mathbf{r}|\hbar\mathbf{k}_0\rangle = e^{i\mathbf{k}\cdot\mathbf{r}} |\mathbf{r}\rangle \langle\mathbf{r}|\hbar\mathbf{k}_0\rangle = e^{i\mathbf{k}\cdot\mathbf{r}} e^{i\mathbf{k}_0\cdot\mathbf{r}} |\mathbf{r}\rangle = e^{i(\mathbf{k}_0+\mathbf{k})\cdot\mathbf{r}} |\mathbf{r}\rangle = |\hbar(\mathbf{k}_0 + \mathbf{k})\rangle \tag{2.50}$$

Eqn. 2.49 has two terms with momentum-space translation operators: one that translates by $+\mathbf{k}$ and one that translates by $-\mathbf{k}$. The term that translates by $+\mathbf{k}$ has a phase that evolves as $e^{-i\omega_\mathrm{L} t}$, which is the co-rotating term. Similarly, the $-\mathbf{k}$ term evolves as $e^{+i\omega_\mathrm{L} t}$ and is the counter-rotating term. This shows what was discussed in Section 2.6, namely that the co-rotating and counter-rotating terms for a given beam provide opposite momentum kicks. Interestingly it was shown here *without* second quantization; there was no mention of photons or their momenta here.

Raman transitions involve two of these single-photon transitions, each one driven by a different beam with its own $\mathbf{k}$. Therefore the two-photon transition involves two applications of the momentum-space translation operator, which can be expressed as

[17] This is akin to the somewhat more common position-space translation operator $\exp(-i\hat{\mathbf{k}} \cdot \mathbf{r})$ in which $\hat{\mathbf{k}}$ is an operator and $\mathbf{r}$ is a vector of real numbers which determines the displacement.

a single momentum-space translation operator which displaces by the sum of those momentum kicks. For simplicity suppose beam 1 drives the up-leg and beam 2 drives the down-leg. The two possible paths shown by Eqn. 2.40 make it clear what these operators should be. The co-rotating path involves absorbing a $+\hbar\mathbf{k}_1$ momentum kick then receiving a $-\hbar\mathbf{k}_2$ momentum kick so that the total momentum change for the atom is $+\hbar\Delta\mathbf{k}$. Similarly, the counter-rotating path involves receiving a $-\hbar\mathbf{k}_1$ momentum kick then absorbing a $+\hbar\mathbf{k}_2$ momentum kick for a total of $-\hbar\Delta\mathbf{k}$. Therefore the co-rotating path couples $|g_1, \hbar\mathbf{k}_0\rangle$ to $|g_2, \hbar(\mathbf{k}_0 + \Delta\mathbf{k})\rangle$ while the counter-rotating path couples $|g_1, \hbar\mathbf{k}_0\rangle$ to $|g_2, \hbar(\mathbf{k}_0 - \Delta\mathbf{k})\rangle$. Notably, although both of the final states have the same internal state, they are different states overall because they have different external states. In writing Eqn. 2.35 I omitted the external state portion for simplicity. Adding it in shows why the terms that provide different momentum kicks should be summed separately: they couple to different final states.

From Eqn. 2.49 it is clear that the full operator, which involves both the internal and external states, for the Raman coupling with a $+\hbar\Delta\mathbf{k}$ momentum kick has the form

$$\mathcal{H}_{\text{eff},g2,g1} = \frac{\hbar\Omega_{\text{R}}}{2} e^{i\Delta\mathbf{k}\cdot\hat{\mathbf{r}}} |g_2\rangle \langle g_1| = \frac{\hbar\Omega_{\text{R}}}{2} |g_2, \hbar(\mathbf{k}_0 + \Delta\mathbf{k})\rangle \langle g_1, \hbar\mathbf{k}_0| \tag{2.51}$$

where it is assumed that only paths that lead to a $+\hbar\Delta\mathbf{k}$ momentum kick are included in the sum when evaluating Ω_{R} from the internal state matrix elements. In other words the $+\hbar\Delta\mathbf{k}$ subscript shown in Eqn. 2.42 is omitted but implied. This is automatically the case when using Eqn. 2.38 for non-degenerate beams because all of those terms involve absorbing from beam 1 and emitting into beam 2. For configurations with degenerate Raman beams where paths with $-\hbar\Delta\mathbf{k}$ momentum kicks are also possible, the coupling has the form

$$\begin{aligned}\mathcal{H}_{\text{eff},g2,g1} &= \left(\frac{\hbar\Omega_{\text{R},+\Delta\mathbf{k}}}{2} e^{i\Delta\mathbf{k}\cdot\hat{\mathbf{r}}} + \frac{\hbar\Omega_{\text{R},-\Delta\mathbf{k}}}{2} e^{-i\Delta\mathbf{k}\cdot\hat{\mathbf{r}}}\right) |g_2\rangle \langle g_1| \\ &= \frac{\hbar\Omega_{\text{R},+\Delta\mathbf{k}}}{2} |g_2, \hbar(\mathbf{k}_0 + \Delta\mathbf{k})\rangle \langle g_1, \hbar\mathbf{k}_0| + \frac{\hbar\Omega_{\text{R},-\Delta\mathbf{k}}}{2} |g_2, \hbar(\mathbf{k}_0 - \Delta\mathbf{k})\rangle \langle g_1, \hbar\mathbf{k}_0|\end{aligned} \tag{2.52}$$

The generalization to include other paths with other possible momentum kicks, e.g. due to the presence of additional beams, is straightforward. Simply calculate the internal state matrix elements for each possible momentum kick separately as discussed in Section 2.6, the multiply each result with the corresponding momentum-space displacement operator. This explicitly shows what was mentioned in the Section 2.6, namely that paths with different momentum kicks should be summed separately.

For a free atom, the eigenstates are simply products of internal states and momentum eigenstates. Eqn. 2.51 is written in that basis, so its couplings in that basis are very clear. An atom in $|g_1\rangle$ with an arbitrary initial external state $|\hbar\mathbf{k}_0\rangle$ is coupled only to the final state $|g_2, \hbar(\mathbf{k}_0 + \Delta\mathbf{k})\rangle$ and that coupling has strength $\hbar\Omega_{\mathrm{R}}/2$. So for a free atom the Raman coupling Rabi frequency is exactly equal to the value of the Raman coupling calculated for the internal matrix elements. The external state is "just along for the ride" in some sense because it doesn't affect the coupling strength between the initial and final states. In the next section it will be shown that this is emphatically *not* generally the case for a trapped atom.

2.7.2 Coupling in a Harmonic Trap

For an atom in a harmonic trap, the external state eigenstates in the absence of Raman coupling are the well-known harmonic trap eigenstates. Those form a complete basis for the external states which can be used in lieu of the momentum eigenstate basis. It is generally more convenient to use the harmonic trap eigenstate basis instead of the momentum eigenstate basis when and atom is in a harmonic trap with trap frequency large compared to the Raman coupling internal state matrix element. On the other hand it is generally more convenient to use the momentum eigenstate basis for untrapped atoms or atoms in a harmonic trap with trap frequency small compared to the Raman coupling internal state matrix element.

From the above, the full Raman coupling matrix element in the harmonic state

eigenbasis is

$$\left\langle g_2, \mathbf{m} \right| \left(\frac{\hbar\Omega_{\mathrm{R}}}{2} \left|g_2\right\rangle \left\langle g_1\right| \right) e^{i\Delta k_x \hat{x}} \left| g_1, \mathbf{n} \right\rangle = \frac{\hbar\Omega_{\mathrm{R}}}{2} \left\langle \mathbf{m} \right| \exp\left(i\Delta\mathbf{k}\cdot\hat{\mathbf{r}}\right) \left| \mathbf{n} \right\rangle \tag{2.53}$$

where e.g. $|\mathbf{n}\rangle = |n_x, n_y, n_z\rangle$ is an eigenstate of the 3D harmonic oscillator with each n_i indexed from zero so that the ground state for each direction is $|0\rangle$. The actual Raman coupling frequency is then the internal matrix element two-photon Rabi frequency multiplied by a numerical factor due to the external state matrix element. The external state matrix elements for Raman coupling in a harmonic trap are of the form

$$\mathcal{H}_{\mathbf{m},\mathbf{n}} = \left\langle \mathbf{m} \right| e^{i\Delta\mathbf{k}\cdot\hat{\mathbf{r}}} \left| \mathbf{n} \right\rangle \tag{2.54}$$

Due to the dot product, each component of $\exp\left(i\Delta\mathbf{k}\cdot\hat{\mathbf{r}}\right)$ acts only on its own direction (i.e. the x-component of $\Delta\mathbf{k}$ only affects the x-direction), so it is sufficient to treat each component separately. For concreteness we will consider the x-direction which will be taken to be along one of the trap's eigenaxes. For simplicity we'll drop the subscripts on the harmonic oscillator states; the math for the other directions will then be the same. The problem effectively becomes 1D, so we need to evaluate matrix elements of the form

$$\mathcal{H}_{m,n} = \left\langle m \right| e^{i\Delta k_x \hat{x}} \left| n \right\rangle \tag{2.55}$$

where Δk_x is the component of $\Delta\mathbf{k}$ along the direction of interest. The derivation for that matrix element is provided in Ref. [72], which evaluates it using the Baker-Campbell-Hausdorff Lemma. This gives the result which can be written in terms of a generalized Laguerre polynomial as (see Eqn.(31) of Ref. [72])

$$\left\langle m \right| \exp\left(i\Delta k_x \hat{x}\right) \left| n \right\rangle = e^{-\eta_{\mathrm{R}}^2/2} \left(i\eta_{\mathrm{R}}\right)^{|\Delta n|} \left(\sqrt{\frac{n_<!}{(n_< + |\Delta n|)!}} \right) L_{n_<}^{|\Delta n|}\left(\eta_{\mathrm{R}}^2\right) \tag{2.56}$$

where $\Delta n = m - n$ is the change in harmonic trap level index[18], $n_< = \min(m, n)$ is the lesser of m and n, $\eta_{\mathrm{R}} = \Delta k_x x_0$ is the Lamb-Dicke parameter for the Raman

[18]Note that the definition for Δn here is slightly different than the one used in Ref [72], differing by the presence or absence of an absolute value. Here $\Delta n = m - n$ but there $\Delta n = |m - n|$.

transition along the x-axis [19], $x_0 = \sqrt{\hbar/(2m\omega_x)}$ is the harmonic oscillator length, m is the atomic mass, and ω_x is the angular trap frequency in the x-direction. For reference, the generalized Laguerre polynomial is defined as[20]

$$L_b^a(x) = \sum_{j=0}^{b} \frac{(-x)^j}{j!} \binom{a+b}{b-j} \tag{2.57}$$

where the large parentheses denote a binomial coefficient.

Eqn. 2.56 is somewhat complicated so it is worth discussing some of its properties. Firstly, the matrix elements are generally less than one. That means that they can actually suppress the Raman coupling rate and make it less than the value calculated from the internal matrix elements alone. This is different than the free-space result and is due to the atom's velocity oscillating in the trap and affecting the velocity-selective Raman coupling. Another point worth noting is that all of the experimental parameters enter the equation only through the dimensionless parameter η_{R} which is a measure of how strong the momentum kick is compared to how strongly the trap confines the atom. Its magnitude can also be expressed as $\sqrt{E_{\mathrm{r}\Delta\mathbf{k}}/(\hbar\omega_x)}$, so it is the square root of the ratio of a $\Delta\mathbf{k}$ recoil kick energy to the energy spacing between harmonic trap levels. It thus captures how hard it is for the given momentum kick to drive changes between harmonic oscillator states. Note that the actual change in the atom's energy can be very different than $E_{\mathrm{r}\Delta\mathbf{k}}$ because it is non-stationary while it is in the trap, as demonstrated Eqn. 2.2.

Note that as defined here, η_{R} is positive if Δk_x is positive and negative is Δk_x is negative. The matrix elements for $+\Delta k_x$ and $-\Delta k_x$ have the same sign if $|\Delta n|$ is even but have the opposite sign if $|\Delta n|$ is odd due to the $\eta_{\mathrm{R}}^{|\Delta n|}$ factor. Therefore the different paths discussed in Section 2.6 which produce opposite $\Delta\mathbf{k}$ will add with the same sign for transitions that involve changing the trap level by an even number, but will add with opposite signs for transitions that change the trap level by an odd number.

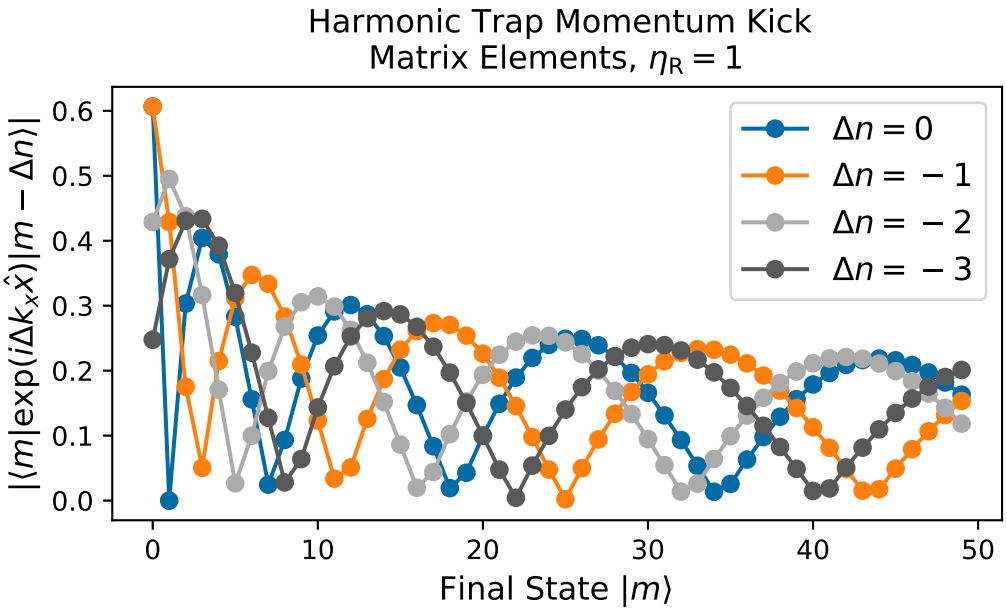
[19] Note that the momentum kick and trap frequency can be different for different directions, so each direction effectively has its own Lamb-Dicke parameter.

[20] Note that Ref. [72] has a small typo in their definition. The lower index on the Laguerre L there should be n, not m.

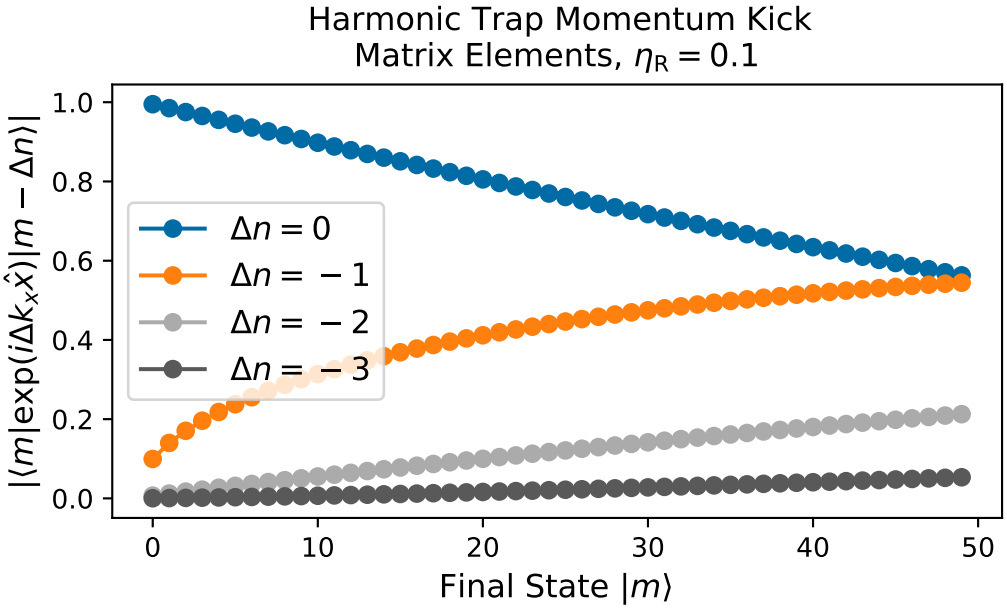
This can lead to cancellations between those terms. In particular, if the two beams driving a Raman transition are degenerate and have the same polarizations, then the path where beam 1 drives the up-leg and the path where beam 2 drives the up-leg will have the same internal state matrix elements but provide opposite momentum kicks. That leads to a complete cancellation for odd Δn transitions. When cooling with unresolved sidebands, this is not much of a problem because the transitions to adjacent states involving even Δn then have twice the matrix element amplitude because the two paths add instead of cancel. However, when cooling with resolved sidebands this prohibits cooling on transition with odd Δn, including the $n \rightarrow (n-1)$ transition required to cool atoms in the $|1\rangle$ state (the ground state is $|0\rangle$). This is a reflection of a general property of Raman transitions: highly symmetric setups often lead to cancellations. Simply rotating the polarization of one of the beams can avoid this cancellation by changing the $\hat{d} \cdot \hat{\epsilon}$ term in the single-photon Rabi frequencies for each path so that the internal state matrix elements are no longer equal. Doing so does not guarantee Raman coupling though, as other cancellations may occur.

The magnitude of the matrix elements given by Eqn. 2.56 are plotted in Fig. 2-8 for various parameters with curves for various values of Δn. Fig. 2-8(a) shows matrix elements with $\eta_{\mathrm{R}} = 1$ and Fig. 2-8(b) shows results with $\eta_{\mathrm{R}} = 0.1$. Each curve shows the magnitude of the matrix elements for a given Δn, and the x-axis represents the final state of that transition. For example the value of the orange $\Delta n = -1$ curve point at 10 on the x-axis is $|\langle 10| \exp(i\Delta k_x \hat{x})|11\rangle|$. When $\eta_{\mathrm{R}} \gtrsim 1$, many different Δn transitions generally have non-negligible matrix elements. However, when $\eta_{\mathrm{R}} \ll 1$ the matrix elements with nonzero Δn are suppressed and transitions with larger Δn are more suppressed, particularly for small $|m\rangle$. This is a consequence of the intuitive fact that a very weak momentum kick will have a hard time moving the atom to a different harmonic oscillator state. The fact that this is less of a problem for more highly excited harmonic oscillator states will become more clear soon.

Fig. 2-9 shows the matrix elements between fixed initial states (one per curve) and a range of final states. Fig. 2-9(a) shows matrix elements with $\eta_{\mathrm{R}} = 1$ and Fig. 2-9(b) shows results with $\eta_{\mathrm{R}} = 0.1$. The plots demonstrate which final states a given initial

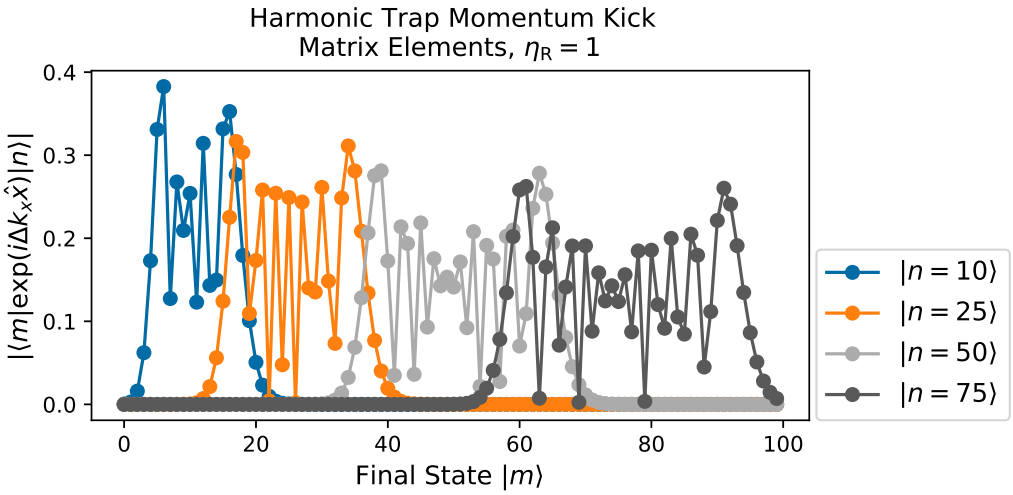


(a)

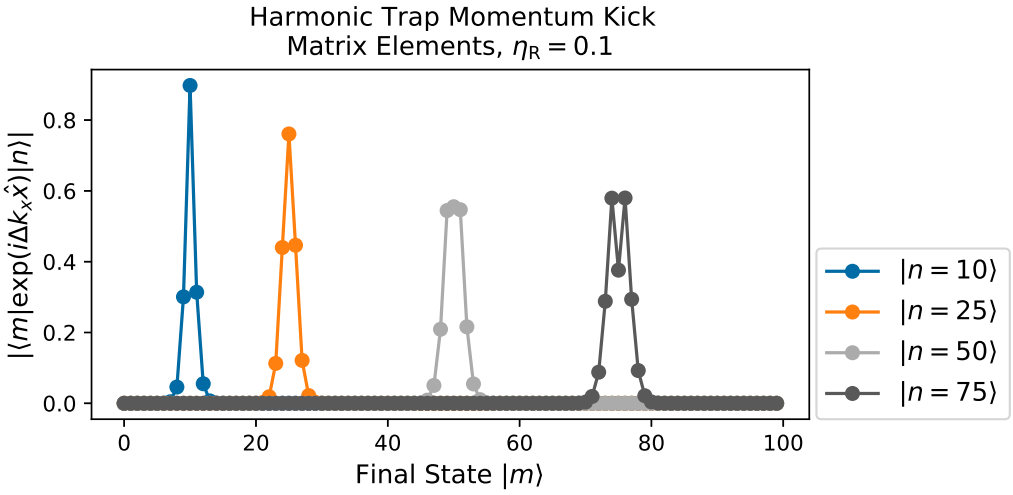


(b)

Figure 2-8: Magnitudes of the matrix elements from Eqn. 2.56 for (a) $\eta_R = 1$ and (b) $\eta_R = 0.1$. The x-axis shows the harmonic oscillator level of th final state and curves are plotted for various Δn. For example, the orange curve is for $\Delta n = -1$ so it represents the magnitude of the matrix element between $|m+1\rangle$ and $|m\rangle$. The matrix elements are defined only at discrete values marked with circles; the lines are included only as guides to the eye.



(a)



(b)

Figure 2-9: Magnitudes of the matrix elements from Eqn. 2.56 for (a) $\eta_R = 1$ and (b) $\eta_R = 0.1$ from given initial states. The x-axis shows the final state's harmonic oscillator level and curves are plotted for various Δn. For example, the orange curve shows the magnitude of matrix elements from the initial state $|n = 25\rangle$. The matrix elements are defined only at discrete values marked with circles; the lines are included only as guides to the eye.

state couples to strongly. Each initial state mainly couples to states nearby it which is why the region of large values moves towards larger $|m\rangle$ as $|n\rangle$ is increased. Matrix elements between states with large enough Δn are suppressed, but there is a region of $|m\rangle$ where matrix elements are not generally suppressed. The range of strongly connected final states is much smaller for smaller η_{R}. Again this makes sense because a small momentum kick will have a hard time significantly changing the harmonic trap level.

Notably the region where the matrix elements are not generally suppressed in Fig. 2-9 becomes wider for larger $|n\rangle$. The reason for this can actually be understood using Eqn. 2.2. That equation shows that the amount of kinetic energy added to or removed from the atom is approximately equal to $\hbar\delta_{\mathrm{D}}$ for large $|n\rangle$ where the Doppler shift is large compared to the $\Delta\mathbf{k}$-recoil energy. The energy of harmonic oscillator state $|n\rangle$ is $(n+1/2)\hbar\omega_x \approx n\hbar\omega_x$. For a classical particle in a harmonic oscillator, that energy oscillates back and forth from kinetic energy to potential energy. The maximum momentum that the particle has during an oscillation is then set by its total energy $p_{\mathrm{m}}^2/(2m) = E_{\mathrm{tot}}$. It should not be surprising then that the probability that a quantum particle in that harmonic oscillator state has a momentum larger than $p_{\mathrm{m}} = \sqrt{2mn\hbar\omega_x}$ (which satisfies $p_{\mathrm{m}}^2/(2m) = n\hbar\omega_x$) is suppressed. Because the Doppler shift is proportional to the atom's momentum, this sets a soft limit on the maximum amplitude of the Doppler shift $|\delta_{\mathrm{D}}| \lesssim (p_{\mathrm{m}}/m)\Delta k_x$. That then limits the amount of energy that can be added or removed during the Raman transition, which limits Δn. To be clear, the energy can be either added or removed. One can imagine this as being due to the atom oscillating along $\Delta\mathbf{k}$ then against $\Delta\mathbf{k}$, flipping the sign on its Doppler shift. Of course for a harmonic oscillator eigenstate the phase of that oscillation is completely uncertain, so essentially both shifts occur simultaneously in superposition. Because the harmonic trap levels differ in energy by multiples of $\hbar\omega_x$, the approximate maximum change in trap level can be calculated by dividing the max energy exchanged by $\hbar\omega_x$. After some simple algebra, this gives $|\Delta n| \lesssim (2\sqrt{n})\eta_{\mathrm{R}}$. This shows that the region of unsuppressed matrix elements scales linearly with η_{R} and with the square root of the trap level, which explains why the region is wider at

larger $|n\rangle$. For $\eta_R = 1$ and $n = 75$, this gives $|\Delta n| \lesssim 17$, so that state should couple mainly to states from $|61\rangle$ to $|92\rangle$, which agrees reasonably well with the dark gray curve in Fig. 2-9(a).

It is also worth pointing out some of the behavior of Eqn. 2.56 at small η_R, which is referred to as the Lamb-Dicke regime. In that limit, the value of the Laguerre polynomial is dominated by its zeroth order term, which gives a constant factor that depends only on m and n. Ignoring constant factors and factors which only depend on m and n gives the following approximate scaling

$$\langle m \,|\, \exp\left(i\Delta k_x \hat{x}\right) |\, n\rangle \propto e^{-\eta_R^2/2} \eta_R^{|\Delta n|} \quad \text{for } \eta_R \ll 1 \tag{2.58}$$

This shows that matrix elements for transitions with sequentially larger changes in trap state number are each suppressed by an additional factor [21] of η_R when $\eta_R \ll 1$. In other words the matrix element for a $\Delta n = 2$ transition is roughly a factor of η_R smaller than the matrix element for a $\Delta n = 1$ transition. This is often important for Raman sideband cooling which typically involves fairly tight confinement and therefore small η_R. In that case, the Raman coupling rate is slowed by this effect. Tuning the system to cool on the first red sideband (i.e. $\Delta n = -1$) will result in faster Raman coupling than if the system were tuned to the second red sideband ($\Delta n = -2$), but even the first red sideband is suppressed by a factor of η_R.

For Raman cooling with unresolved sidebands, i.e. in the free-space limit, the external state matrix elements still affect the Raman coupling Rabi rate, but to a lesser extent. Looking back at Fig. 2-8(a), there are several states for which the matrix element for a $\Delta n = -1$ transition is close to zero, which may make it seem like the cooling would not work there. However, for the states where the $\Delta n = -1$ matrix element is nearly zero, the other $\Delta n = -2$ and/or $\Delta n = -3$ matrix elements are much larger. In the free-space limit those transitions are not spectrally resolved, so the cooling proceeds through those other couplings. This makes it more akin to

[21] The ratio of the matrix elements for different Δn will generally also depend on m and n through the factors that were dropped here, but for small enough η_R the majority of the change in matrix element for different Δn will be due to the $\eta_R^{|\Delta n|}$ factor.

truly free-space cooling, i.e. cooling without a trap, where there is no suppression of the Raman coupling due to external state matrix elements.

2.7.3 Lamb-Dicke Regime

As mentioned in the previous subsection, the Lamb-Dicke regime is the limit $\eta_R \ll 1$. In this limit, transitions with large Δn are suppressed due to their small matrix elements. The Lamb-Dicke regime often goes hand-in-hand with the resolved-sideband regime, however they are different and it is possible to be in one of those regimes without being in the other. For example, the work in Ref. [46] demonstrated resolved Raman sideband cooling outside the Lamb-Dicke regime.

Interestingly the authors of Ref. [46] had to go out of their way to drive various Δn rather than only driving e.g. $\Delta n = -1$ transitions. This was done to avoid getting stuck in states which do not strongly couple with a $\Delta n = -1$ transition. Such states can be seen to exist in Fig. 2-8(a) which shows the matrix element amplitudes for $\eta_R = 1$. For example, the coupling between $|25\rangle$ and $|26\rangle$ is $\approx 2 \times 10^{-3}$ there. Driving other Δn transitions makes the cooling somewhat similar to cooling in the unresolved-sideband regime since that entails driving multiple Δn transitions as well. Driving transitions other than $\Delta n = -1$ typically isn't required when in the Lamb-Dicke regime. As can be seen from Fig. 2-8(b), none of the first 50 states have a $\Delta n = -1$ coupling near zero for $\eta_R = 0.1$. States like that do exist but only at much higher trap levels. Because the Lamb-Dicke regime is generally associated with large trap frequencies, even modest initial temperatures can result in atoms primarily populating lower states initially, thus avoiding the issue altogether. This is one advantage of cooling in the Lamb-Dicke regime.

Thus far the discussion has focused on the external matrix elements in a harmonic trap due to the momentum kick during a stimulated Raman transition. However, the momentum kicks during the optical pumping lead to essentially the same matrix elements. The only difference is that the $\Delta\mathbf{k}$ are not the same for the Raman transition and the optical pumping transition, so the Lamb-Dicke parameters are different. Part of the reason that the $\Delta\mathbf{k}$ differ is that often (but not always) different wavelengths of

light are used to drive those two processes. The other reason that they are different is that the wavevector of the emitted photon during the stimulated Raman transition is controlled because it is emitted into one of the Raman coupling beams. However, during optical pumping the scattered photon is emitted into a random direction. For this reason it is often easier to treat the recoils due to absorption and emission during optical pumping separately. Thus we can define the Lamb-Dicke factor for the optical pumping process as $\eta_{\mathrm{p}} = |\mathbf{k}_{\mathrm{pump}}|\, x_0$ for the x-direction, and similar for the other directions. As before, the problem can be treated separately along each trap eigenaxis and I'll omit the direction subscript on the Lamb-Dicke parameter for notational simplicity.

Of course as with the stimulated Raman transition, only the component of $\mathbf{k}_{\mathrm{pump}}$ in the x-direction contributes to the coupling between harmonic oscillator states in that direction, so that projection must be included before calculating any matrix elements. The momentum kick due to the photon absorption is set by the optical pumping beam's wavevector and so we could have defined η_{p} similarly to η_{R} in the sense that we could have accounted for its projection along the x-direction and then defined η_{p} to be $k_{\mathrm{pump},x} x_0$. However, the spontaneously emitted photon's direction is random so its projection along x is not fixed and instead one must average over its emission pattern. Therefore it is easier to just specify η_{p} as $|\mathbf{k}_{\mathrm{pump}}|\, x_0$ for the x-direction and keep in mind that one must account for the projection of the wavevector onto the x-axis when calculating matrix elements. The value of η_{p} is then the maximum value that η can take in a given direction, which occurs in the case that the photon is emitted along that axis as that maximizes the recoil momentum projection along the axis. Defining the Lamb-Dicke parameter this way is also convenient because then it can be re-expressed as $\eta_{\mathrm{p}} = \sqrt{E_{\mathrm{r}}/(\hbar\omega_x)}$ where $E_{\mathrm{r}} = (\hbar\mathbf{k}_{\mathrm{pump}})^2/(2m)$ is the recoil energy due to one optical pumping photon.

Although this doesn't necessarily have to be the case, typically $|\Delta\mathbf{k}| \sim |\mathbf{k}_{\mathrm{pump}}|$ since both processes are generally driven by optical photons. That implies that if $\eta_{\mathrm{R}} \ll 1$ then it is usually also the case that $\eta_{\mathrm{p}} \ll 1$. Therefore often there is no distinction made between being in the Lamb-Dicke regime for the Raman coupling or

being in the Lamb-Dicke regime for the optical pumping transition. Unless otherwise specified, it is usually the case that cooling in the Lamb-Dicke regime implies both $\eta_{\mathrm{R}} \ll 1$ and $\eta_{\mathrm{p}} \ll 1$.

The consequences of being in the Lamb-Dicke regime are somewhat different for the optical pumping step than for the Raman coupling step. The issue of certain Δn transitions potentially having very small matrix elements is usually not relevant for the optical pumping for neutral atoms. This is because the optical pumping can proceed with a $\Delta n = 0$ transition, which isn't suppressed when $\eta_{\mathrm{p}} \ll 1$.

A common misconception about that Lamb-Dicke regime with regards to optical pumping is that people often think that it reduces the average amount of recoil heating during a scattering. This is incorrect. The average heating during a scattering is still $2E_{\mathrm{r}}$ as it would be in free space, with one E_{r} due to the absorption recoil and one due to the emission recoil. What does change in the Lamb-Dicke regime is how often an atom which receives a recoil kick stays in the same trap level despite having received a recoil kick. That happens much more frequently in the Lamb-Dicke regime than outside the Lamb-Dicke regime and is essentially the same phenomenon as that of the Mössbauer effect [73]. The explanation for this behavior is fairly simple. Consider just one of the momentum kicks during a scattering for now, either from the absorption or the emission of a photon. When η_{p} isn't small, a recoil kick will generally change the atom's trap state by many levels. This is demonstrated by the large width of the curves in Fig. 2-9(a) where the Lamb-Dicke parameter is one. Now suppose $\eta_{\mathrm{p}} = 0.1$ and consider just one of the recoil kicks during scattering for now. When $\eta_{\mathrm{p}} = 0.1$, the trap frequency (times $\hbar$) is 100 times larger than the recoil energy. Therefore an atom can average $1E_{\mathrm{r}}$ of heating per cycle after 100 cycles by staying in the same state after the momentum kick 99 times and moving one state higher in energy once. In actuality, the atom can also increase its trap level by more than one, or even decrease its trap level as shown by the matrix elements with $\Delta n > 1$ in Fig 2-8(b) and Fig. 2-9(b). However this slight oversimplification still demonstrates the main point: in the Lamb-Dicke regime the atom receives the same recoil heating on average as outside the Lamb-Dicke regime, but it still remains in its same initial trap state

more frequently in the Lamb-Dicke regime than when it is outside of the Lamb-Dicke regime. This is just a reflection of the fact that when the energy of the recoil is small compared to other energy scales in the problem, it produces comparatively smaller effects. Despite producing smaller effects, the actual recoil energy is the same. The difference in the Lamb-Dicke regime isn't that the recoil is smaller, it's that the other energy scales are larger.

At this point it may seem that the Lamb-Dicke regime doesn't affect the amount of recoil heating during cooling. However, this isn't quite true either. This is because although the average heating per scattering is the same, the Lamb-Dicke regime can reduce the number of scatterings required to get an atom into the ground state. Suppose that during Raman cooling an atom is moved from $|g_1\rangle$ to $|g_2\rangle$ by the Raman transition and in the process is placed into the ground state of the trap. The atom then needs to be optically pumped back to $|g_1\rangle$ to complete the cooling cycle. If the atom is outside the Lamb-Dicke regime, then it is exceedingly likely that the optical pumping recoils will move the atom to a more highly excited trap state. In order to cool to the ground state, cooling cycles would have to be repeated until the atom just happened to end up in the ground state after the optical pumping. This leads to many iterations of cooling followed by recoil heating. On the other hand, if the atom is in the Lamb-Dicke regime then it is likely that it will remain in the ground state after optical pumping. At that point the atom is in its lowest energy state possible, so the cooling is complete and no more Raman cooling cycles are necessary. This reduces the total amount of recoil heating and makes cooling to the trap ground state much more feasible in the Lamb-Dicke regime.

Chapter 3

Raman Cooling Apparatus

The work presented in this thesis was performed in a lab that was initially built about twenty years beforehand. The lab has hosted several generations of grad students, not to mention post docs and other contributors, over the years and thus is an accumulation of the work of many individuals. For this reason, much of the apparatus has been documented in other theses [74, 75, 76, 77, 78, 79]. In the interest of completeness, some of the more central components documented in those theses will briefly be summarized here as well. This is intended to provide sufficient context for the reader to understand the experimental work presented here. Components which are newly built or upgraded will be described in more detail. For lab members, many more details about the system which are not relevant for this thesis are instead recorded on the lab wiki, which is the subject of Appendix B.

3.1 Experiment Chamber

The entirety of the lab is designed to control the electromagnetic field in a small region inside of a vacuum chamber. This section focuses on that vacuum chamber and the components surrounding it. Section 3.1.1 discusses the chamber itself. The coils surrounding the chamber used to control the magnetic field are discussed in Section 3.1.2, and the beam paths for light directed to and from the atoms are presented in Section 3.1.3.

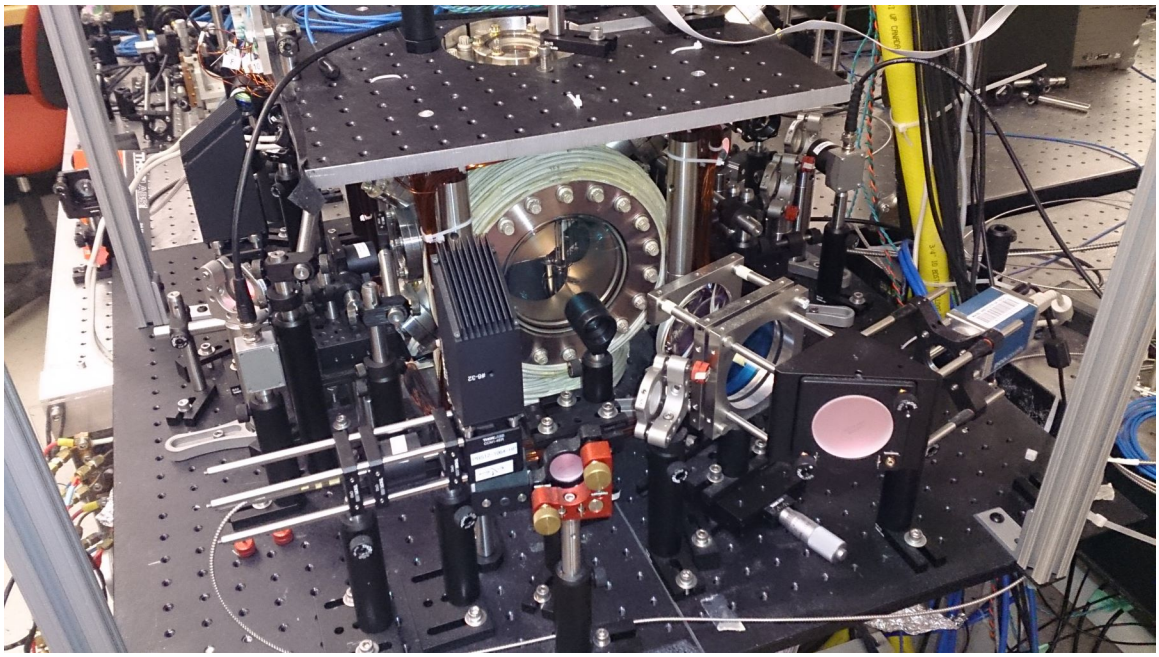
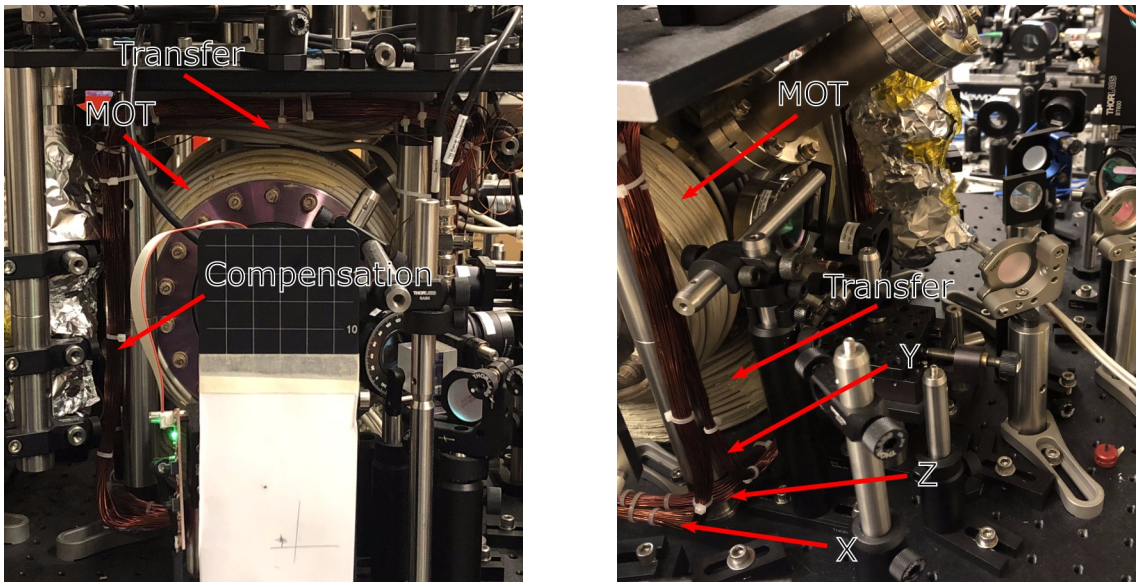


Figure 3-1: Photograph of the vacuum chamber. The ultracold clouds are prepared inside of the vacuum chamber in order to protect them from collisions with atmospheric molecules. Also visible are some of the magnet coils as well as many of the optics and optomechanics used to guide light to the atoms. This photo was taken before the addition of the Z Downward trap beam and before the trap beams were intensity stabilized, which will be discussed in Section 3.1.3.

3.1.1 Vacuum Chamber

The heart of the experiment is an ultra-high vacuum (UHV) chamber, inside of which we prepare our ultracold atoms. A photograph of the chamber, which is still the original chamber from the lab's initial construction, is presented in Fig. 3-1. It has ten ports, including two large windows on 6.75" CF flanges along a horizontal axis which we refer to as the x-axis. Three pairs of smaller windows on 2.75" CF flanges are arranged in the y-z plane, including two pairs of diagonal windows used for the y-z magneto-optical trap (MOT) beams. The axes of the diagonal flanges are 35° from horizontal. The top and bottom flanges are both 4.625" CF. The top flange originally supported the atom chip apparatus used by previous generations [74, 75, 76, 77, 78, 79], but we replaced it with with a zero-length 4.625"-2.75" reducer and a 2.75" viewport. This was opportunistically done to improve optical access when we needed to bring the chamber to atmosphere in order to set up the rubidium vapor cell anyway,



(a) (b)

Figure 3-2: Photos of magnet coils surrounding the chamber. Each edge of each compensation coil lies against the edge of another compensation coil, which makes them hard to distinguish in the photos. In (a) one edge with two compensation coils is labeled "Compensation." In (b) the compensation coils are labeled individually with arrows pointing to a corner where they can be distinguished more easily.

as will be discussed in Section. 3.5.3. The bottom 4.625" CF flange connects to another section of vacuum chamber which houses an ion pump, titanium sublimation pump, ion gauge, wiring feedthrough for the (now empty) rubidium dispensers, and a valve used to close off the chamber after pumping it down. An additional viewport is also present at the very bottom of the chamber, allowing further optical access in the vertical direction.

3.1.2 Magnet Coils

The chamber itself is surrounded by pairs of magnet coils, as shown in Fig. 3-2. One pair of coils, called the MOT coils, consists of two water-cooled coils arranged in an anti-Helmholtz configuration along the horizontal x-axis. Not surprisingly, those coils are used to generate the quadrupole field for the MOT. Another pair of water-cooled coils, called the transfer coils, are oriented along the vertical z-axis and historically were used to displace a magnetic trap when transferring atoms to the

atom chip. Those coils were originally arranged in a Helmholtz configuration, but they were rewired to an anti-Helmholtz configuration to create magnetic field gradients. Despite initially having plans to use the magnetic field gradient from the transfer coils, they were not used in the work presented in this thesis and are only mentioned for completeness. Three additional pairs of coils, called the compensation coils, are arranged orthogonally around the chamber. They are Helmholtz coils used to control the magnetic field at the atoms and are named after the axis along which they create a magnetic field, e.g. the X coil for the x-direction. Each pair is made from solid core Kapton-insulated wires and they are passively air cooled. Each compensation coil is wound in a square and they are arranged in a cube around the vacuum chamber. Each face of the cube is outlined by one coil and two coils meet along each edge of the cube. Some improvements were made to the design of the compensation coil controllers, which will be discussed in Section 3.3.

3.1.3 Beam Paths

The apparatus has many different beam paths for guiding light to and from the atoms. This subsection describes each of these beam paths, while the descriptions of the lasers used to generate this light is deferred to Section 3.2.

The first beams that the atoms interact with are the 780 nm MOT beams shown in Fig. 3-3, each of which is retroreflected onto itself. Light from one fiber containing both the MOT and repumper light is split then directed through the diagonal viewports of the vacuum chamber. Light from a second fiber, containing only MOT light, is directed in the horizontal plane through the large viewports of the vacuum chamber. That MOT beam is not quite orthogonal to the chamber windows. Instead it is rotated $\approx 10^\circ$ in the horizontal plane to make space for other beams.

After atoms are captured by the MOT, they are transferred to a far-off-resonant trap generated with 1064 nm light. Four trap beams at this wavelength pass through the chamber, one of which is retroreflected. These beam paths are shown in Fig 3-4. Each beam can be turned on or off independently of the others and has its own analog power control with intensity stabilization. As described below, this makes for a very

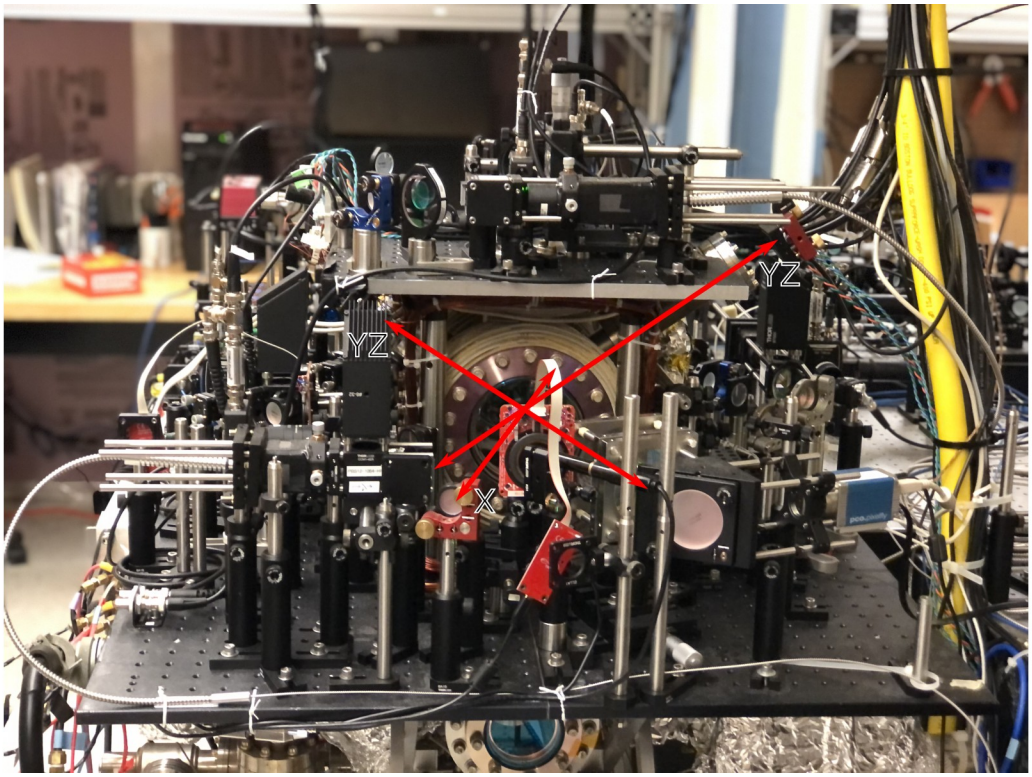


Figure 3-3: The MOT beam paths. Two paths propagate diagonally in the plane of this image and are labeled "YZ." One path propagates nearly parallel to the x-axis, approximately perpendicular to the plane of this image, and is labeled "X." Each beam is retroreflected, and two of those retroreflection mirrors are visible in the photo (Red Radiant Dyes mirror mounts). The other retroreflection mirror and the forward-direction mirrors are not visible in the photo.

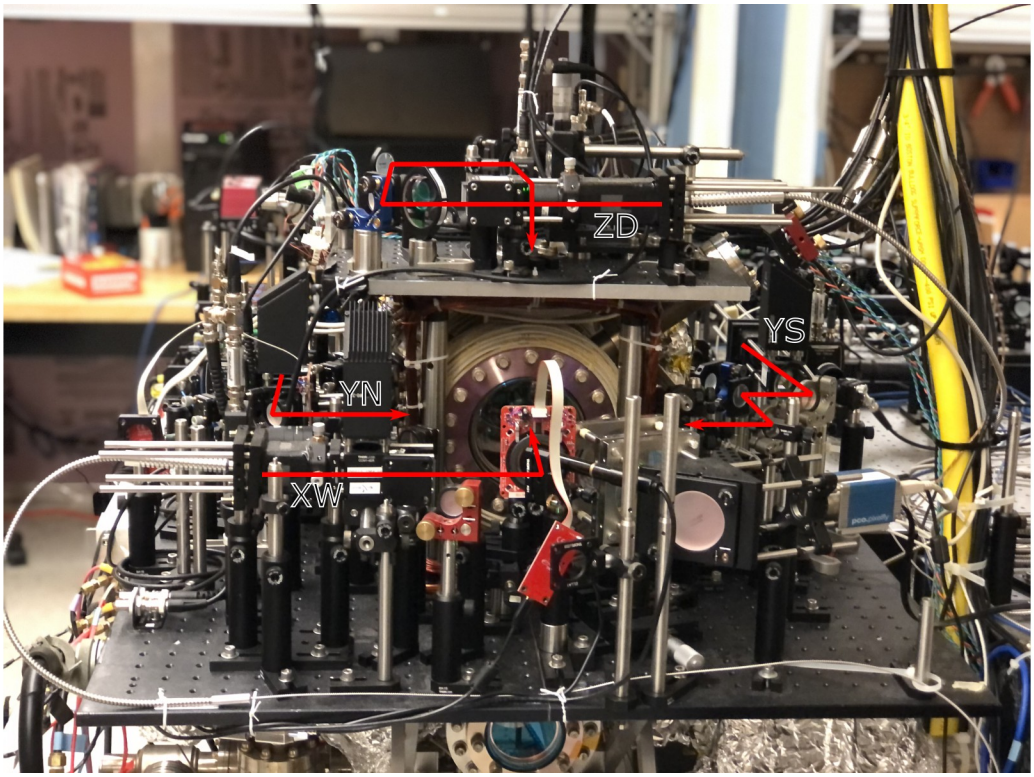


Figure 3-4: The trap beam paths, labeled with the path name shorthand notation. XW is X Westward, YS is Y Southward, YN is Y Northward, and ZD is Z Downward. The X Westward retroreflection optics are on the far side of the chamber and are not visible in this photo, but can be seen in Fig. 3-5.

flexible setup which can generate a single optical dipole trap (ODT), a crossed optical dipole trap (cODT), a 1D lattice, a 2D lattice, or a 3D lattice. The beams can also be turned on and off mid-sequence to transfer atoms between these different types of traps, e.g. when performing a cloud compression procedure by retrapping [26, 45, 24].

The trap beams were added at various times and occasionally their lenses were changed to adjust the trap beam waist. Not all of these beams are used in all of the work presented in this thesis; most of the experiments only made use of a subset of these beams. In the lab we generally label these beams by the axis along which they propagate and their propagation direction. For example, the beam propagating along the y-axis in the southward direction is referred to as "Y Southward," which is often shortened to "Y South" or "YS." The other beams are "X Westward," "Y Northward," and "Z Downward." All of these beams are derived from the same laser, but they are passed through AOMs which shift their frequency by ± 80 MHz. The Y Northward and Y Southward beams use the $+1$ AOM diffraction order, while the X Westward and Z Downward beams use the -1 order. This configuration makes it so that the Y beams interfere with each other to form a 1D lattice when they are both on, but they do not interfere with the X Westward and Z Downward beams. Similarly the X Westward beam and Z Downward beam interfere with each other but not the Y beams.

The X Westward beam can also be retroreflected upon itself, and the retroreflection can interfere with the beam's forward direction and Z Downward as well. The retroreflection can be blocked either manually or electronically with a high-power shutter (Thorlabs SH05). This makes it possible to create a horizontal 1D lattice (X Westward on, retroreflection on, Z Downward off), a diagonal 1D lattice (X Westward on, retroreflection blocked, Z Downward on), or a 2D lattice (X Westward on, retroreflection on, Z Downward on). The lattice dimension can also be increased by one by turning on the Y beams as well. For example, turning on all of the beams generates a 3D lattice. Single ODTs can be generated by turning on any individual beam alone. Furthermore, a cODT can be generated by turning on one of the Y beams and either the X Westward or Z Downward beam. The work in Ch. 4 was performed using the Y

Southward beam and X Westward beam (with its retroreflection blocked). The work in Ch. 5 was performed using the Y Southward beam and the Z Downward beam, in part because the fiber for the X Westward beam had to be returned for repair. The Y Southward beam was generally preferred over the Y Northward beam for ODTs because it has a smaller beam waist. That gave it better axial confinement due to the smaller associated Rayleigh range.

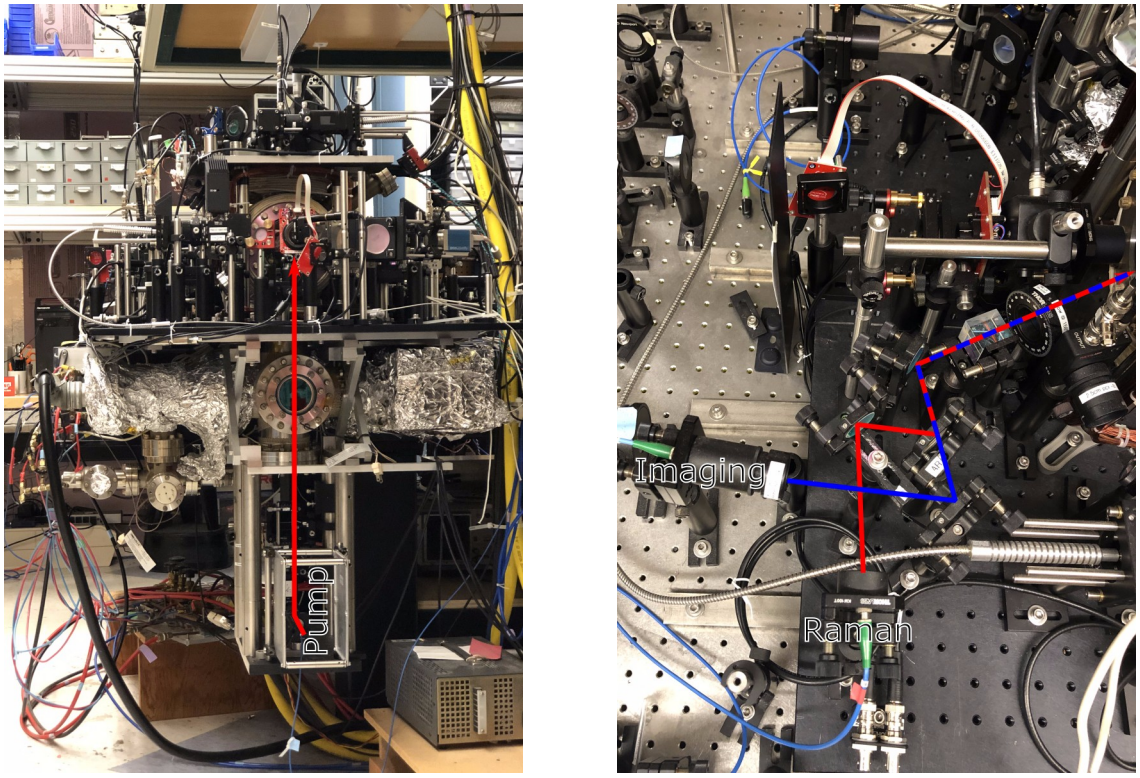
Each of the trap beams is angled slightly relative to the vacuum viewport that it passes through. This was done because we had seen effects of residual lattices. In particular we observed that atoms could end up pinned in place in a optical dipole trap as if there were a lattice, even when the other beams were turned off. This could be observed by cooling atoms in the Y lattice then displacing them from the waist of the Y Southward ODT. That was done with an "atomic conveyor belt" by slightly tuning the drive frequency of one of the Y beam AOMs [80, 81]. After extinguishing the Y Northward beam, most atoms would be released from the lattice and would begin to oscillate around the center of the Y Southward ODT. However, some atoms would remain offset from the ODT waist, pinned in place as if the lattice were still on. This could occur even if Y Northward were blocked with a shutter, so the residual lattice was not due to a finite extinction ratio of the Y Northward light. We suspected that this may have been due to partial reflections off of the AR-coated viewports (as the Y Southward beam leaves the chamber) interfering with the forward-propagating beam, thus creating a lattice. It was also possible that the reflections occurred off of other surfaces, such as the tip of the Y Northward fiber itself, as at the time the two beams were exactly counterpropagating. The effects of the residual lattice were alleviated after angling the beams, presumably because this misaligned stray reflections from the forward propagating beam. To avoid the issue in the future, all of the trap beams were angled with respect to their viewports. The Y beams were both angled upwards slightly to form a slight "X" shape. With that configuration, the light from one would not hit the optics of the other, which further reduced the number of possible accidental retroreflection surfaces.

Some of the trap beam paths have waveplates to provide polarization control.

This enables tuning the interference angle between trap beams, which tunes both the lattice frequency and the Raman coupling provided by the trap beams themselves. These waveplates were originally installed in manually-adjustable mounts. Later the manual mounts were replaced with motorized mounts (Thorlabs ELL14), which were then integrated with the control system. In particular, the Y Northward beam and the forward direction of the X Westward beam have half-wave plates to allow rotating their linear polarization. The X Westward retroreflection beam path has a quarter-wave plate, which acts as a half-wave plate because the retroreflected light passes through it twice. Thus the Y Northward beam, the X Westward beam, and the X Westward retroreflection all have independently adjustable linear polarizations. The Y Southward beam has a fixed linear polarization along the z-axis, and the Z Downward beam has a fixed linear polarization along the y-axis so that it can interfere with X Westward and its retroreflection.

Our Raman cooling requires 795 nm light, so two 795 nm beam paths are included in the apparatus as shown in Fig 3-5. The first is the "optical pumping" beam which propagates approximately vertically upwards from the bottom of the chamber. When using the trap light for Raman coupling, as we typically do during Raman sideband cooling, that light alone is sufficient. When using the 795 nm light to perform the Raman coupling, as we typically do when performing Raman cooling in an ODT or cODT, then a second 795 nm beam is needed. The second beam is referred to as the "Raman beam" and it drives the up-leg of the Raman transition. Although the optical pumping beam contributes to the Raman coupling by driving the down-leg, we generally still refer to it as the optical pumping beam. The Raman beam propagates in the horizontal plane at about a 17° angle with the x-axis. This was chosen because it then overlaps with the imaging path, which will be discussed below. The Raman beam was aligned with the imaging path primarily because that made it easy to align it to the atoms.

The last beam that passes through the chamber is the absorption imaging beam. Its path before passing through the chamber can be seen in Fig. 3-5(b) and its path after the chamber can be seen in Fig 3-6. As mentioned above, it passes through



(a) (b)

Figure 3-5: The 795 nm light beam paths. (a) The optical pumping beam propagates upwards from the bottom of the chamber. (b) The Raman beam propagates in the horizontal plane and is overlapped with the absorption imaging beam. The X Westward retroreflection optics are visible in (b), just above the region with the overlapped path.

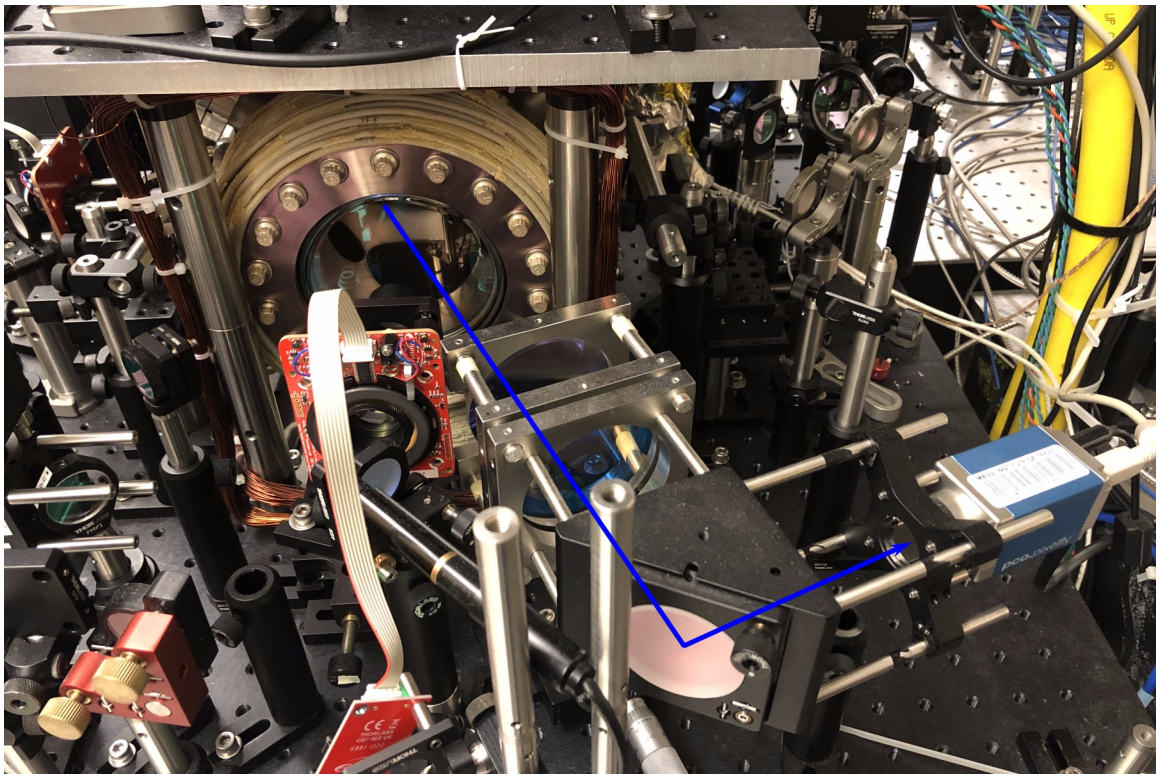


Figure 3-6: Photo of the camera side of the absorption imaging beam path. The absorption imaging beam passes through the chamber from the other side, following the path shown in Fig. 3-5(b).

the chamber in the $x-y$ plane, approximately 17° off of the x-axis. After passing through the chamber it passes through a few lenses and onto a camera (PCO PixelFly USB PF-M-QE-PIV) [79]. To avoid unnecessary background in the image, no light is included in the absorption beam to repump the atoms. Instead the repumper light in the MOT beam paths is used to repump the atoms during imaging. The repumper light can be sent to the atoms without sending the MOT light as well thanks to shutters placed before the optical fiber.

3.2 Lasers Systems

A critical component of nearly all atomic physics experiments is the laser system, which is detailed in this section. The experiments discussed here use systems centered around three different wavelengths: 780 nm, 795 nm and 1064 nm. A subsection is included for each wavelength below with details about their designs.

3.2.1 780 nm System

Light at 780 nm is used for generating the ^{87}Rb MOT and for imaging the cloud. The design of 780 nm light system is primarily the same as in previous generations of the experiment, though some of those lasers have been repaired or replaced. The overall design consists of five lasers, which are depicted in Fig. 3-7. The first is the reference laser, which is frequency-stabilized (aka "locked") to a vapor cell and acts as a frequency reference for the other lasers. The light from a second laser, the MOT master laser, is used to perform absorption imaging of the atoms as well as injection lock two other lasers in a master-oscillator-power-amplifier (MOPA) arrangement. The light from those injection-locked lasers is then used to generate the MOT. Details about each laser and the system overall are provided below.

The reference laser is locked to an ^{85}Rb transition via Doppler-free polarization spectroscopy [83] on a room temperature vapor cell [75, 76]. The light from the reference laser is not sent to the atoms, rather it is used as a reference for other lasers which are locked to it. Originally a homemade grating laser was used for this purpose

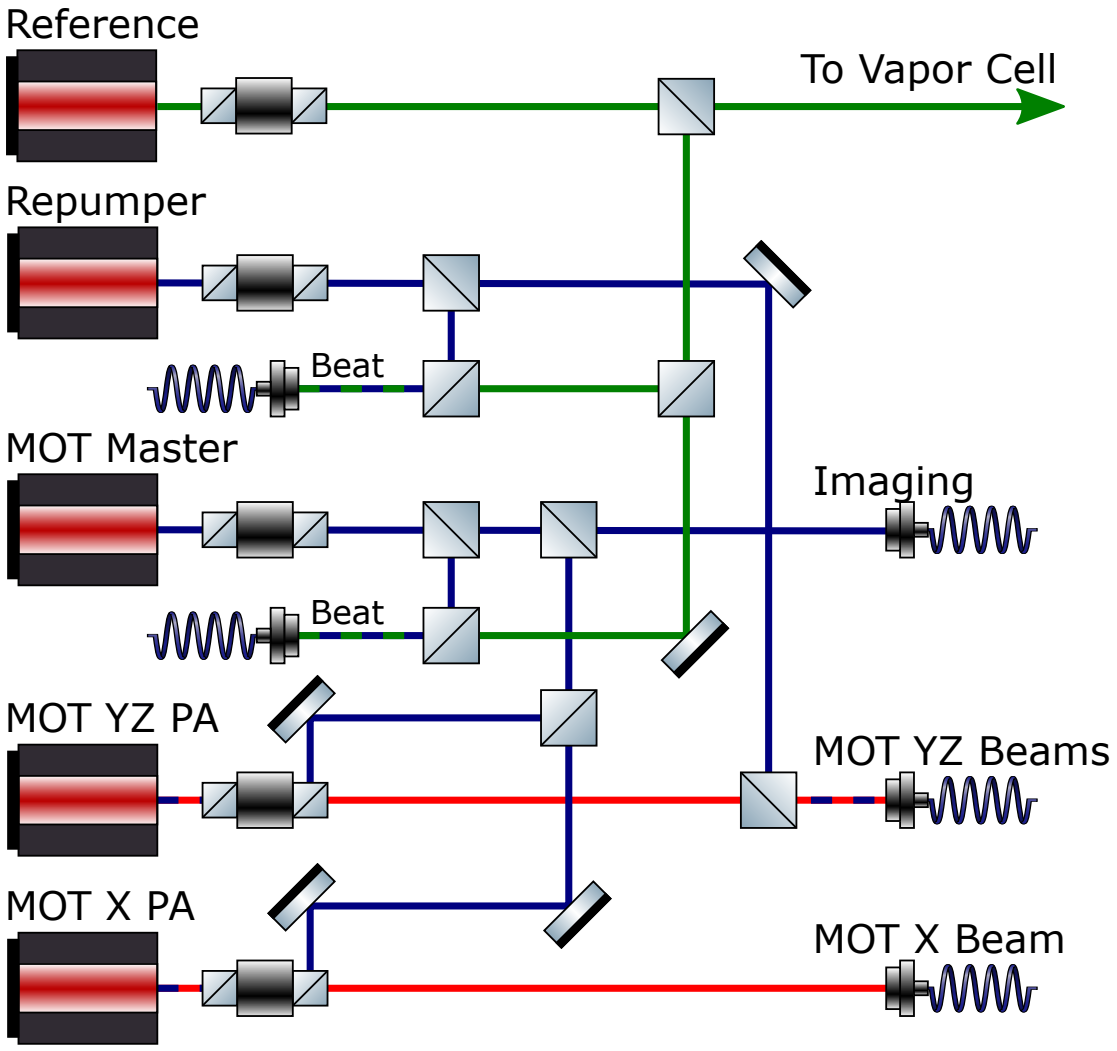


Figure 3-7: 780 nm laser system, drawn with ComponentLibrary [82]. The reference laser is locked to a vapor cell and the MOT master and repumper lasers are beat note locked to the reference laser. The MOT Power Amplifier (PA) lasers are injection to the MOT Master. The beams entering optical fibers on the right go the the experiment vacuum chamber. The beams entering fibers on the left go to the beat note lock photodiode and electronics. Shutters, acousto-optic modulator (AOMs), waveplates, and some other optics are omitted for clarity.

following the design in Ref. [84] and discussed further in Refs. [74, 75]. Unfortunately that laser would mode hop a few times per day, causing it to go out of lock each time [1]. A custom-made "mode stabilizer" circuit was created to reduce the frequency of mode hops, which helped the laser stay in lock. Additional details on the mode stabilizer are included in Appendix B. Unfortunately even with the mode stabilizer the laser would still go out of lock roughly once every few days. This was likely due to its output railing in which case it could not drive the homemade reference laser's piezo to a large enough voltage in order to keep it in the same mode. The homemade reference laser was used for the duration of the work presented in Ch. 4 and for the initial part of the work presented in Ch. 5.

Restricted lab access during the COVID-19 pandemic provided enough motivation for us to replace the aging homemade laser with a commercial cateye laser from Moglabs, which was used for the rest of the work presented in Ch. 5. The hope was that the commercial laser would be more stable and would stay in lock longer, making it easier to run the machine remotely. This turned out to be true to some extent; often the laser would stay in lock for two weeks or more. However it was found that the alignment of the lasing cavity was prone to drift, which would drive mode hops that would become more frequent as the alignment became worse. A replacement laser from the company showed the same behavior, requiring internal realignment roughly every six weeks. As of this writing we are in contact with the company in an effort to resolve this issue and may or may not continue to use this laser in the future. Ideally I would have incorporated the laser into Lock Monitor, a plugin I wrote for blacs, which is a component the Labscript suite [85, 86, 87]. Lock Monitor will be discussed further in Appendix C, but in short it is designed to monitor laser locks and automatically relock lasers when necessary. Unfortunately the Moglabs laser controller does not have external control inputs that can adjust its mode position. That means that re-locking the laser after its cateye alignment has drifted typically cannot be done via the electronic controls. Instead, some potentiometers on the controller must be

[1] That is on a typical day. It could go out of lock more frequently on days where the weather changed significantly, particularly when the outdoor temperature crossed past room temperature.

manually adjusted in order to tune the laser current separately from its piezo voltage and bring the laser back to the desired mode. That makes it impossible to relock the laser remotely.

The MOT master and repumper lasers are both distributed feedback (DFB) lasers with similar construction and lock setups, as detailed in Ref. [76]. Both are locked to the reference laser via tunable beat note locks. The beat notes [2] are prescaled down to ∼300 kHz then fed into a LM565 phase-locked loop (PLL) integrated circuit (IC). The PLL adjusts the control of a voltage-controlled oscillator (VCO) to match its frequency to that of the ∼300 kHz prescaled beat note. The VCO control signal is then a measure of the input signal's frequency. That measure can be used to feedback and adjust the laser's frequency to the desired value relative to the reference laser. The frequency setpoint can be adjusted dynamically by adjusting an offset added to the VCO control signal, allowing the frequencies of the MOT master and repumper lasers to be adjusted over the course of an experimental sequence. The light from the repumper is used directly in the experiment. It is applied to the atoms during the MOT and typically during imaging as well.

A portion of the light from the MOT master is injected into two power amplifier lasers [3]. These are made with cheaper laser diodes which can operate at higher power than the laser diode in the MOT master, but they have large free-running linewidths. However, when they are correctly adjusted, the light injected from the MOT master causes them to lase at exactly the MOT master's wavelength. This generates light with the linewidth of the expensive low-power laser diode but with the power of the broad high-power diodes. This gives the emitted light the best of both worlds, narrow linewidth and high power. Unfortunately, injection locked systems are notoriously prone to environmental drifts which can cause the power amplifier lasers to begin lasing at their own wavelength rather than at that of the MOT master. Additionally,

[2] The repumper's beat note is actually first mixed with a 6.8 GHz signal before going to the prescaler. This reduces the frequency to a range that the prescaler can handle.

[3] These are often called "slave" lasers in the current nomenclature, though that term may be on its way out of common usage due to the sordid history of slavery and racism in the United States. For this reason I'll use the naming convention "master laser" and "power amplifier" which matches the master-oscillator-power-amplifier (MOPA) phrase commonly used to describe similar systems.

these lasers have historically died more often than the others, likely due to us running them closer to their nominal maximum current as that generally decreases the lifespan of laser diodes. For this reason, some upgrades were made to these lasers and how we use them.

The first change was the development of a daily alignment procedure. Previously the injection paths were rarely realigned because the alignment procedure based on injection pulling was difficult, tedious, and time-consuming. The new, much easier, procedure is as follows. First, the fiber coupling of power amplifier lasers to single mode fibers are reoptimized. This is a standard procedure in AMO labs which can be done in a matter of minutes or less. Due to their single mode nature, each fiber then acts as a very precise marker for the beam position and direction. After optimizing the fiber coupling, the power amplifier lasers are turned off but the MOT master is left on. Some of the MOT master light on the injection path then reflects off of the power amplifier's diode and then proceeds along its beam path to the fiber. The alignment of the injection beam is adjusted to optimize its fiber coupling, thereby aligning it to the power amplifier's beam path and optimizing the injection alignment. Importantly, only mirrors that are on the injection path but not on the power amplifier's beam path are adjusted in this step, otherwise the fiber would no longer mark the power amplifier's beam path. After this alignment, the power amplifier lasers are turned back on and allowed to warm up. This alignment procedure can be done in $\approx$5 minutes and then the lasers would warm up after $\approx$20 minutes, which makes it fast enough to do on a daily basis. That means that it is possible to keep the injection locking path well-aligned, which makes the injection locks much more stable.

The next upgrades to the power amplifier system were mechanical. Originally the collimating lenses had to be carefully positioned using the 3D translation stage setup, then glued into place. This was also a difficult, tedious, and time-consuming procedure. It also had to be redone from scratch whenever the laser diode was replaced and the results were very fragile [4]. The lasers now have their collimators mounted in

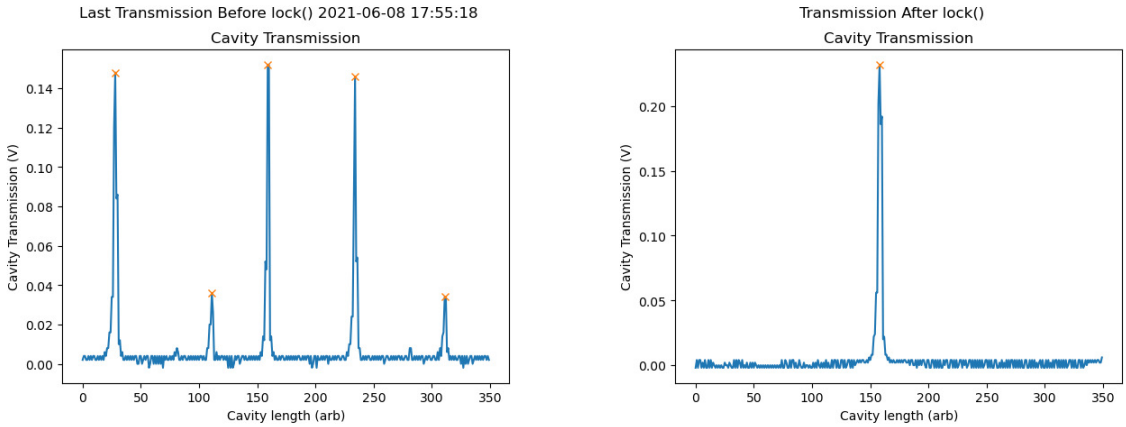
[4] I, and others in our group, have accidentally knocked a collimator off of a laser while making other repairs. Putting the collimator back in place required removing the laser from the system, cleaning the old glue off of the components, then realigning and gluing the collimator. The laser

standard Thorlabs SM05 lens tubes. This makes it much easier and faster to align the collimators, and they are much less prone to falling off. The quick alignment is particularly helpful when replacing the laser diodes.

Although it is much faster now to replace the laser diodes in the power amplifier lasers, doing so still takes some time. In an effort to reduce how frequently this needs to be done, we changed the type of laser diode in use. Previously our lab used Sharp GH0781JA2C laser diodes which were designed with CD drives in mind. That meant they were reasonably cheap and were rated for 120 mw of continuous optical power. However, they would die every few years in our lab because we would run them for several hours a day at several tens of milliwatts of power, which is over half their rated maximum power output. Although this was within their safe operating limits, running laser diodes at a large fraction of their maximum rated power generally reduces their lifespan. The power amplifier lasers now use Thorlabs LD785-SH300 laser diodes which are rated for 300 mW continuous optical power. We still run them at about the same power that as we did when we used the Sharp GH0781JA2C diodes, but now this is a much smaller fraction of the diode's maximum power rating. We hope that this will significantly increase their lifespan and reduce how frequently the power amplifier laser diodes need to be replaced.

The final upgrade to the power amplifier lasers was a control system upgrade: they were integrated into Lock Monitor. For these lasers, a laser locking class was written to communicate with an oscilloscope that measures the transmission through a scanning confocal Fabry-Perot cavity. That cavity is used to observe the power amplifier lasers' frequencies while locking them. Importantly the light from both power amplifier lasers goes to the same cavity. The light from the lasers couples to different, but degenerate, cavity modes so that they do not interfere with each other but they still see the same cavity mode resonant frequencies. When both power amplifier lasers are locked, they operate at the same frequency and transmit through

then had to be installed back into the system and all of its beam paths then had to be realigned. This would set us back a few days of work and generally prohibited us from taking any data until after finishing that repair, not to mention the other repairs that were already underway when the collimator was knocked off.



(a) (b)

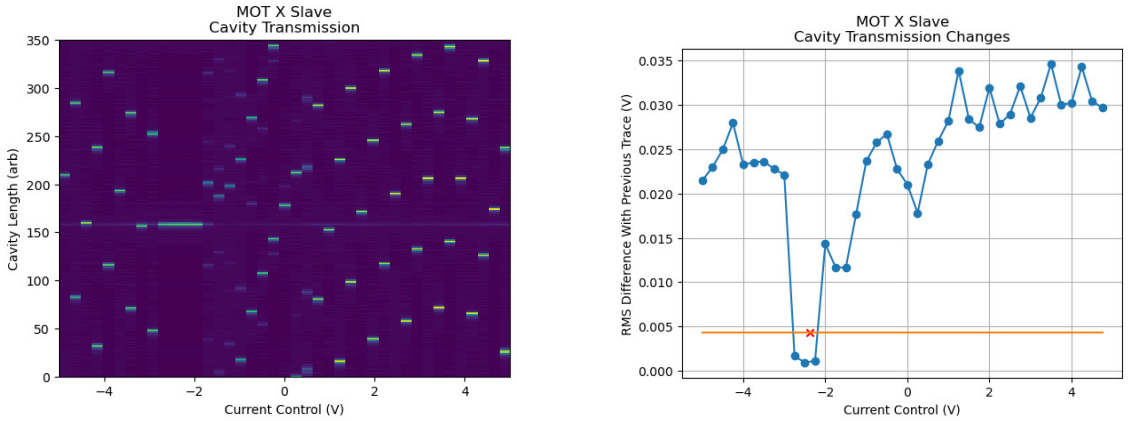
Figure 3-8: Scanning Fabry-Perot cavity transmission signals used for locking the MOT power amplifier lasers. (a) Transmission signal when a laser is out of lock and multimode. (b) Transmission signal after relocking the laser. The absolute position of the peaks drifts due to temperature fluctuations in the lab which affect the cavity length. The scan range is slightly more than one FSR, so for some cavity lengths only one peak appears near the center of the scan when the lasers are locked. Sometimes two peaks will appear when the lasers are locked; one near each end of the scan separated by one FSR. The traces here are from an actual automatic relocking event. The trace in (a) was acquired and the laser locking class determined that at least one laser was out of lock, so the experiment was paused. The locking procedure was then automatically performed and then trace (b) was taken, from which it was determined that the relocking was successful. The experiment was then allowed to continue.

the scanning cavity at the same time as the cavity's length changes. In that case, the transmission peaks seen in the oscilloscope signal always occur one free spectral range (FSR) apart [5]. However, if one of the power amplifier lasers goes out of lock or begins to go multimode, then additional peaks will generally appear [6] and the minimum space between two peaks will be smaller. Two cavity transmission traces are shown in Fig. 3-8. Before each shot of the experiment, the laser locker class acquires an oscilloscope trace and finds all of the peaks in it. If they are all separated by about one FSR, then it allows the shot to start. If it detects that there are peaks that are too close together then it automatically relocks the lasers.

The locking procedure for the power amplifier lasers is straightforward and mimics what we do when locking the lasers manually. When manually locking one of the lasers, we scan the diode current and watch the cavity transmission signal. When the laser is out of lock, its output frequency changes as the current is tuned, so the peaks in the cavity transmission signal move as the current is adjusted. Once the laser enters lock, its frequency is independent of the diode current, so the peaks stop moving. This remains the case until the current is adjusted too far, at which point the laser jumps out of lock and the peaks start moving again. The locking class implements this procedure automatically. It adjusts the diode current for one of the lasers, then acquires an oscilloscope trace, and repeats those two steps many times. A series of such traces is presented as a color plot in Fig. 3-9(a). The locking class takes each oscilloscope trace, subtracts it from the previous trace, then takes the RMS value of the difference. When the laser is locked, the peaks do not move and two consecutive traces only differ due to noise so the RMS difference is small. When the laser is out of lock, the peaks move between consecutive traces so the RMS difference is much larger. A plot of the RMS difference as a function of the laser diode current control signal

[5] Actually because we use a confocal cavity, the peaks are one half of an FSR apart. I'll neglect that factor of two in this discussion though for simplicity; it doesn't affect any of the principles of operation.

[6] It is possible for other frequency components in the light to differ from the injected light wavelength by a multiple of the cavity's FSR. In that case their peaks would line up with the carrier and it would look as if the laser were in lock. Fortunately this is a rare coincidence. Furthermore, the laser frequencies tend to quickly drift right after they go out of lock, so this coincidence likely wouldn't remain the case for long



(a) (b)

Figure 3-9: Data taken during the automatic relocking event discussed in Fig. 3-8. (a) The cavity transmission signal as a function of cavity length and laser diode current (in units of control voltage). Each vertical slice of this plot corresponds to one cavity transmission oscilloscope trace like those in Fig. 3-8, each one taken with a different laser diode current. Nearest neighbor interpolation is used, which is why the peak positions appear to jump around rather than move continuously. Note the region near -2.5 V on the x-axis where the peak remains in the same location as the current is changed. That behavior indicates that the laser is in lock at those currents. (b) RMS of the difference between each trace and the previous one. The dip below the horizontal orange line corresponds to currents where it is determined that the laser is locked. The red "x" marks the current to which the laser will be set.

is presented in Fig. 3-9(b). The locking class then identifies the region of currents where the RMS differences are low and assumes that corresponds to the region of currents which lead to successful injection locking. The locking current is taken to be the center of the locking region [7]. The locking code then ramps the laser's current to the target value, ensuring that it ramps the current in the same direction used previously so as to account for the hysteresis associated with injection locking. This procedure is performed once for each of the lasers, then a final cavity transmission trace is taken to ensure the the lock attempt was successful. The SNR of this process can be improved by blocking the light of one laser with a shutter while locking the other, but this is not necessary as long as one laser's current remains fixed while the other's is scanned. The full locking procedure to lock both lasers takes about one minute, primarily limited by how fast the oscilloscope can acquire and transfer traces to the computer.

3.2.2 795 nm System

The experiment requires 795 nm light in order to perform optical pumping on the D_1 line during Raman cooling. It is used instead of D_2 light because the $5P_{1/2}$ state does not have an $F = 3$ manifold. That makes it so that the $|F = 2, m_F = -2\rangle$ state remains dark to σ^- optical pumping light even when that light is detuned far from the $F = 2 \rightarrow F' = 2'$ resonance. If this were attempted on the D_2 line, then the $|2, -2\rangle$ state would scatter off of $|3', -3\rangle$, ruining the dark state.

Before the work presented in this thesis, the 795 nm laser system was composed of one laser. It used a DBR laser diode (Photodigm PH795DBR120T8), which was locked to a vapor cell using optical heterodyne saturation spectroscopy [88, 79]. Both the pump and probe beams are each passed through a fiber electro-optic modulator (EOM). The probe beam's fiber EOM has a 15 MHz radio frequency (RF) signal

[7] The first current where the laser is in lock has a large RMS difference with the previous trace where the laser is out of lock, so that point is not part of the dip in RMS differences. That is accounted for when determining the center of the locking region, which is why the red "x" in Fig. 3-9(b) is slightly off-center from the dip. The current is scanned from large to small control voltages, which is why the red "x" is offset to the right from the dip's center.

applied to it used to create the phase modulation required for the heterodyne saturation spectroscopy. Additionally, an adjustable microwave frequency signal is applied to both fiber EOMs such that the microwave sidebands can be locked to the atomic transition. That makes it possible to offset the carrier from the atomic transition while locking the laser. The offset is equal to the microwave frequency (assuming that the first order sideband is locked to the transition). That gave the system the flexibility to lock the laser to a wide range of frequencies around the atomic resonance.

Over time the initial laser slowly started producing less and less power, so a second 795 nm DBR laser diode was constructed with a Photodigm PH795DBR180TS laser diode. This second laser was beat note locked to the original and this configuration was used for the work presented in Ch. 4. The beat note lock gave a very simple error signal, which made it easy to scan the second laser's frequency over a large range of values. That capability was particularly important for the scan over optical pumping detunings which will be presented in Ch. 4.

After more use, the original 795 nm laser eventually fully died and would not lase at all. At that point some light from the second laser was diverted to the vapor cell lock setup so that it could be locked directly to the cell, again with an offset set by a tunable microwave frequency. This was the configuration used for the work presented in Ch. 5.

After the second laser was locked to the vapor cell, a laser locking class was written for it to incorporate it into Lock Monitor as well. This was useful because the laser would occasionally go out of lock. When it went out of lock, the error signal offset had typically drifted so far that the dispersive locking feature did not cross zero. Presumably this is why the feedback failed; there was no longer an operating point that would zero the error signal so the PID controller railed. We expect that this was due to residual amplitude modulation (RAM) as the error signal offset could be changed by toying the with fibers to the fiber EOM. A more straightforward solution may have been to temperature stabilize the fiber EOM so as to reduce RAM drifts [8],

[8]Insulating the fiber EOM by surrounding it with some packing foam seemed to help somewhat by passively stabilizing the temperature. Additionally, maximizing the amplitude of the microwave sideband also helped by increasing the signal amplitude without increasing the drift. Notably that

but the software approach was easier to develop remotely which made it a more favorable option due to the reduced lab access during the COVID-19 pandemic. The software approach also had the advantage of handling other locking failures and would make it possible to automatically relock the laser after having turned it off and on again.

The laser locking class for the 795 nm laser also mimics the approach used to lock the laser by hand. When manually locking the laser, the PID controller feedback is disabled and a function generator is used to scan the laser diode's current and thereby scan the laser's frequency. An oscilloscope is used to record the vapor cell error signal as the laser is scanned. The desired spectroscopic feature is identified in the oscilloscope trace and the offset of the scan is adjusted to center the feature in the scan range. The scan range is then reduced to zoom in on the locking point, and then that procedure is repeated several times. At some point during the procedure, the setpoint input (aka the error offset) is adjusted so that the error signal around dispersive locking feature is centered about zero. Once the setpoint has been adjusted and the scan range has been reduced to the point that it is smaller than the dispersive spectroscopic locking feature, the feedback is engaged. The laser locking class code follows this same procedure. There are several spectroscopic features within the scan range of the laser, so the locking code must be able to distinguish the desired one. The spectroscopic feature of interest is the largest in the scan range [9], so it is found by taking the position halfway between the highest and lowest points in the oscilloscope trace. The required setpoint adjustment is determined in the same way; the setpoint is changed to ensure that the highest and lowest points in the error signal are centered around zero. The laser center frequency tends to drift as the scan amplitude is changed, so the scan range is reduced in several steps rather than one big one. That ensures that the desired spectroscopic feature remains within the scan range through

is different than increasing the 15 MHz RF sideband amplitude, as that would increase both the signal amplitude and the RAM drift by the same factor.

[9]This is no accident. The various transitions and crossover signals produced by the optical heterodyne saturation spectroscopy have different amplitudes. By choosing the correct microwave sideband frequency, the laser can be locked to the strongest feature. The microwave frequency is larger than the scan range of the laser, so the carrier error signals do not appear in the scan.

the duration of the locking procedure. Plots showing the oscilloscope traces of the ramp signal and error signal during an automatic locking run are presented in Fig. 3-10. The locking procedure takes approximately 30 seconds, primarily limited by the speed at which the oscilloscope can transfer data and respond to commands.

3.2.3 1064 nm System

1064 nm light was used to generate the trapping potentials for the work presented in this thesis. That wavelength is far to the red of the ^{87}Rb D_1 and D_2 lines, so the atoms are drawn towards regions where the 1064 nm light intensity is the highest. The 1064 nm wavelength was chosen for the trap light in large part due to the availability of high power commercial lasers and fiber amplifiers at that wavelength.

The 1064 nm laser system consists of two primary components: the seed laser and the fiber amplifier. The seed laser is an NP Photonics FLM-50-1-1064.175-1 fiber laser. It is a very low-frequency-noise laser with a specified linewidth of $\sim$2 kHz. It does have significant intensity noise at its relaxation oscillation frequency of $\approx$1.5 MHz though, but this was not prohibitive for our experiments. The detuning of the 1064 nm light is $\sim$100 THz from resonance, so the experiment is fairly immune to frequency drifts of the laser. Therefore the laser was left free-running; i.e. its frequency was not locked. The seed laser does not generate sufficient power to trap the atoms, but it generates sufficient power to seed a fiber amplifier.

The fiber amplifier used in this work was an ALS-IR-50-SF, which was lent to our lab courtesy of Martin Zwierlein's group. Although this was already mentioned in the acknowledgments, I would like to reiterate my significant gratitude towards Martin for his kindness in lending us this fiber amplifier, particularly for such a long period of time. None of the Raman cooling research presented in this thesis would have been possible without it. The fiber amplifier amplifies the light from the seed laser to an adjustable power. It was set to various values for various periods of research but was generally in the range of 10 W to 20 W.

The light from the fiber amplifier was then distributed into several beam paths using half wave plates and polarizing beamsplitters. Each beam path used by our

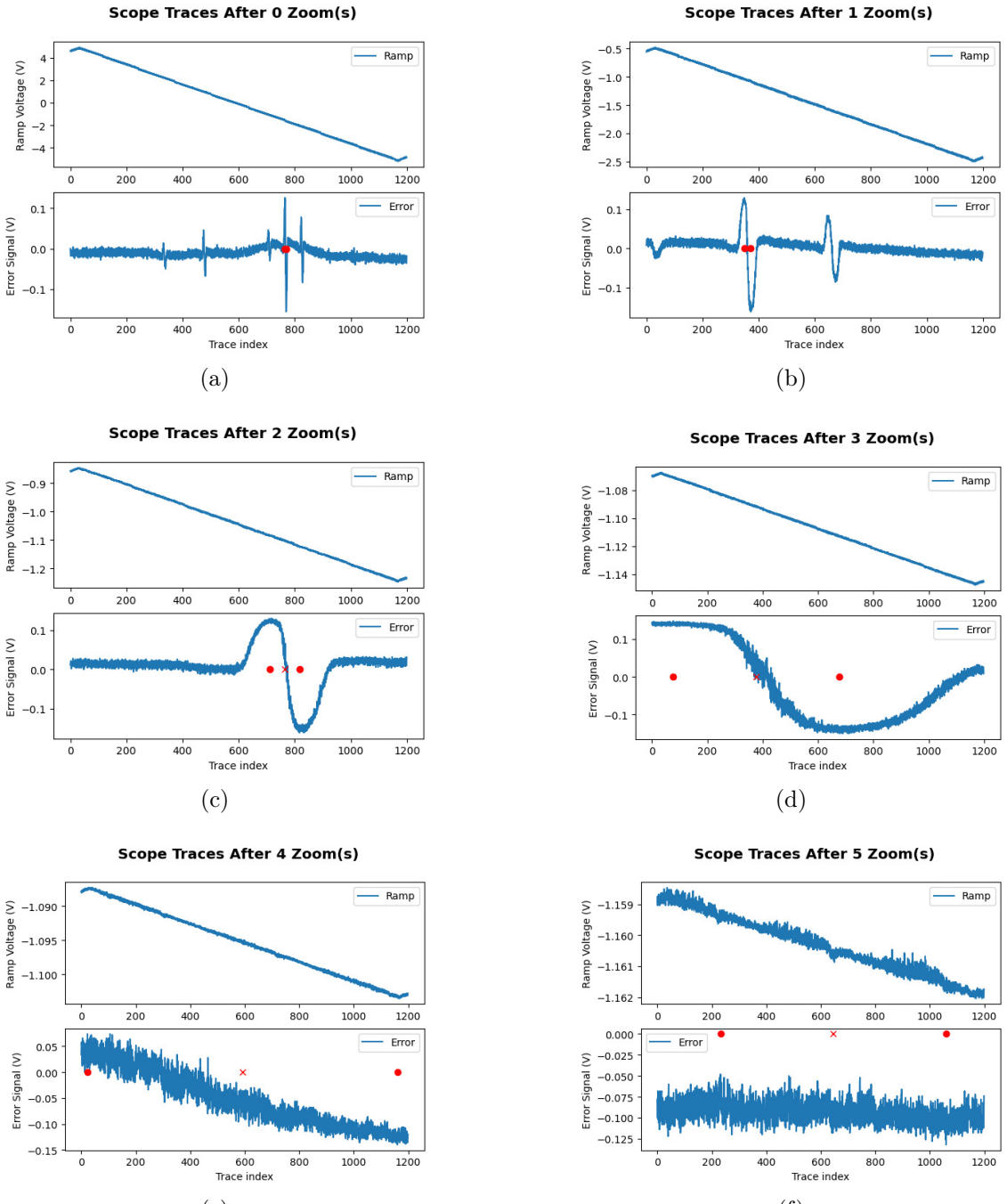


Figure 3-10: Oscilloscope traces of the ramp signal (upper) and the spectroscopic error signal (lower) taken by the automatic 795 nm laser locking code. The ramp range is reduced between each trace to hone in on the target spectroscopic feature, identified by its large peak values. The red circles on the error signal plots mark where in the scan the highest and lowest error signals are measured, marking the peaks of the target error signal. The red x on the error signal plots marks the target laser setting.

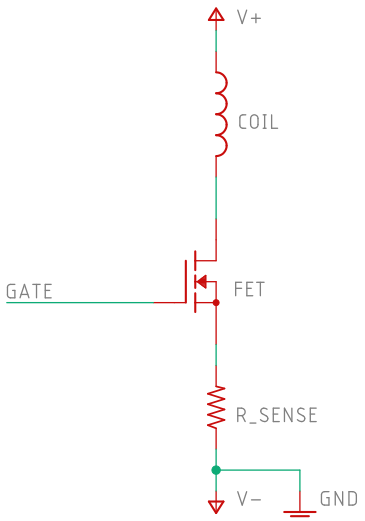


Figure 3-11: Schematic showing component connections for a typical unipolar coil driver with grounded power supply. The PID controller and connections are not shown. The compensation coils in our lab are configured for bipolar drive which uses a slightly different setup, but the unipolar setup is simpler to understand and many of the same principles apply to the bipolar driver as well.

lab [10] was passed through an AOM for individual intensity control then the light was coupled into photonic crystal fibers and transported to the chamber. On the chamber side the light was collimated then focused onto the atoms. A pickoff was installed on each trap beam path to divert a fraction of the light to a Thorlabs PDA05CF2 InGaAs photodiode. The photodiode signal was then used to stabilize the beam intensity by feedbacking back to the AOM driver. This reduced trap intensity noise and drifts, and it also had the benefit of linearizing the relation between the control signal (which adjusted the feedback loop's setpoint) and the beam intensity.

3.3 Magnet Coil Control

The compensation magnet coils around the vacuum chamber mentioned in Section 3.1.2 use a custom current control system. The basic design for these systems is straightforward and is explained in other theses from our group [75, 89, 90]. It consists of a

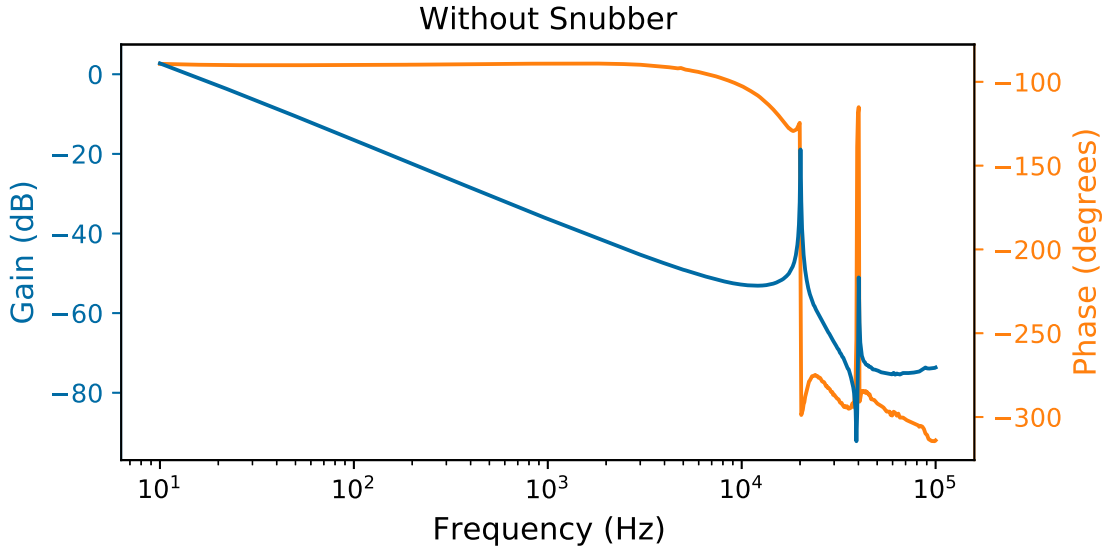
[10] Some light was also sent to the adjacent cesium lab for their experiment.

power supply driving the coil with a current-sensing shunt resistor and a MOSFET in series. A PID controller measures the voltage across the shunt resistor, which is proportional to the current, and compares it to a control voltage. The controller then adjusts the voltage applied to the MOSFET's gate to tune the current so that the shunt resistor voltage matches the control signal. A typical unipolar coil setup showing the components external to the PID controller is depicted in Fig. 3-11. The compensation coils now all use a bipolar drive setup which will be discussed later on, but the unipolar design is simpler to understand and demonstrates issues which apply to the bipolar design as well. This simple design works, but has some room for improvement, particularly with regards to feedback loop bandwidth. This section focuses on some enhancements made to the design. Improving this system is a good project for visiting students and undergrads. It is a very instructive project for them because it involves basic analog electronics and control theory, which are two sets of knowledge essential for AMO physicists. For that reason, much of the work presented here was done in tandem with various visitors to the lab, including Dawid Paszko and Haoyang Gao. Haoyang in particular had a few clever insights and contributions to the design, and as of this writing Michelle Chong is working on further improvements.

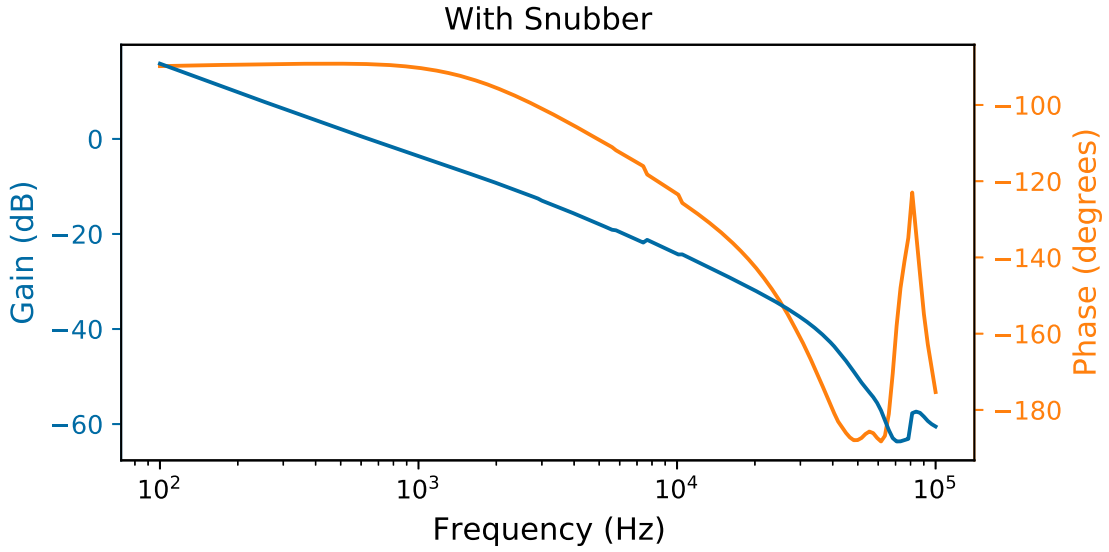
3.3.1 LC Resonance Snubber

The first issue resolved regarding the coil drivers was a resonance in its response. With the simple design in Fig. 3-11, the typical loop gain of the feedback loop looks like the results shown in Fig. 3-12(a), measured with a Stanford Research Systems SR780 network analyzer. A large resonance appears at ≈20 kHz where the gain sharply increases by about 40 dB and the phase changes dramatically. The loop gain must be reduced to keep the peak of this resonance below 0 dB in order for the loop to be stable. Reducing the gain comes at the cost of reducing the bandwidth of the feedback loop.

The resonance is caused by an interplay between the inductance of the coil and some parasitic capacitance in the MOSFET. This can be understood by including the capacitance between the MOSFET's drain and source in the schematic, as depicted

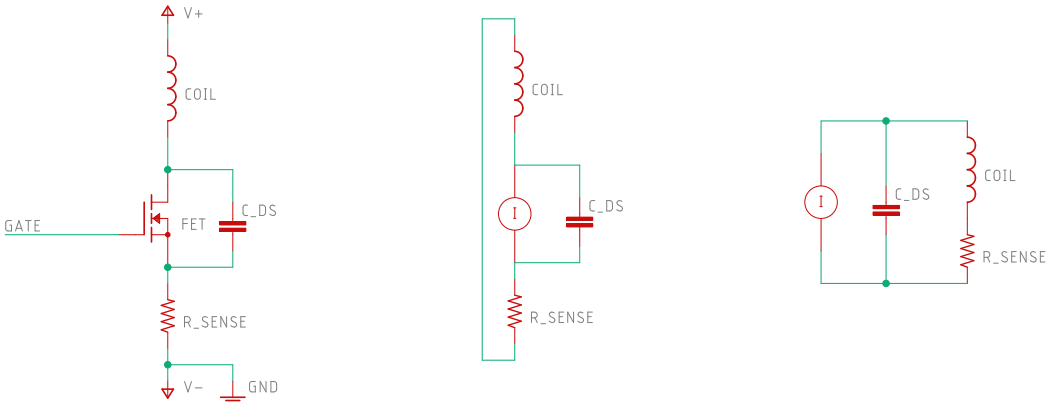


(a)



(b)

Figure 3-12: Feedback loop gain and phase for the compensation coil current control. (a) Results without a snubber installed, showing a clear resonance with large resonant gain and a dramatic change in phase over a narrow range of frequencies. (b) The same measurement after installing a snubber, which removes the resonant peak and makes the phase change more gradually with frequency.



(a) (b) (c)

Figure 3-13: Schematics showing how a parasitic capacitance in the MOSFET of a coil driver system can lead to an LC resonance with the coil. (a) Schematic of the unipolar coil setup with the parasitic drain-source capacitance included. (b) Equivalent schematic in which the power supply and MOSFET have been redrawn as a current source. (c) Identical schematic to (b) with the connections but organized in such a way to more clearly show the LC resonator. The Q of the resonator can be large when the resistance of the shunt resistor and the parasitic resistance of the inductor are low.

in Fig. 3-13. The presence of the LC resonator is made more clear when the power supply and MOSFET are considered as a current source and the schematic is slightly reorganized, as shown in Fig. 3-13(c). The Q of the resonance is spoiled slightly by the resistance of the current-sensing shunt resistor, as well as any resistance of the coil itself. However, those resistances can be small which can lead to a fairly high-Q resonance as depicted in Fig. 3-12(a). Interestingly, the drain-source capacitance of the MOSFET generally depends on the voltages applied to it, so the frequency of the resonance changes for different operating currents.

The effects of the resonance can be mitigated by "de-Q-ing" it with the addition of a "snubber." The idea is that a resistor is added in parallel with the components in Fig. 3-13(c) with its resistance chosen to roughly match the impedances of the inductor and capacitor at the LC resonance [11]. That reduces the Q of the LC resonator (i.e. "de-Q's it") to about one, which removes the sharp peak from the loop gain and

[11] The capacitor and inductor has the same impedance at the LC resonance frequency.

makes the phase change more smoothly over a wider range of frequencies. For the configuration in Fig. 3-11, this snubbing resistor is connected between the MOSFET's drain and source to help keep the current through the coil and the shunt resistor the same. The loop gain after adding the snubber is depicted in Fig. 3-12(b) where it is clear that the sharp resonance is gone and instead a simple second-order pole is present where the resonance used to be. That makes it possible to increase the loop gain without causing the system to oscillate at the LC resonance, and doing so increases the bandwidth significantly. For the system tested in Fig. 3-12, the peak is reduced 20 dB which alone makes it possible to increase the bandwidth by about two orders of magnitude. By adding a derivative filter with its corner frequency at the LC resonance, one order of the pole can be canceled to leave a first order pole. Then the gain and bandwidth can be increased further. For reference, the impedance of the coil (5.3 mH inductance) at 20 kHz is ≈700 Ω, and the snubber was made with a 250 Ω resistor. This was to ensure that the LC resonator would remain well-damped even if it moved to lower frequencies at a different current operating point due to the changing MOSFET capacitance. An exact match between the snubber and resonator impedances at the resonant frequency is not necessary; making the snubber resistance smaller simply overdamps the LC resonator.

In practice the resistance required to de-Q the LC resonator can be somewhat small. That implies that the power dissipation through that resistor can be large. To reduce the power dissipation a DC-blocking capacitor is typically installed in series with the snubber's resistor, as depicted in Fig. 3-14. The capacitance of the DC-blocking capacitor is chosen to be large enough so that the RC corner frequency between it and the snubber's resistor is small compared to the LC resonant frequency. When that is the case, the impedance of the capacitor and resistor in series is dominated by the resistor at the frequency of interest, namely at the LC resonant frequency. The snubber's DC-blocking capacitor then has little effect at the resonance. However it blocks DC currents, which typically greatly reduces the power dissipation in the snubber resistor. It is worth pointing out that AC currents still go through the snubber resistor, so it can need to dissipate a lot of power when the coil current

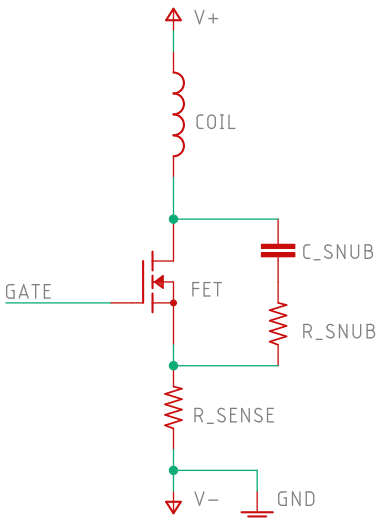


Figure 3-14: Connections for a unipolar coil driver with grounded power supply and snubber.

is changed by a large amount quickly. This is particularly common when tuning the PID parameters of the feedback loop, as the system will undergo large oscillations in current if the feedback loop becomes unstable while adjusting the system's gains. I have burnt out snubbing resistors in the past (despite using a DC-blocking capacitor) due to the system oscillating while tuning PID parameters. Therefore it is still wise to use high-power resistors even when a DC-blocking capacitor should be used. Ideally the resistor should be rated for a power greater than V^2/R where V is the power supply voltage and R is the resistor's resistance, in which case it can dissipate as much power as the power supply is capable of providing it.

3.3.2 Floating Power Supply

The schematics shown in Fig. 3-11 and Fig. 3-14 show circuits in which the MOSFET is placed between the coil and sense resistor and the power supply is grounded, i.e. its negative terminal is connected to ground. The ordering of those components can be swapped in principle, but in practice this is the only arrangement that keeps the gate voltage and the voltages at the shunt resistor low. This is helpful because the coil supply voltage is typically larger than the voltages that the PID controller can

accept as inputs or generate as outputs without the use of special (and often noisy) high-voltage op amps and other components.

Unfortunately this arrangement has a few disadvantages. As will be discussed below, it does not generalize well to a bipolar coil driver configuration. Even for the unipolar arrangement, there is a subtle issue. The voltage across the shunt resistor is intended to be proportional to the current through the coil. That is the case for ideal components, but real MOSFETs have parasitic gate-drain and gate-source capacitances. That means that some of the current going through the coil can go through the drain-gate parasitic capacitor and into the PID controller. Similarly, current from the PID controller can go through the gate-source parasitic capacitor and then through the shunt resistor. This implies that the current through the shunt resistor is not exactly equal to the current through the coil, which can cause the feedback loop to aim for the wrong operating point. Of course the impedance of these capacitors is infinite at DC so they do not affect the DC operating point. However, they can affect the dynamic behavior of the system whenever the setpoint is changed, as it is throughout a typical experimental sequence.

Haoyang Gao was the first person to point out this effect to me and showed that simulations including those capacitances produced results that matched the actual system behavior quite well. He also came up with a solution: float the power supply. Doing so makes it possible to have other orderings of the components while still keeping the voltages going to and coming from the PID controller within a reasonable range of ground. The setup for a unipolar driver with a floating power supply is depicted in Fig. 3-15. In that arrangement, the current through the coil and shunt resistor is exactly the same, even when accounting for the effects of parasitic FET capacitances. That helps ensure that the PID control is provided an accurate measure of the coil current.

The grounded unipolar circuit of Fig. 3-11 can be regarded as using a bit of a trick to keep the gate and shunt resistor voltages low: it floats the coil up to the supply voltage. Unfortunately that trick doesn't generalize to a bipolar coil design. In that case the circuit used in the past was that presented in Fig. 3-16(a), which

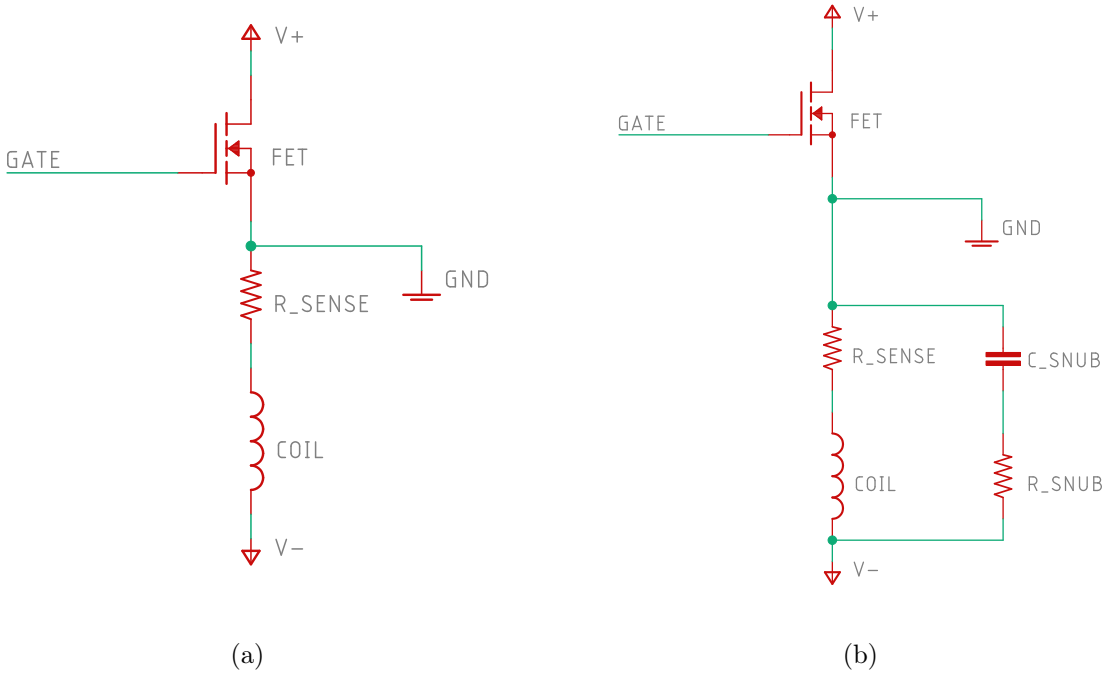
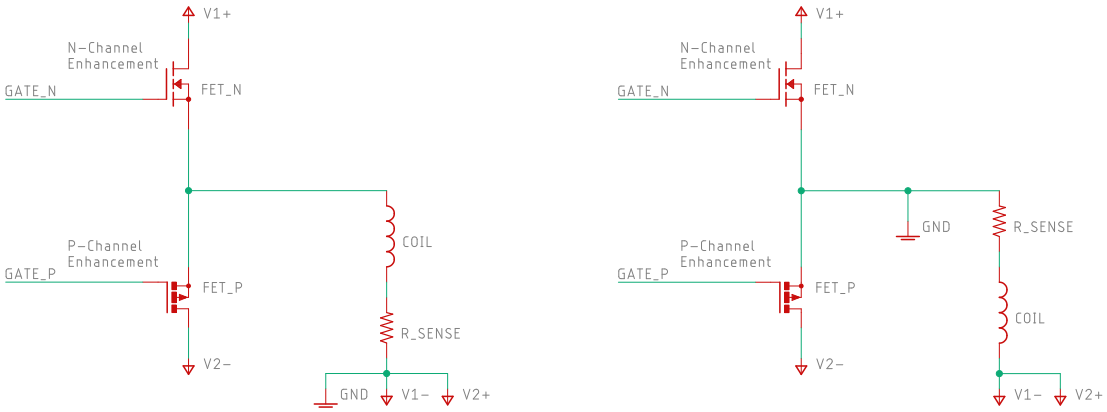


Figure 3-15: Schematics showing the connections for a floating supply unipolar coil driver (a) without a snubber and (b) with a snubber. Alternatively the snubber can be placed in parallel with the MOSFET, between its drain and source.



(a) (b)

Figure 3-16: Schematics showing the connections for a bipolar coil driver (a) with a grounded power supply and (b) with a floating power supply.

is more demanding for the PID controller to drive. If the maximum output voltage of the PID controller is less than the coil supply voltages, then the range of voltages that can be applied to the MOSFET gates are limited [12], which limits the current that can be driven through the coil. A direct workaround for this is to engineer a PID controller that can output voltages up to the coil power supply rails. This is not ideal though as it typically requires connecting extra power supply lines to the PID controller and using noisy high-voltage op amps. Fig. 3-16(b) shows a schematic which mitigates the PID controller voltage range issues by floating the power supply. In that configuration the MOSFET gate voltages only need to be driven to slightly above their threshold voltage. Thus a floating power supply design can be superior for bipolar coil drivers as well.

[12]Our recent PID controller designs for this system have two outputs which only differ by and adjustable offset voltage. One output is connected to the N-Channel MOSFET and the other is connected to the P-Channel MOSFET. The voltage offset is then tuned to minimize the crossover distortion while avoiding excessive current flow at low-current setpoints where both MOSFETS could be turned on if the offset were too large.

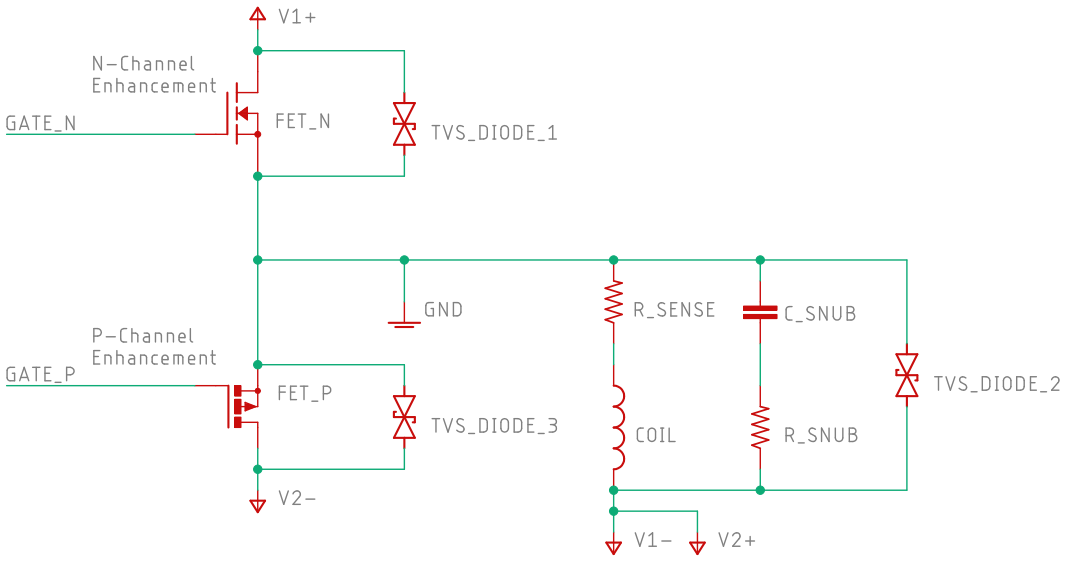


Figure 3-17: Connections for a bipolar coil driver using a floating power supply and a snubber. Suggested locations for transient voltage suppression (TVS) diodes are included as well.

For reference, a schematic showing a bipolar coil driver setup with floating power supply and snubber is presented in Fig. 3-17. It also includes suggested connections for transient voltage suppression (TVS) diodes. Those can be added to limit the large voltages across devices which otherwise could occur when the current is changed quickly due to a coil's large inductance.

3.3.3 Future Improvements

Although the design changes discussed in the previous subsections have greatly improved the performance of our coil drivers, there are still ways in which they can be improved. This subsections discusses a few of those limitations.

A limiting factor on how fast the coils can change current is set by the gate

capacitance of the MOSFETs. Changing the current requires changing the MOSFET gate voltage, which in turn requires adding or removing charge from the MOSFET gate. The amount of charge needed is determined by the MOSFET gate capacitance, and the rate at which charge can be moved is limited by the output capabilities of the PID controller, particularly its maximum output current. Therefore the PID controller maximum output current and MOSFET gate capacitance set a limit on how quickly the coil current can be changed. Additionally, the PID controller output impedance can interact with the capacitances of the MOSFET in a somewhat strange way. As mentioned in Section 3.3.2, the gate-drain and gate-source capacitances of the MOSFET couple the gate to the drain and source. Combining that with the finite output impedance of the PID controller creates an RC filter for the gate voltage. At low frequencies, the gate voltage follows the PID controller's output voltage. However, above some cutoff frequency, the gate voltage starts to be more strongly determined by the drain and source voltages. That can lead to strange behavior because the drain and/or source voltages depend on the gate voltage. Therefore a PID controller with larger output current capabilities and lower output impedance will generally do a better job of controlling the coil current.

Our initial PID controllers used INA111 or INA114 instrumentation amplifiers to drive their outputs, and those ICs could only source or sink currents up to a limit of 15 to 30 mA. We are currently in the process of designing a PID controller that can drive larger currents to the MOSFET gates. Dawid Paszko produced a design that uses LT1010 power buffers, which have low output impedance and are rated to drive up to 150 mA. However, that circuit is prone to bursts of oscillations during changes in the setpoint. This is possibly due to insufficient power supply bypassing for those large-current devices, but further investigation is required. A simpler approach is to use high-current op amps to drive the MOSFET gates, which is another avenue that we are investigating.

Another imperfection of the current system is associated with the nonlinear relationship between gate voltage and drain-source current for MOSFETs. The rela-

tionship is approximately quadratic [13], so the derivative goes to zero at small gate voltages but is larger at larger gate voltages. This changing derivative implies that the feedback loop gain is nonlinear; the loop gain (defined for small perturbations) is larger at large coil current setpoints and smaller at small coil current setpoints. That in turn implies that the bandwidth is larger at larger currents and smaller at smaller currents. This is not ideal because the bandwidth has some upper limit at large currents, and then the bandwidth at small currents is much lower. That slows the response of the system at small currents. The gain could be increased to improve the performance at small currents, but then that would increase the gain at high currents as well which would make the system unstable.

One way to improve the bandwidth at small currents is to introduce a nonlinear element to (at least partially) cancel out the nonlinearity of the MOSFET. One possibility would be to introduce an element with a square-root relation between its input and output voltages, which could cancel the quadratic relation of the MOSFET. That is possible with certain ICs [14] but may require some bandwidth and noise considerations. A simpler possibility may be to use a non-linear gain stage built with Zener diodes in a typical op amp amplifier. For example, consider a Zener diode in series with a resistor. That has large impedance below the Zener diode's breakdown voltage, then an impedance approximately equal to the resistor's resistance above the breakdown voltage. Using one or more Zener diode and resistor pairs in place of single resistors in a standard inverting or non-inverting op amp circuit can make the gain approximately step between different discrete values. That occurs whenever the voltage crosses the breakdown voltage of one of the Zener diodes. Appropriate choice of Zener diodes and resistors can allow for a rough cancellation of the MOSFET's nonlinearity, which can be sufficient and may be much easier to implement than a true square-root circuit.

As an aside, there is a bit of an upper bound on the rate at which the magnetic field at the position of the atoms can be easily changed. That rate is limited by eddy

[13] See e.g. Section 3.04 of Ref. [91].
[14] This is possible e.g. with an MPY634. The AD538, AD543, and AD535 are other options.

currents in the vacuum chamber and other components, which will oppose changes in magnetic field due to Lenz's law. Those eddy currents then die down with a timescale set by the resistivity of the component's material, and that determines how quickly the magnetic field settles to a new value. This is likely a limiting factor for our system for setpoints at which the coil current feedback loop bandwidth is high. However, we are likely still limited by the feedback loop bandwidth at setpoints where the bandwidth is lower, so increasing the feedback bandwidth may still lead to improved magnetic field control. If magnetic field changes faster than the eddy current timescales are required, then feedforward control signals can be used to compensate for the eddy currents if they are well-characterized. Alternatively, in-vacuum coils can potentially circumvent that issue altogether, though adding those to an experiment is a nontrivial task.

3.4 Control System

Previous generations through our lab have used a custom LabVIEW-based application for to control the experiment. The application allowed users to specify output waveforms by individually entering the times at which an output should change and to what value it should change. Half of the analog channels would always ramp linearly between consecutive values, and the other half of channels would discretely jump to new values. The times and outputs values were entered in a spreadsheet-like interface in which only one row of the spreadsheet was viewable at a time. This system made it possible to do all the research that our lab produced before this thesis.

However, the old control system left a lot to be desired in many ways. While it had some tools for editing sequences, such as the ability to shift all edges after a certain time by some fixed amount, doing anything but simple edits to a sequence was very tedious. It also was not possible to combine parts of different sequences. For example, at one point I managed to perform a Stern-Gerlach type of experiment, separating different Zeeman states with a magnetic field gradient during imaging which was a helpful diagnostic for Raman cooling. However, doing so required manually entering

over one hundred pairs of output times and values. It was too tedious to insert and remove that from sequences, so we had to keep two copies of the sequence, one with and one without the Stern-Gerlach imaging, and then we had to manually synchronize any changes between the two copies. Of course this didn't last long and we soon lost the ability to perform that kind of measurement. Additionally, the LabVIEW-based control system was written based on the deprecated NI Legacy DAQ drivers, which aren't available on versions of Windows newer than Windows XP. This meant that our control system computer had to run 32 bit Windows XP, which stopped receiving security updates and could no longer be connected to the network safely. This was problematic as it implied that the computer controlling the camera and the computer used for data analysis could not talk to the system that controlled the rest of the machine. That had many negative consequences, among which was the fact that the software used to communicate with the camera had to be manually synchronized with the control software on the Windows XP machine by starting both in the proper order at nearly the same time. Even if that was done successfully, other issues could desynchronize the two machines and corrupt the data. Furthermore, any information about the sequence which was needed for the analysis had to be entered manually, which was very prone to errors.

For these reasons and many others, it was decided that a control system upgrade would be well worth the investment. After looking at many different options I settled on the free and open-source Labscript [85, 86, 87] suite. It has support for our NI cards [15] and is extensible, so we could incorporate our camera and any other desired hardware into it. All of that meant there was no financial investment to install Labscript and try it out. Labscript also allows users to write their sequences in Python, which makes it particularly easy to do. Variables, functions, control flow, and other standard programming capabilities can be used to generate sequences, which is powerful. Using Python also makes it possible to leverage the power of many tools built for software development when writing sequences. This includes version tracking sys-

[15] Support for the NI cards which were still supported by the new NI DAQmx drivers. Not surprisingly our NI PCI-6034E analog input card, which NI dropped support for about 20 years ago, had to be abandoned.

tems, debuggers, linters, syntax highlighters, third party libraries, and more. All of this together makes writing sequences much easier and more flexible, which makes researchers more productive. The Labscript suite also includes several easy-to-use GUI components which are used to build, view, run, and analyze sequences. This makes not only easier to construct sequences, but easier to use them as well. Furthermore, Labscript can run on more modern operating systems, so we could upgrade the control computer and that enabled us to more safely connect it to the internet. This turned out to be critical during the COVID-19 pandemic when the lab had to be run remotely. For these reasons and more, the upgrade to Labscript significantly boosted lab productivity and made more research possible.

Labscript's flexibility made it possible to implement a few other upgrades as well. One of which was the implementation of Lock Monitor, a plugin I wrote for blacs (a software component of the Labscript suite) designed to keep lasers in lock. Lock Monitor was mentioned earlier in this chapter and is the subject of Appendix C. Additionally, some of the Labscript developers had written code [92] to use Labscript with M-LOOP [32, 33], an open-source Bayesian optimization package developed for AMO experiments. Both M-LOOP and the Labscript integration code for it were installed onto our system as well. This was a critical upgrade for a few reasons. Firstly, M-LOOP found parameters for our MOT loading and compression sequence which significantly outperformed the ones we found by hand. That reduced our MOT loading time from about 1 second to 100 ms, a factor of ten improvement which significantly sped up our data acquisition. That new optimized sequence also loaded more atoms than the slower old one. Furthermore, a lot of our Raman cooling research is essentially an optimization problem where our goal is to find a good set of cooling parameter values in a high-dimensional parameter space. M-LOOP performs that task automatically. This meant that the control system could effectively do our research on its own, which freed me to perform other tasks and significantly boosted productivity. It even meant that research continued while I slept.

Primarily thanks to these upgrades, the data production rate of the lab skyrocketed. For comparison, the backups of all of our data from 2015 to 2019 require

≈400 GB of hard drive space. By contrast, the backups for our data from February to May in 2021 alone require ≈1.25 TB of hard drive space. The hard disk space usage per shot has increased by about a factor of two because we now save more information with each shot, but we still managed to run more shots of the machine in four months of 2021 than we had in four previous years combined. What's more is that we typically used to run the machine with two people [16] in lab continuously interacting with it. The 2021 data was acquired either with one person running the machine remotely, or with no human intervention at all for hours on end. That meant that these upgrades gave us ten times the data rate with less than half the person-hours despite restrictions from the COVID-19 pandemic.

3.5 Other Upgrades

In addition to the upgrades mentioned in the previous few subsections, many other smaller or less-critical upgrades were also made. A selection of those upgrades are listed in this section.

3.5.1 Shutter Driver Upgrades

Our lab switched from using commercial shutters to using a 3D-printed shutter design developed by our group [93, 90]. The shutters themselves work very well and are economical and easy to assemble. However, the shutter driver circuit was found to imprint significant noise into the rest of the system.

The shutters themselves are made with an electric motor with a blade on its axel and stops. When current runs through the motor in one direction, the motor rotates the blade clockwise until it hits one of the stops. Reversing the current rotates the motor the other way until the blade hits the other stop. This design allows the motor (which would normally spin continuously) to instead rotate between two fixed positions, changing from one to the other when the current through it is reversed. The

[16] It was possible for one person to run the machine on their own, but this slowed data taking due to the amount of manual intervention required. For this reason we typically ran it with two people.

beam is then directed through a region where the blade blocks it when the shutter is in one of its resting positions, but doesn't block it in the other resting position. The shutter driver's job is then to drive current through the motor in one direction or the other based on the state of a TTL logic control input. This is achieved with an H-bridge and a NOT gate.

It was found that the design of the driver presented in Ref. [93] sometimes would imprint volts of noise onto the grounds of the TTL control lines in our experiment. This occurred with bursts of high-frequency noise when the shutter changed state. The TTL control signals are generated by an NI card (NI PCI-DIO-32HS) which is used to control many other devices in our lab, including the camera used for acquiring absorption images of our atomic clouds. The noise level from the shutter driver was large enough to occasionally mistrigger our camera, which was problematic. The noise level was not always repeatable; it would vary when moving around the shutter power supply cables, changing the shutter drive voltage, and would vary for different drivers with identical designs.

The original driver design did not include power supply decoupling capacitors, neither for the motor supply lines nor for the logic supply. A large contribution to the noise when switching states seemed to be associated with the under-voltage protection of the H-bridge, which is designed to disconnect its outputs if the motor supply voltage input drops below a threshold level. When a shutter would change state it would suddenly draw a spike in current, which could drop the supply voltage because it was not adequately bypassed. The drop in supply voltage could trigger the under-voltage protection of the H-bridge, disconnecting the motor. Once the motor was disconnected, the supply line voltage would increase again, disabling the under-voltage protection and reconnecting the motor. That could then drop the supply voltage again, retriggering the under-voltage protection. This process could happen many times before settling down, causing significant noise in the process.

Adding decoupling capacitors on the power supply lines significantly reduced the noise imprinted onto the control lines, enough so that mistriggering the TTL inputs of other devices in the system was no longer an issue. However, some noise was still

imprinted onto the control lines which could potentially affect analog signals in the system. It was deemed worthwhile to more carefully isolate the TTL control inputs to the shutter drivers from the noisy signals inside the controller. Doing so would keep the noise contained to the shutter driver where it was not causing any harm and protect the rest of the system from it. This was achieved using TLP2367 optocouplers.

Optocouplers consist of an LED and a phototransistor (or other light-sensitive device) assembled into one IC package. An input drives the LED, turning it on or off, which then turns the phototransistor on or off. This setup makes it possible to send a signal from one side of the IC to the other, even when the two sides of the circuit are referenced to different grounds. This enables sending a TTL signal from one system to the other while keep the systems well-isolated. The TLP2367 includes an internal Faraday cage and can handle common mode voltage differences of 3.75 kV, which demonstrates its isolation capabilities.

Many optocouplers require external drive components on the receiving side in order to convert the phototransistor's conductance into a voltage logic level. However the TLP2367 was chosen because it includes those components inside the IC. It can directly source or sink current without any additional external components, which greatly simplifies the design. The input side of the TLP2367 requires only a resistor, the resistance of which is chosen to obtain the appropriate LED current for the given input voltage logic level. For example, a 3.3 V logic system and a 5 V logic system would require a different resistance in series with the LED, which is why it is left as an external component rather than included in the IC. The TLP2367 datasheet suggests that small capacitor is added in parallel with the resistor to improve high-frequency response, though this is likely not necessary for the shutter driver as it intrinsically has a slow response due to the time it takes for the motor to rotate.

The upgraded shutter driver design consists of four identical channels and the schematic for one channel is included as Fig. 3-18. The resistor and capacitor in series with the TLP2367's LED are chosen for 5 V TTL logic in accordance with the suggested values from the TLP2367 datasheet; see the datasheet's suggested values for other logic levels if necessary. The design also includes a few additional passive

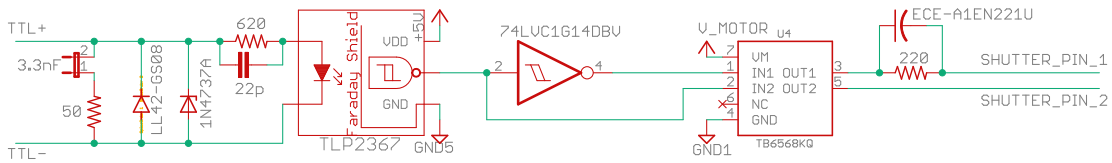


Figure 3-18: Schematic for a single channel of the upgraded shutter driver. The TTL logic control input is applied on the left. A 3.3 nF capacitor in series with a 50 Ω resistor are used to provide AC-coupled 50 Ω line termination, which should work well for control lines much less than 300 m long. A Schottky diode (LL42-GS08) protects the optocoupler (TLP2367) input from reverse-biasing and a 7.5 V Zener diode (1N4737A) protects it from over-biasing. The 620 Ω resistor and 22 pF capacitor are the components suggested from the optocoupler datasheet for 5 V TTL logic operation. The optocoupler output and a NOT gate (74LVC1G14DBV) control an H-bridge (TB6568KQ), which drives the shutter motor connections on the right of the schematic. The 220 Ω resistor sets the holding current for the motor after it has rotated and the ECE-A1EN221U bipolar 220 μF capacitor allows for a large spike in current to accelerate the motor quickly when the shutter changes state.

components at the TLP2367 input to provide line termination and to protect it from accidental reverse-biasing or over-voltage conditions. The design uses AC-coupled termination to reduce the current draw during the long periods when the control input isn't changing. Frequency components well above the RC corner frequency of those components ($\approx(2\pi)1$ MHz for the values in Fig. 3-18) will see an $\approx$50 Ω load and be properly terminated. In practice this means that transmission lines with lengths small compared to $c/(1\text{ MHZ}) \approx 300$ m will be well-terminated. The schematic suggests a 1N4737A 7.5 V Zener diode for over-voltage protection, though a TVS diode may be better suited for the task. A LL42-GS08 Schottky diode is suggested for reverse bias protection, though its round packaging can make it somewhat tedious to keep in place while soldering. Other Schottky diodes would work equally well and may be easier to solder; the LL42-GS08 was selected primarily because our lab had it in stock. The H-bridge suggested in Ref. [93] (TA8428KS) has gone out of production, but the TB6568KQ and TB6643KQ are pin-identical drop-in replacements.

3.5.2 Shutter Mounting

When shutters are actuated, they tend to produce mechanical vibrations which can affect the rest of the apparatus on the optics table. To mitigate that issue, vibration damping material is typically placed between the shutter and the optics table.

Previously, our lab would insert a vibration damping layer by cutting out Sorbothane pieces in the shape of the optomechanics bases used to mount the shutters, e.g. Thorlabs BA1S. Those pieces were then inserted between the base and the optics table, then a nylon screw was used to secure the base to the table. This approach had some issues. Firstly, the nylon screw is more rigid than the Sorbothane and likely increased the coupling of vibrations to the table. Secondly, the Sorbothane could get wrapped around screw, making it difficult to turn and tearing up the Sorbothane piece in the process. Furthermore, this approach made it very difficult to align the shutters. The Sorbothane sticks to the table and makes it difficult to slide the base around to position the shutter horizontally. Tightening the nylon screw compresses the Sorbothane, which moves the shutter up and down. If the screw is not perfectly centered, then tightening the screw also tilts the shutter mount as well. Adjusting anything on the mount usually involves pushing or pulling on the mount slightly, which could also displace the shutter. These effects made it very tedious and time consuming to align a shutter to a beam.

A solution to the issues with the Sorbothane base pads was to use vibration-damping sandwich mounts instead, as depicted in Fig. 3-19. These mounts have two threaded sides separated by a rubber spacer. The threaded components are not mechanically connected by metal; they only connect through the vibration-damping rubber. They are available with various thread types, as well as various spacer material and geometry. One side of a sandwich mount can be screwed into the top of a standard optics post, then the shutter can be mounted to the other side. This secures the shutter in place while providing a damping layer between it and the optics table. Unlike the Sorbothane base pads, the sandwich mounts are placed above all of the components that are adjusted to position the shutter. This alleviates all of the issues

(a)

(b)

Figure 3-19: Vibration damping sandwich mounts. (a) A single sandwich mount. (b) A shutter mounted with a sandwich mount between it and the optics post to provide vibration damping while allowing for easy alignment.

mentioned in the previous paragraph, which makes it much faster and easier to align the shutter to a beam.

When testing different sandwich mounts it was found that very soft sandwich mounts rated for small loads had some hysteresis in their positioning. If the shutter was pushed to one side or other, then it would spring back somewhat, but not all the way to its previous position. That effect could misalign a shutter from its beam if it were accidentally bumped, which is undesirable. Somewhat stiffer sandwich mounts did not have this issue. We found that McMaster 9213K12 Neoprene or the McMaster 9376K17 natural rubber sandwich mounts are good choices. They are rated for 20 lb loads and do not show significant hysteresis when they are pushed. They also have have male #8-32 threads on each side, which can be threaded into standard Thorlabs 1/2 inch optics posts and can fit through the shutter's mounting hole. Tightening the sandwich mounts can be difficult because they do not provide a solid section to turn with a wrench. Instead a single nut is placed on the optics post side, which can be tightened against the optics post to secure the sandwich mount in place. On the shutter side, two nuts are used. One is placed on each side of the shutter's mounting hole, then the nuts can be tightened against each other with the shutter in between to hold it in place. These nuts can be seen upon close inspection of Fig. 3-19(b).

Possible future upgrades include the design of a shutter enclosure. This can prevent users from accidentally pushing or pulling on the shutter itself, thereby avoiding the hysteresis issues mentioned above. That would enable the use of softer sandwich mounts which may provide better damping. The enclosure could also provide a mounting point for the electrical connector, which could alleviate issues associated with tugging on the cable. As of this writing, Luke Stewart is in the process of designing an enclosure along these lines.

3.5.3 Vapor Cell Source

Previously our apparatus used dispensers to provide the rubidium for the experiment. Three of them were mounted in the chamber, but they all eventually ran out of rubidium over ≈15 years of use. A rubidium vapor cell source, which is a glass

tube with a sizable chunk of solid rubidium in it, was also attached to the chamber. However, previous efforts to produce a workable vapor pressure from it were not successful.

The vapor cell was at the end of a long bellows, and it was suspected that the bellows restricted the flow of rubidium too much so the vapor pressure in the MOT region was too low. To remedy that, the chamber was opened and the vapor cell was repositioned. A tee was added to one of the diagonal MOT beam flanges, then the vapor cell's valve was attached to the side of the tee. This brought the vapor cell much closer to the MOT region of the chamber. After this change it was able to produce sufficient rubidium vapor pressure to run the experiment, even without heating the cell. The flow of rubidium can be restricted by partially closing the valve if desired. Doing so reduces the rubidium vapor pressure in the chamber, which increases the vacuum lifetime at the cost of reducing the MOT loading rate. The valve can also be used to seal off the vapor cell when the camber is brought to atmosphere.

3.5.4 Cable Management

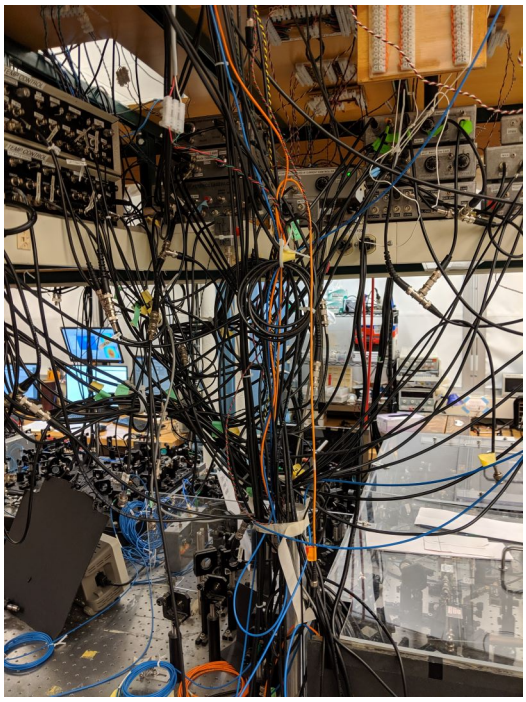
Cable management in a lab is often regarded as unimportant; it seems like it shouldn't matter where a cable goes as long as it starts and ends at the right places. Over the years our lab has repeatedly learned the hard way that this is an oversimplification. Proper cable management is not a matter of aesthetics, it is a matter of flexibility and reliability.

A first step towards better cable management is to label cables, particularly those carrying signals. Label makers are cheap and easy to use, so there is little cost or effort involved in labeling cables. Doing so frequently comes in handy, particularly when a device needs repair. When that happens, the cables typically need to be disconnected so that the device can be removed. If the cables are not labeled, then reconnecting the device is a very slow process. Each cable needs to be traced back to its source to determine where it should be connected on the reinstalled device, which is tedious and error prone. In some situations components can even be damaged if the wrong connection is made. When cables are labeled, reinserting a repaired device

goes much more quickly and smoothly and is less prone to misconnections. Given that testing and repairing devices is a frequent occurrence in research labs, labeling cables is a time investment that pays for itself rather quickly and many times over.

In addition to labeling cables, routing them properly is important as well. Tangled nests of cables are difficult to work with. I've found that people will often pull on cables and devices to move them rather than carefully disentangling and rerouting the cables. That pulling often damaged connectors and jacks, which cause intermittent connection issues. Such issues are notoriously difficult to debug because they are not easily reproducible. Simply connecting an oscilloscope or other device can involve perturbing a sensitive connection, causing the issue to appear or disappear randomly. That makes such debugging efforts very time consuming. When cables are more neatly organized I've found that this is less of an issue. If someone attempts to move a device and one of the well-organized cables is too short, they will simply see that this is the case and disconnect it, then safely move the device. By contrast, when the cables are disorganized, there's a possibility that pulling on one harder might drag more slack out of the tangled nest of cables. In that case people will often try yanking on cables and inflict damage in the process. Messes of cables can also lead to noise issues. Fig. 3-20 shows images of the cabling below our 780 nm laser electronics before and after reorganization. Before reorganizing the cables, there was ∼100 mV of 80 MHz noise on much of the 780 nm control lines. This was presumably due to some undesired coupling to the 80 MHz signals in our AOM driver systems. After some amount of searching I was not able to find the source of the coupling. However, the noise disappeared after organizing the cables and has yet to show up again even years later.

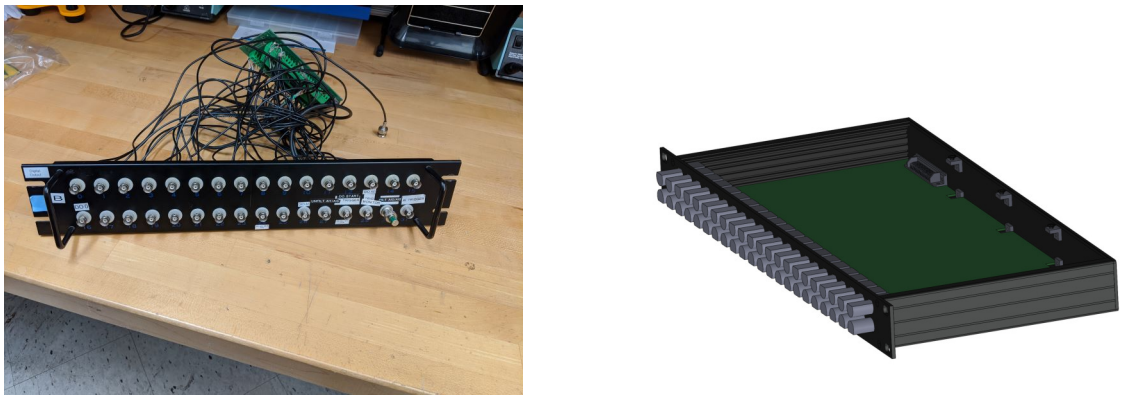
Another set of components with cabling issues were the breakouts for the NI card outputs. These constituted the interface between the control system and the experiment, but the connections were poorly labeled, sometimes broken, and occasionally missing completely. The original homemade versions were made using screw terminal connectors (NI CB-68LPR) without wire ferules and pigtailed BNC connectors without strain reliefs, making them prone to failure. They also took up more more rack



(a)

(b)

Figure 3-20: Before and after photos of the 780 nm laser system cabling reorganization. Although the reorganized cables are certainly more appealing aesthetically, the more important improvement is performance-related. The large amount of 80 MHz noise which was present on the 780 nm laser electronics signals before the reorganization disappeared after the cables were organized.



(a) (b)

Figure 3-21: (a) Old BNC breakout for our NI PCI-DIO-32HS card used for digital outputs. The cabling is less than organized, is missing connections for some pins, and can be prone to connection failures. The front panel has limited labeling and the numbers printed on the panel do not match the channels that they are connected to. (b) CAD rendering of the replacement breakout with the top panel hidden to make the PCB visible. The breakout takes half the rack space of the original yet provides connections for all of the NI card's pins. The connections are done via a PCB so no wiring is required. Though not visible in this image, the front panel has engraved labels which match the pin names.

space than was necessary, which was not ideal due to the limited rack space available on the upper canopy where the control computer resides. One such breakout is shown in Fig. 3-21(a). Some commercial breakouts (NI BNC-2110) were also used, but they did not provide BNC jacks for all of the desired connections. They also were not rack mounted and instead were left on the canopy in front of the control computer, prohibiting us from pulling out the computer on its rack slides.

To alleviate some of those issues, a new custom breakout for our NI PCI-DIO-32HS digital output card was designed. Its electrical connections are made with a PCB to avoid wiring difficulties and to keep things organized. An enclosure was designed which supports and protects the PCB and has a rack-mountable front panel. The front panel also has engraved labels which match the names of the corresponding card pins, and a connection is provided for every pin. The enclosure panels are designed to be produced using a CNC mill and can be ordered from Front Panel Express. The other components are stock parts, also available from Front Panel Express. This makes it easy to produce additional breakouts if another card is added, or if another

lab needs a breakout as well. The end result is a compact, reliable, easy-to-produce, well-labeled breakout. More recently, Michelle Chong has made a similar breakout for our NI PCI-6713 analog output cards, which will further mitigate the cabling issues around our control computer.

Chapter 4

Direct Raman Cooling to BEC

The following is a reproduction of Ref [28].

4.1 Main Text

4.1.1 Abstract

We present a method for producing three-dimensional Bose-Einstein condensates using *only* laser cooling. The phase transition to condensation is crossed with 2.5×10^4 ^{87}Rb atoms at a temperature of $T_c = 0.6$ μK after 1.4 s of cooling. Atoms are trapped in a crossed optical dipole trap and cooled using Raman cooling with far-off-resonant optical pumping light to reduce atom loss and heating. The achieved temperatures are well below the effective recoil temperature. We find that during the final cooling stage at atomic densities above 10^{14} cm^{-3}, careful tuning of trap depth and optical-pumping rate is necessary to evade heating and loss mechanisms. The method may enable the fast production of quantum degenerate gases in a variety of systems including fermions.

4.1.2 Introduction

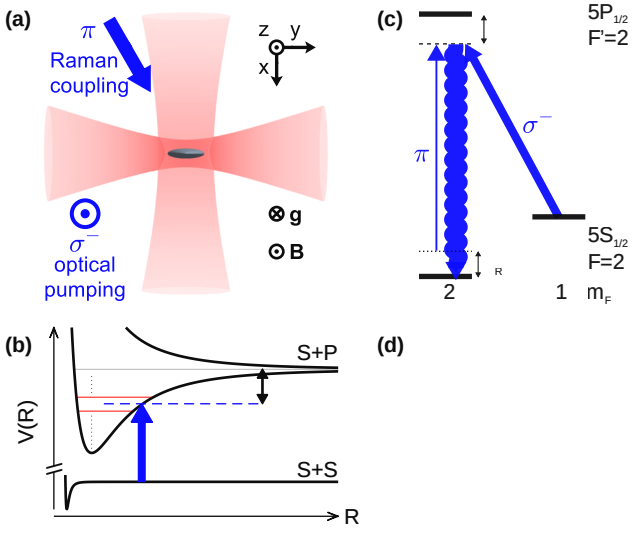
Quantum degenerate gases provide a flexible platform with applications ranging from quantum simulations of many-body interacting systems [9] to precision measure-

ments [12]. The standard technique for achieving quantum degeneracy is laser cooling followed by evaporative cooling [57] in magnetic [6, 7, 94] or optical traps [95]. Evaporation is a powerful tool, but its performance depends strongly on atomic collisional properties and it requires removal of most of the initially trapped atoms. Laser cooling gases to degeneracy could alleviate those issues, but it has proven difficult to implement.

The elusiveness of laser cooling to Bose-Einstein condensation (BEC) [6, 7] for more than two decades [19, 42, 39, 45] can be understood as follows: optical cooling requires spontaneous photon scattering that moves entropy from the atomic system into the light field. Such scattering sets a natural atomic temperature scale T_r associated with the recoil momentum from a single photon of wavelength $2\pi\lambda\!\!\!^{-}$ and an associated atomic thermal de Broglie wavelength $\lambda_\mathrm{dB} = \sqrt{2\pi\hbar^2/(mk_\mathrm{B}T_\mathrm{r})} \sim \lambda\!\!\!^{-}$. Here $\hbar$ is the reduced Planck constant, m the atomic mass, and k_B the Boltzmann constant. BEC must then be achieved at relatively high critical atomic density $n_\mathrm{c} \sim \lambda_\mathrm{dB}^{-3} \sim \lambda\!\!\!^{-3}$, where inelastic collisions result in heating and trap loss. In particular, light-induced collisional loss becomes severe when $n \gtrsim \lambda\!\!\!^{-3}$ [27, 96].

For strontium atoms, where cooling on a spectrally narrow transition is available, a strongly-inhomogeneous trapping potential has been used to cool a lower-density thermal bath while decoupling the emerging condensate from the cooling light [23]. Very recently, based on similar principles to the ones presented here, an array of small, nearly one-dimensional condensates has been prepared using degenerate Raman sideband cooling [51, 34, 45] of ^{87}Rb atoms in a two-dimensional optical lattice [24].

In this Letter, we demonstrate Raman cooling [22, 41, 42, 43, 44] of an ensemble of ^{87}Rb atoms into the quantum degenerate regime, without evaporative cooling. Starting with up to 1×10^5 atoms in an optical dipole trap, the transition to BEC is reached with up to 2.5×10^4 atoms within a cooling time of $\sim$1 s. As discussed in detail below, the essential components of our technique are (i) the use of carefully far-detuned cooling light to reduce atom loss and heating at high atomic densities [27, 97, 98], (ii) a reduced optical pumping rate during the final stage to avoid heating by photon reabsorption (*festina lente* regime [58, 59, 60]), (iii) a final cooling of atoms


(a)
π
Raman
coupling
z
y
x
⊗ g
⊙ B
σ⁻
optical
pumping
(c)
5P1/2
F'=2
π
σ⁻
5S1/2
F=2
2
1
mF
R
(b)
S+P
V(R)
S+S
R
(d)

in the high-energy wings of the thermal velocity distribution to achieve sub-recoil cooling, and (iv) careful control of the final trap depth to reduce heating induced by inelastic three-body collisions.

Raman cooling of the optically trapped atoms is a two-step process where kinetic energy is first removed via a stimulated two-photon Raman transition that simultaneously changes the internal atomic state. Subsequently, entropy is removed in an optical pumping process that restores the original atomic state via the spontaneous scattering of a photon [see Fig. 4-1(c)-(d)]. The optical pumping into the state $|5S_{1/2}; F{=}2, m_F{=}-2\rangle$ along the z-axis is performed with σ^--polarized light. We reduce light-induced loss by using optical pumping light with large negative detuning $\Delta/(2\pi) = -4.33$ GHz from the $|5S_{1/2}; F{=}2\rangle$ to $|5P_{1/2}; F'{=}2\rangle$ transition of the D_1-line, choosing a detuning far from molecular resonances [see Fig. 4-1(b-c) and 4.2]. The far-detuned σ^--polarized beam and a π-polarized beam of similar detuning which propagates in the $x-y$ plane [see Fig. 4-1(a)], drive the stimulated Raman transition to the state $|5S_{1/2}; 2, -1\rangle$, simultaneously changing the atomic momentum by the two-photon recoil $\hbar(\Delta\mathbf{k})$ 4.2. To cool all three directions, we choose $\hbar(\Delta\mathbf{k})$ to have a non-zero projection along any trap eigenaxis.

4.1.3 Methods

Cooling is performed in several stages to allow optimization of the cooling as the atomic temperature and density change. Within each stage, the trapping beam powers, optical pumping rate Γ_{sc}, Raman coupling Rabi frequency Ω_R, and Raman resonance detuning δ_R are held constant. The first two cooling stages, S1 and S2, are performed in a single-beam optical dipole trap (sODT), after which the atoms are transferred to a crossed optical dipole trap (xODT), where we perform three more cooling stages [X1 to X3, see Fig. 4-2(a)]. During each stage, we characterize the cooling performance using time-of-flight absorption imaging, extracting the atom number N and temperature T. (For the final cooling stage close to the BEC threshold, we exclude the central part of the time-of-flight image from the temperature fit.) We quantify the cooling performance by the classical phase space density

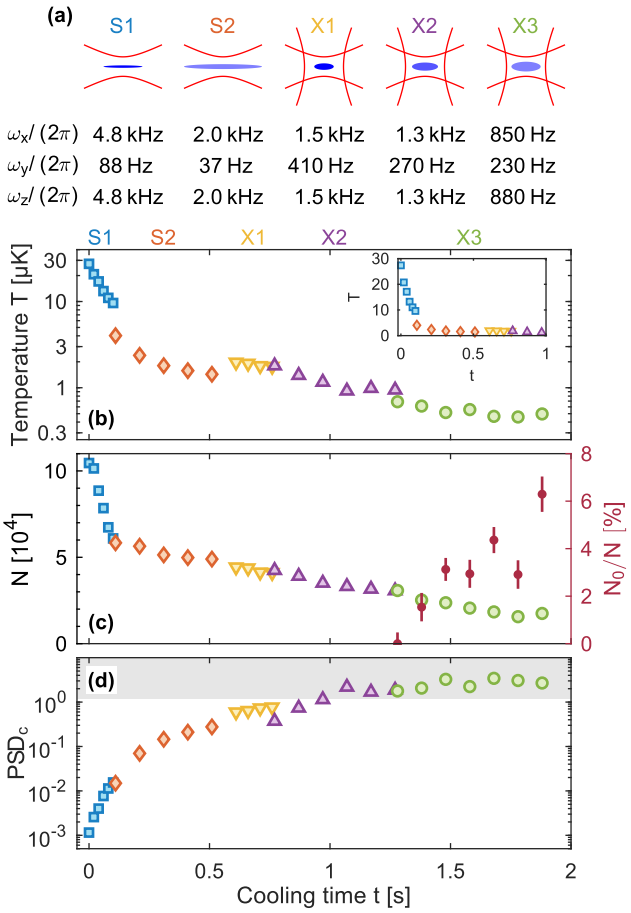


Figure 4-2: (a) Schematic of the trapping potential during the cooling sequence, along with the values of the trapping frequencies for each cooling stage. (b) Atomic temperature T as a function of cooling time t. Discrete jumps are caused by changes of the trapping potential between the cooling stages. Inset: Temperature on a linear scale. (c) Atom number (open symbols) and condensate fraction N_0/N (solid circles) during the cooling sequence. (d) Classical phase-space density $\mathrm{PSD_c}$ (see main text) as a function of cooling time t. The grey shaded area denotes the quantum degenerate region. Subfigures (b-d) are all plotted along the same time axis.

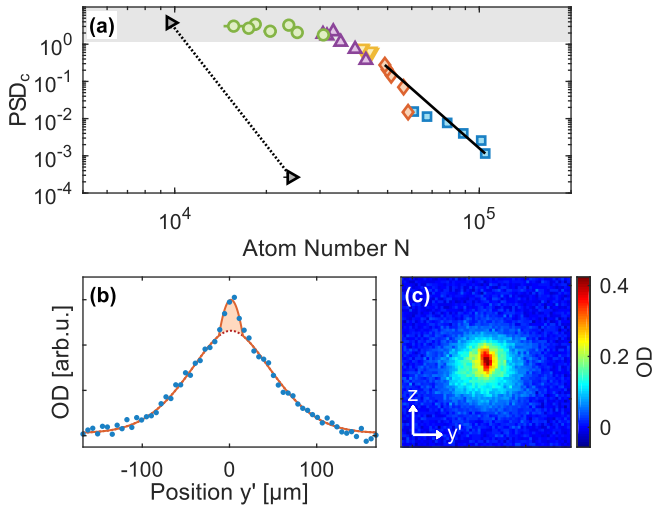


Figure 4-3: (a) Classical phase-space density $\mathrm{PSD_c}$ as a function of remaining atom number N. The cooling is very efficient until $\mathrm{PSD_c}{\sim}1$ is reached. The black symbols denote the performance of the same sequence with the initial atom number reduced by a factor 5. The final atom number is only reduced by a factor 2, indicating a density limit in the cooling. The solid (dashed) black line indicates the S1-S2 (S1-X3) path with an efficiency $\gamma = 7.2$ ($\gamma = 11$) for each case. (b) Line optical density (dots) of the cloud in time-of-flight along the y'-direction (slightly rotated from the y-direction in the $x-y$ plane, see SM 4.2). The data is taken after 1.6 s of cooling and fitted with a $g_{5/2}$ Bose distribution with a Thomas-Fermi distribution superimposed (orange line). The shaded area indicates the condensed fraction. (c) False color image of the same cloud (before integration along the vertical direction), showing the anisotropic expansion of the condensed fraction in the center.

$\mathrm{PSD_c} = N\left(\hbar\bar{\omega}/(k_\mathrm{B}T)\right)^3$, where $\bar{\omega} = (\omega_x\omega_y\omega_z)^{1/3}$ is the geometric mean of the three trapping frequencies. Far from degeneracy, i.e. for a classical gas, $\mathrm{PSD_c}$ is equal to the true PSD at the center of the cloud, $\mathrm{PSD} = n(0)\lambda_\mathrm{dB}^3$. The parameters of each stage are optimized to yield the highest $\mathrm{PSD_c}$ at the end of the stage 4.2.

4.1.4 Results

We load $1{\times}10^5$ atoms from a magneto-optical trap into the sODT propagating along the y-direction with a 10 μm waist [Fig. 4-1(a)]. After cooling in stage S1 (see Fig. 4-2), the trap power and vibration frequencies are reduced, thereby lowering the density and therefore the loss in stage S2. For all stages, we verify that the trap remains sufficiently deep to keep evaporative cooling negligible.

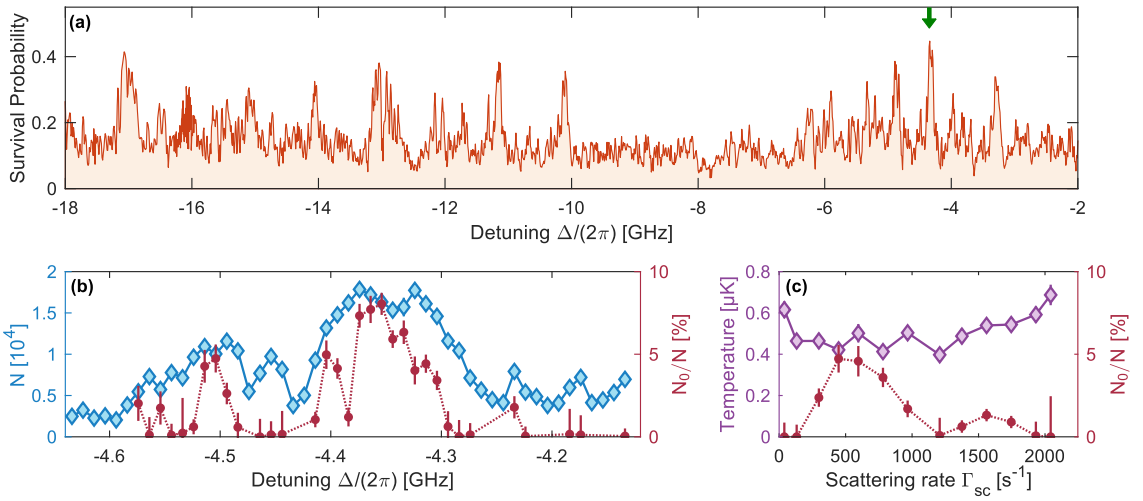


Figure 4-4: (a) Photoassociation loss spectrum. Survival probability of trapped atoms as a function of the detuning Δ of the optical pumping beam, when scattering $\sim$100 photons. In substantial portions of the spectrum, the atomic loss is large, due to photoassociation resonances, whereas the peaks correspond to gaps in the photoassociation spectrum away from resonances. The green arrow indicates the detuning used for the data in Fig. 4-2 and 4-3. (b) Performance of the full cooling sequence as a function of optical pumping detuning Δ near a locally optimal detuning. A condensate fraction N_0/N is visible only when the losses are limited. To keep the Raman resonance detuning δ_{R} constant between data points, the magnetic field is adjusted to compensate the change in light shift associated with varying Δ. (c) Temperature and condensate fraction as a function of the scattering rate Γ_{sc} in the final cooling stage. Here, the intensity of the π-polarized beam is adjusted to keep the Raman coupling strength constant between data points, and δ_{R} is adjusted to optimize the cooling performance for each data point.

During stages S1 and S2 we perform fast cooling over 500 ms from 30 μK down to 1.5 μK, and up to $\mathrm{PSD_c}$ just below unity. The larger loss rate during S1 relative to the other stages [see Fig. 4-2(c)] occurs because the Raman cooling cycle, and hence light-induced collision rate, is faster. The initial cooling at high temperatures $T \approx 30$ μK and densities $n < 7\times10^{13}$ cm^{-3} is quite efficient, with a logarithmic slope of $\gamma = -d(\ln \mathrm{PSD_c})/d(\ln N) = 7.2$ [see Fig. 4-3(a)].

After stage S2, the ensemble is sufficiently cold so that it can be efficiently transferred in the xODT by ramping up the second trapping beam (18 μm waist, propagating along x), and applying a short initial cooling (stage X1). In a similar fashion to the S1-S2 sequence, we reduce the confinement of the xODT and cool further during X2, after which the ensemble is at the threshold to condensation. No condensate appears in X2 despite $\mathrm{PSD_c}$ reaching the ideal-gas value of 1.2 [99], which we attribute to a combination of the finite size effect [100], the interaction shift [101] and small calibration errors. After further reduction of the xODT, we are able to cross the BEC transition during X3, as shown by the appearance of a condensed fraction in the velocity distribution [see Fig. 4-3(b)]. The onset of BEC is further confirmed by the anisotropic expansion of the central part of the cloud due to trap confinement anisotropy [see Fig. 4-3(c)].

In order to achieve BEC, the trap depth during X3 must not be too large. For trap depths much larger than $\sim k_B \times 20$ μK we observe a strong, density-dependent anomalous heating when the Raman cooling is turned off. (Heating rates of up to 10 μK/s are observed at $n \approx 2\times10^{14}$ cm^{-3} with a trap depth of $k_B \times 250$ μK.) We surmise that at high atomic densities $n \gtrsim 10^{14}$ cm^{-3}, recombination products of inelastic three-body collisions undergo grazing collisions with trapped atoms, depositing heat in the cloud 4.2. This limit on the maximum trap depth is akin to the necessity to maintain, even in the absence of evaporative cooling, a sufficiently low trap depth by applying a so-called "RF-shield" in magnetic traps which allows highly energetic nonthermal atoms to escape [66, 67].

The BEC transition is crossed with $N \approx 2.5\times10^4$ atoms at a critical temperature of $T_c = 0.61(4)$ μK. We are able to reach condensate fractions N_0/N up to 7%. We

also verify that if the Raman cooling is turned off during stage X3, the PSD does not increase, and a condensate does not appear. Furthermore the condensate can be maintained for ∼1 s after creation if the cooling is left on, but if the cooling is turned off, the condensate decays within ∼100 ms. This confirms that evaporation is insufficient to create or maintain a condensate in this trap configuration, and that the laser cooling is responsible for inducing the phase transition.

For most laser cooling methods, the requisite spontaneous photon scattering sets a recoil temperature limit. Nonetheless, we achieve sub-recoil temperatures by addressing atoms in the high-energy wings of the thermal distribution. The optical pumping $|2,-1\rangle \rightarrow |2,-2\rangle$ requires on average the spontaneous scattering of three photons, and therefore imparts $6E_\mathrm{r}$ of energy, where $E_\mathrm{r} = \hbar^2/(2m\lambda^2) = h\times 3.6$ kHz is the recoil energy of a 795 nm photon. As a result, only atoms with $K_{\Delta\mathbf{k}}/h \gtrsim 29$ kHz of kinetic energy along the $\Delta\mathbf{k}$-direction can be cooled at all 4.2. This sets an effective recoil temperature $T_\mathrm{r}^\mathrm{eff} = 2.8\ \mu$K. We achieve cooling below this effective recoil limit down to 0.5 μK, i.e. a mean kinetic energy $\langle K_{\Delta\mathbf{k}}\rangle/h = 5.2$ kHz, by detuning the Raman coupling so that atoms with more than the average kinetic energy are addressed by the cooling light 4.2. However, this slows down the cooling for temperatures below $T_\mathrm{r}^\mathrm{eff}$ (stage S2 onwards), while inelastic collisions add an increased heat load at high densities. In X3 when we cross T_c, we find that under optimized cooling the Raman transition removes 15 kHz of kinetic energy, 30% less than the expected $6E_\mathrm{r}$ of heating (see SM 4.2). This could indicate that the cooling is aided by bosonic stimulation into the condensate during the photon scattering process, in a similar fashion to Ref. [102].

4.1.5 Optimization

The improved performance of our scheme compared to previous Raman cooling results [42, 43] is primarily due to the flexibility to perform optical pumping to a dark state at large detuning from atomic resonances by operating on the D_1-line. To identify suitable detunings, we separately characterized light-induced losses over a 16 GHz frequency range to the red of the bare atomic transition, as shown in Fig. 4-4(a), and further detailed in the SM 4.2. Fig. 4-4(b) displays the final atom number and con-

densate fraction of the optimized sequence as a function of detuning around the value of -4.33 GHz chosen for the experiment. We find that a suitable detuning has to be maintained within ± 50 MHz to ensure good cooling performance.

Another parameter that needs to be optimized is the photon scattering rate Γ_{sc} for optical pumping into the $|2,-2\rangle$ dark state. Despite the large detuning, the reabsorption of spontaneously scattered optical pumping photons by other atoms is a resonant two-photon process that can lead to excess heating. However, it was shown theoretically [58, 59], and confirmed experimentally [61], that the excess heating can be suppressed at sufficiently low scattering rate Γ_{sc}, such that the confinement and two-photon Doppler broadening reduce the reabsorption probability. This limit is known as the *festina lente* regime. The degradation of the performance at larger Γ_{sc} in Fig. 4-4(c) is consistent with increased rescattering, as the calculated reabsorption probability approaches unity 4.2. A too small value of Γ_{sc}, on the other hand, leads to higher temperatures as parasitic heating mechanisms cannot be compensated when the cooling is too slow.

While ^{87}Rb has relatively favorable collision properties (low two-body inelastic loss rate coefficient, and moderate three-body loss rate coefficient in the upper hyperfine manifold), these properties are not unique, and other atomic species may also be suitable for direct laser cooling to BEC. Since the cooling is not *deeply* subrecoil, relatively high densities $n \sim \lambda\!\!\!^{-3}$ are required for reaching BEC. Thanks to the fast cooling, the effect of inelastic loss is small enough if a cloud is stable at these densities (typ. lifetime $\gtrsim 1$ s at 10^{14} cm^{-3}). Inelastic processes can be further reduced in an effectively one-dimensional geometry [24], where fermionization of the bosonic wavefunction reduces collisional processes. The demonstrated technique could also be directly applied to fermionic atoms [94], as well as to laser cooled molecules [103, 104].

4.1.6 Conclusion

In conclusion, we have realized the decades-old goal of BEC purely by laser cooling by creating a single, moderately sized Bose-Einstein condensate in a standard crossed optical dipole trap. Notably, the method is consistent with the general theoretical

recipe put forward by [60]. Further work is needed to explore the limits of this new technique in terms of speed, condensate size and final temperature. [We observe that the performance is still density-limited, see Fig. 4-3(a), and verified that when allowing for moderate final evaporation, nearly pure condensates can be created within a cooling time of 1 s.] It may also be interesting to investigate if this technique can be used to experimentally realize an atom laser, where the condensate is created by bosonic stimulation into the atomic final state during the spontaneous photon scattering [105, 106], rather than through thermalizing elastic collisions.

We would like to thank Cheng Chin, Wolfgang Ketterle, and Martin Zwierlein for stimulating discussions and insightful comments. A. A. acknowledges support from the SNSF grant No. P2ELP2-181926. This work was supported by the NSF, NSF CUA, NASA, and MURI through ONR.

4.2 Supplement

4.2.1 Optical Pumping Detuning

To find an optical pumping detuning that avoids losses from molecular resonances as predicted in Refs. [27, 97], we start with a cold cloud of atoms and measure the loss spectrum below threshold (i.e. bare atomic resonance) for light-atom detunings Δ above -18 GHz. This is closer to resonance than previously explored experimentally on the D_1-line of ^{87}Rb [107, 108, 109]. After executing a sequence which scatters the same number of photons, about 100, regardless of detuning, we obtain a fraction of surviving atoms as given in Fig. 4-4(a) of the main text, with frequency regions where the remaining atom number is relatively large shown in Fig. 4-5. The scan is performed with a resolution of 10 MHz. The optimum around -4.33 GHz in Fig. 4-5(b) was chosen due to its good performance. Other optima could be used as well, and better optima may exist outside the range of this scan.

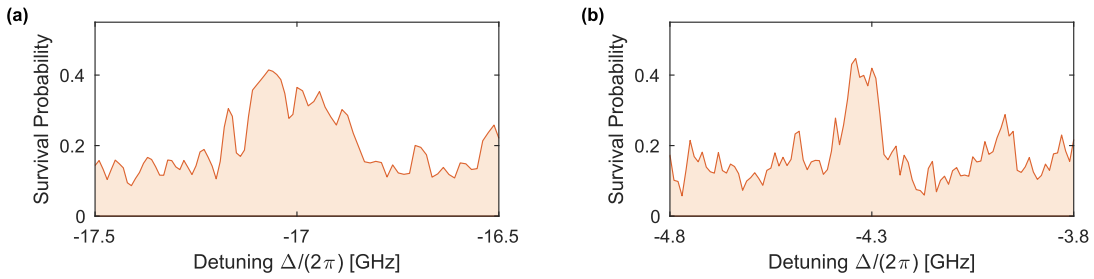


Figure 4-5: Survival probability of the atoms in the trap as a function of detuning after repeated cycles of optical pumping, corresponding to the scattering of about 100 photons, magnified on two exemplary detuning ranges. Fig. 4-5(b) is centered around the value used for most of the cooling data presented in the main text.

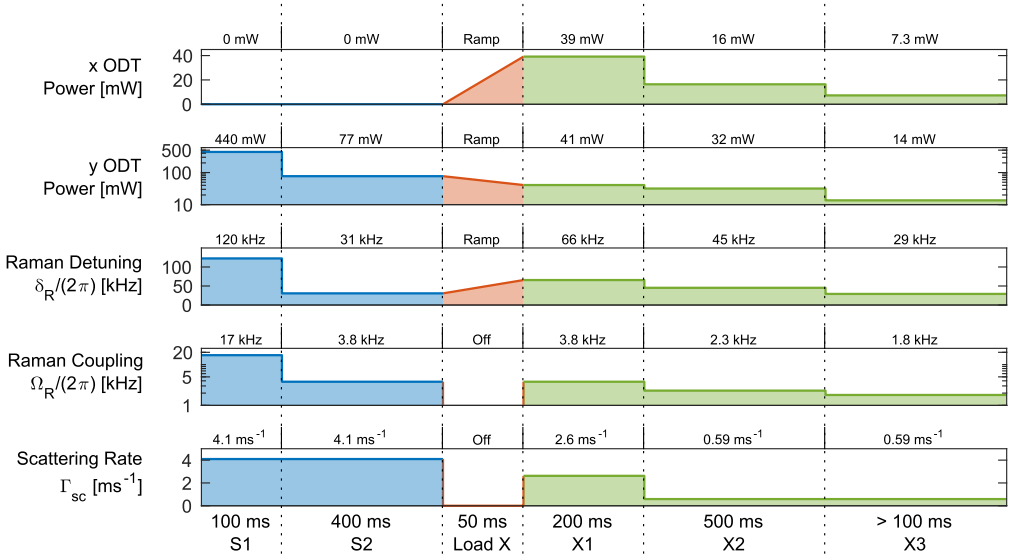


Figure 4-6: Optimized multistage Raman cooling sequence. Initially we cool in single beam ODT (blue), then ramp up a second trapping beam (red) to load into a crossed ODT for the final stages of cooling (green).

4.2.2 Experimental Details

The cloud is prepared in a similar way as in Ref. [24]. Rb atoms are loaded into a MOT from a thermal vapor, followed by a compressed MOT stage where the optical pumping power is strongly reduced, such that the atoms occupy the $F = 1$ ground state manifold. The ODT is turned on at all times and about $1{\times}10^5$ atoms are loaded into it when the MOT fields are switched off.

The ODT propagating in the y-direction (see Fig. 4-1 in the main text), in which the atoms are originally loaded, is focused to a waist of about 10 μm, while the second ODT propagating in the x-direction has a waist of about 18 μm at the position of the atoms. To avoid interference, the beams differ in frequency by 160 MHz. The powers of each trapping beam throughout the sequence are shown in Fig. 4-6. The (calculated) total trap depths, excluding the influence of gravity but including counter-rotating terms, are 430 μK in stage S1 and 14 μK in stage X3.

The imaging axis is the same as that of the π beam, which is slightly rotated by $\approx 17^\circ$ from the x-axis in the $x-y$ plane. In Fig. 4-3(b)-(c), we denote the horizontal axis of the image as y', which is rotated from the y-axis in the $x-y$ plane by the same angle.

The σ^- optical pumping beam at 795 nm has a highly elliptical shape at the position of the atoms, with waists of 30 μm along the x-direction and ~ 1 mm along the y-direction, to optimally address atoms along the sODT. The pumping beam creates a sizable light shift δ_{LS} of the $|2,-1\rangle$ state, given by $\delta_{\mathrm{LS}}/\Gamma_{\mathrm{sc}} = \Delta/\Gamma$, where Γ is the natural linewidth of the $5\mathrm{P}_{1/2}$ excited state. Since the state $|2,-2\rangle$ is dark for the σ^- light, there is no appreciable light shift on this state. The light shift is at its largest in stage S1, reaching $\delta_{\mathrm{LS}}/(2\pi) = 500$ kHz. It is determined experimentally by measuring the shift in the Raman resonance. From the measured values of δ_{LS} we deduce the scattering rates Γ_{sc} for $|2,-1\rangle$, shown in Fig. 4-6. Since the detuning Δ is comparable to the hyperfine splitting of the ground state, the same light also pumps atoms out of the $|5\mathrm{S}_{1/2}; F{=}1\rangle$ manifold.

The π Raman beam propagates in the $x-y$ plane (see Fig. 4-1 in the main text),

with a waist of 0.5 mm. Its light is derived from the same laser that generates the σ^- beam, but it is detuned by 2 MHz from the σ^- using acousto-optic modulators to avoid interference. This makes laser frequency noise common mode between the two beams, thereby loosening the requirements on laser linewidth necessary to have a narrow Raman transition. The Raman beam follows the same path as the light used for absorption imaging, and therefore it is circularly polarized. Over the course of the cooling only its π-polarized component (polarization along the z-axis, i.e. half its power), contributes to the Raman coupling. The other half of the power only adds a negligible amount of scattering and light shift to the $|2,-2\rangle$ state. Very little power ($\lesssim 100$ μW) is required to obtain significant Raman coupling (several kHz, see Fig. 4-6), so the light shift and scattering rate induced on the $|2,-2\rangle$ state are limited to $\lesssim 0.3$ kHz and $\lesssim 2$ s^{-1}, respectively, even at the largest powers used in the cooling sequence.

For all the data presented, each data point is evaluated as an average of 3 to 5 time-of-flight absorption images.

4.2.3 Effective Recoil Limit

At trap frequencies in the range 0.1−5 kHz, smaller than the Raman coupling, the cooling operates mostly in the free space limit with unresolved motional sidebands and outside the Lamb-Dicke regime [22, 42, 43]. Each Raman transition from $|2,-2\rangle$ to $|2,-1\rangle$ transfers $-\hbar\Delta\mathbf{k}$ of momentum to the atoms, where $\Delta\mathbf{k}$ is the difference between the wave vectors of the σ^- and π photons. The kinetic energy removed during one Raman transition for an atom of initial velocity $\mathbf{v}$ is:

$$\begin{aligned}\Delta K_{\text{Raman}} &= \hbar\Delta\mathbf{k}\cdot\mathbf{v} - \frac{\hbar^2|\Delta\mathbf{k}|^2}{2m} \\ &= \hbar\delta_{\mathbf{v}} - 2E_{\text{r}}\end{aligned} \tag{4.1}$$

where $\delta_{\mathbf{v}} = \Delta\mathbf{k}\cdot\mathbf{v}$ is the atom's two-photon Doppler shift. The orthogonality of the π and σ^- beams results in $|\Delta\mathbf{k}|^2 = 2|\lambda|^{-2}$, and $E_{\text{r}} = \hbar^2/(2m\lambda^2)$ is the recoil energy of a 795 nm photon. The branching ratio of the optical pumping to the dark state $|2,-2\rangle$

is 1/3, which results in the scattering of an average of 3 σ^- photons to optically pump the atoms following a Raman transition. In the limit of low scattering rate ($\Gamma_{\rm sc} \ll \omega_{x,y,z}$, which is required to limit reabsorption heating, see below), only the recoil energies of the scattered σ^- photons add, on top of those of the spontaneously emitted photons, resulting in an average net recoil heating of $6E_{\rm r}$ per optical pumping cycle. The average net energy removed per cooling cycle is then given by:

$$\begin{aligned} \Delta K_{\rm total} &= \Delta K_{\rm Raman} - 6E_{\rm r} \\ &= \hbar\delta_{\mathbf{v}} - 8E_{\rm r} \end{aligned} \tag{4.2}$$

Therefore there is net cooling if and only if the two-photon Doppler shift is larger than $8E_{\rm r}$, which translates to a kinetic energy in the $\Delta\mathbf{k}$ direction of:

$$K_{\Delta\mathbf{k}} \geq 8E_{\rm r} = h\times 29 \text{ kHz} \tag{4.3}$$

As a result, for temperatures below $T_{\rm r}^{\rm eff} = 2.8$ μK ($\langle K_{\Delta\mathbf{k}}\rangle = \frac{1}{2}k_{\rm B}T_{\rm r}^{\rm eff}$ along $\Delta\mathbf{k}$), the cooling speed drops since only a small fraction of the atoms have Doppler shifts above $8E_{\rm r}$ where they can be cooled.

4.2.4 Optimization of the Cooling and Limiting Factors

For each stage the Raman transitions are tuned to a particular velocity class by adjusting the magnetic field, which differentially shifts the $|2,-2\rangle$ and $|2,-1\rangle$ states. In stage S1, where the mean Doppler shift is much larger than the $8E_{\rm r}$ limit, we find as expected that the optimal Raman detuning is near the rms Doppler shift, see Fig. 4-7(a). This is a compromise between removing a large amount of energy per cooling cycle ($\Delta K_{\rm total} \approx \hbar\delta_{\rm R}$) and having a sufficiently large probability of finding atoms at those velocities (which drops as $\delta_{\rm R}$ increases). In stage X3 the optimization yields an optimal resonant Doppler shift of $\delta_{\rm R}/(2\pi) = 29$ kHz. As shown in Fig. 4-7(b), the probability of finding an atom at these velocities is very small, and the amount of energy removed $\Delta K_{\rm total}$ also becomes very small. By averaging $\Delta K_{\rm Raman}$ over the actual

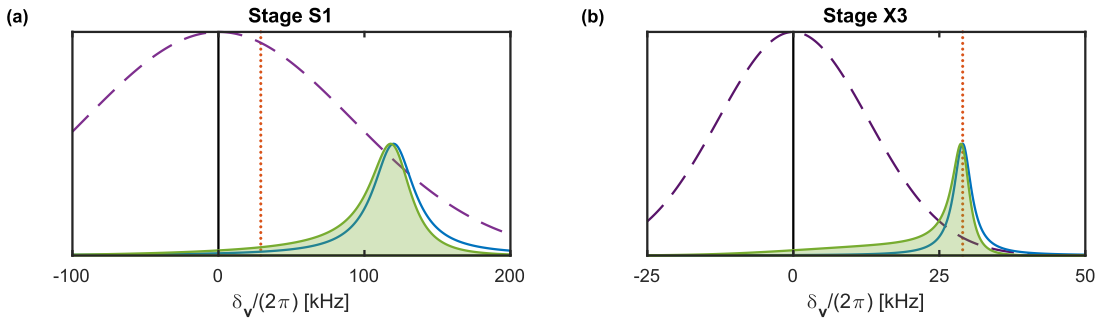


Figure 4-7: Thermal velocity distribution (dashed purple), time averaged Lorentzian Raman excitation profile expected given the Raman Rabi frequency (solid blue line), rescaled product of the thermal distribution and the excitation profile (shaded green), and effective recoil limit of 29 kHz (dotted red) for cooling sequence parameters in (a) stage S1 and (b) stage X3. On average, a cooling cycle will cool atoms with velocity greater than the effective recoil limit, and heat atoms with velocity below the effective recoil limit.

distribution of addressed atoms [green shaded in Fig. 4-7(b)] in stage X3, we obtain $\langle \Delta K_{\mathrm{Raman}} \rangle / h = 15$ kHz, well below the recoil heating of $6E_{\mathrm{r}}/h = 22$ kHz, which should lead to heating. Yet cooling is observed, which suggests bosonic enhancement of the branching ratio into the $|2, -2\rangle$ state due to the emerging condensate. A better branching ratio would reduce the average recoil heating during optical pumping, and so cooling could be achieved even when the Raman transition removes less than $6E_{\mathrm{r}}$ of kinetic energy.

At each stage, the strength of the Raman coupling Ω_{R} is optimized by scanning the power of the π beam. Too small Ω_{R} lead to a narrow excitation profile, and therefore only a small fraction of the atoms undergo a Raman transition. However, if Ω_{R} is too large, already "cold" atoms undergo a Raman transition, due to the broadened excitation profile, and are heated during the optical pumping.

Initially in the sODT it is favorable to have a fairly large scattering rate for cooling speed. The smaller number of atoms available to cool in the Boltzmann tail above the recoil energy in the later stages results in a lower optimal scattering rate as it becomes more favorable to decrease reabsorption heating as seen in Fig. 4-4(c) and discussed below.

For the ODT powers, the main considerations are finding a balance between low

density in order to limit inelastic loss and heating, and maintaining large enough trapping frequencies and therefore critical temperatures. Additionally in the final stage, having a low trap depth has proven crucial to avoid an observed density-dependent heating that increases with trap depth. As pointed out in Ref. [62], the products of three-body recombination of ^{87}Rb atoms in the $|2,-2\rangle$ state to the least-bound molecular state ($h\times$24 MHz of binding energy [110]) can collide with the cold sample before they leave the trap. This was shown to lead to large loss for a collisionally opaque ensemble. In our case, the sample is not collisionally opaque (collision probability ~ 0.1 for s-wave scattering only), nor is the trap deep enough to directly hold recombination products which would dump their energy into the cloud. The presence of a d-wave shape resonance at the energy of the least-bound state for a $^{87}Rb_2$ molecule is expected to enhance the collisional cross-section of recombination products with the trapped atoms, which could lead to strong losses in the collisionally thick regime. However, we mostly observe heating, which we suspect arises from recombination products undergoing grazing collisions with trapped atoms, with the latter remaining trapped and depositing heat into the cloud. We found that lowering the trap power as much as possible during the final stage X3 to minimize heating was necessary to reach condensation.

The optimized values of the relevant parameters throughout the sequence, namely trap power, Raman detuning δ_{R}, Raman coupling Ω_{R} and scattering rate Γ_{sc}, are shown in Fig. 4-6 for reference.

4.2.5 Estimation of the Reabsorption Probability

The reabsorption of a scattered photon causes excess recoil heating. This is especially of concern since the reabsorption cross-section, corresponding to a two-photon resonant process, can take on its maximum possible value $\sim 6\pi\lambda\!\!\!^{-2}$. Several strategies have been laid out for suppressing this [58, 59, 60], boiling down to the use of a low scattering rate Γ_{sc} for optical pumping, a regime known as *festina lente*. With our optimal parameters, we have $\Gamma_{\mathrm{sc}} \ll \omega_{\mathrm{D}}, \omega_{x,y,z}$, leading to a suppression of the reabsorption process on the order of $\Gamma_{\mathrm{sc}}/\omega_{\mathrm{D}}$, where ω_{D} represents the Doppler width. The

reabsorption cross-section σ_{reabs} is given by [59]:

$$\sigma_{\text{reabs}} \approx 4\pi\lambda^2 \frac{\sqrt{\pi}\Gamma_{\text{sc}}}{2\sqrt{2}\omega_{\text{D}}} \tag{4.4}$$

and the reabsorption probability p_{reabs} is given by:

$$p_{\text{reabs}} \approx \sigma_{\text{reabs}} \langle nl \rangle \tag{4.5}$$

where $\langle nl \rangle$ is the mean column density of the cloud. We use the following formula for the mean column density of a thermal cloud, taken from [62]:

$$\langle nl \rangle = \sqrt{\frac{\pi}{8}} n(0) \sigma_z \frac{\text{artanh}[\sqrt{1 - 1/\varepsilon^2}]}{\sqrt{\varepsilon^2 - 1}} \tag{4.6}$$

where $n(0)$ is the peak density, σ_z is the cloud's waist along the short axis of an elongated trap (z-axis here), and $\varepsilon = \omega_y/\omega_z$ is the aspect ratio of the trap. Assuming a classical cloud at the condensation point and a scattering rate $\Gamma_{\text{sc}} = 0.59$ ms^{-1}, we obtain that a probability for reabsorption $p_{\text{reabs}} \sim 0.1$. Hence reabsorption is not expected to be a significant factor, but it could explain why the performance deteriorates at larger Γ_{sc} in Fig. 4-4(c) of the main text, since p_{reabs} varies linearly with Γ_{sc}.

Chapter 5

Combining Raman Cooling with Evaporation

The following is the most recent draft of Ref. [40], which is currently still a work in progress. The draft is included here for completeness, but the user is encouraged to seek out the final version of the paper.

5.1 Main Text

5.1.1 Abstract

As repetition rate is a principal limitation on the performance of quantum gas experiments, techniques to enhance the speed of condensed gas preparation have broad utility. In this letter, we demonstrate the preparation of a Bose-Einstein condensate of 4.7×10^3 ^{87}Rb atoms in a crossed optical dipole trap in 575 ms, which is the fastest preparation of a degenerate gas to date known to the authors. Moreover, our results are achieved with a much simpler apparatus than other sub-second quantum gas preparation experiments. The preparation rate demonstrated here is enhanced over conventional approaches by Raman cooling to high phase space density, which provides faster and more efficient cooling than evaporation for clouds produced by a typical alkali magneto-optical trap. A final rapid evaporation phase produces pure

condensates from the cold and dense clouds prepared by the Raman cooling. Manually optimizing such a complex sequence involving tens of experimental parameters would take prohibitively long and furthermore is difficult for many standard numerical optimization algorithms. Here, it was accomplished using Bayesian optimization, in which a machine learning model is trained to predict system performance, making the high-dimensional optimization tractable and leading to high-performing sets of parameter values with minimal user intervention.

5.1.2 Introduction

In recent years, Bose-Einstein condensates (BECs) have enabled groundbreaking advances in various fields of research, from quantum simulation [9, 10, 13] to sensing and precision measurement [8, 11, 12]. With few exceptions [111], experiments on BECs use a shot-based approach in which data is collected by repeatedly preparing a BEC then destructively measuring it. With this approach the data collection rate is limited by the BEC preparation rate. Thus methods for faster BEC production [112, 29, 113, 28, 30] have broad applicability for accelerating research progress and reducing statistical noise.

Previously, fast BEC production schemes have generally relied on complicated apparatuses. For alkalis, the tight confinement of atom chip magnetic traps has enabled fast evaporation sequences with BEC preparation times as short as 850 ms for clouds with 4×10^4 atoms [29]. However atom chips require lithographic fabrication and restrict optical access. Experiments on species with narrow optical transitions [112, 30] have prepared BECs of 2×10^4 atoms in under 700 ms [30] using those transitions, but generally these approaches require overlapping multiple magneto-optical traps (MOTs) and stabilizing lasers to narrow linewidths. By contrast, the approach we present achieves its enhanced repetition rate using only a single vapor-loaded MOT, a crossed optical dipole trap (cODT), and two Raman cooling beams as depicted in Fig. 5-1(a). Furthermore, the approach works for species with broad (here ∼6 MHz) excited state linewidths. No Zeeman slower, 2D MOT, atom chip [29], or dynamic trap shaping [113] or strobing [114, 30] is necessary, and a laser with modest power

and linewidth can provide the light for Raman cooling. Despite the simplicity of the apparatus used, this approach yields BECs of 4.7×10^3 atoms in 575 ms, which is faster than any other method to date known to the authors. Furthermore, we demonstrate that we can optimize the many experimental control parameters using Bayesian optimization [31, 32, 33], a technique commonly used in machine learning, which significantly simplifies and speeds up the optimization procedure. Our results indicate that with modest effort new and existing experiments could implement this BEC preparation procedure and greatly enhance their data acquisition rates.

5.1.3 Methods

Raman cooling [22] provides sub-Doppler cooling by driving velocity-selective Raman transitions between hyperfine states. These states have negligible intrinsic linewidths, allowing for very narrow velocity class selection and thus circumventing the Doppler limit. The Raman transitions are non-dissipative, so optical pumping light is used to return atoms to their initial hyperfine dark state and entropy leaves the system in the form of spontaneously scattered photons. Light assisted collisions, which typically prohibit laser cooling at high phase space densities (PSDs), are suppressed by detuning the optical pumping light 4.33 GHz red of the D1 $F = 2 \rightarrow 2'$ transition [27, 28]. The detuning is large enough so that the same light used for optical pumping can be used to drive the down-leg of the Raman transitions as well, which simplifies the apparatus. Additionally, the light for the up-leg of the Raman transition is derived from the same laser as the optical pumping light. This ensures that the frequency noise is common mode between the two legs of the Raman transition and makes it possible to resolve Doppler shifts much smaller than the laser linewidth.

The Raman cooling dynamics are manipulated via five controls: (i) the horizontal P_y and (ii) vertical P_z trap beam powers which set the trap depth and frequencies, (iii) the pump beam power P_{p} which sets the optical pumping scattering rate, (iv) the Raman coupling beam power P_{R} which drives the up-leg of the Raman transition and sets the Raman coupling Rabi frequency for a given pump beam power, and (v) the magnetic field B_z which determines which velocity class is resonant for the Raman

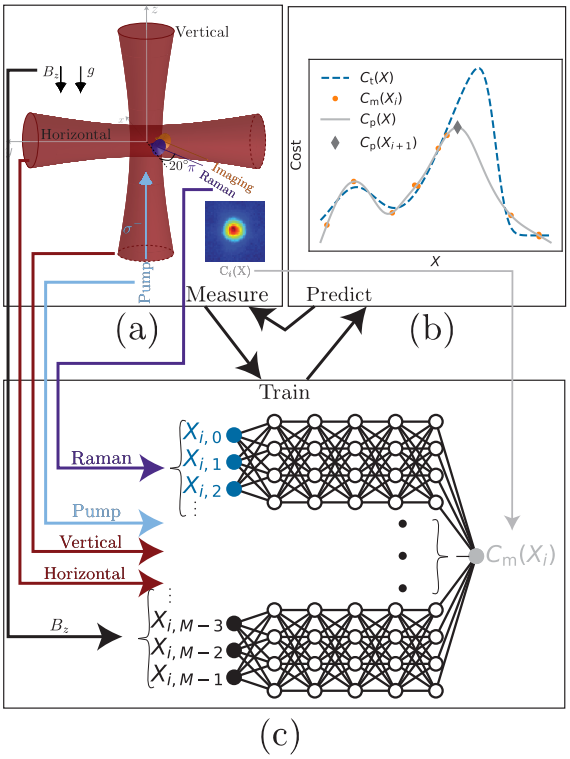


Figure 5-1: (a) Beam setup showing 1064 nm horizontal (18 μm waist) and vertical (14 μm waist) cODT beams, 795 nm Raman coupling (500 μm waist) and optical pumping (elliptical profile with 30 μm waist along x and $\sim$1 mm waist along y) beams, and 780 nm absorption imaging beam. (b-c) Bayesian optimization works by developing a model, here a neural network, to predict (gray solid line) system performance $C(X)$ (blue dashed line) for untested optimization parameter values X based on measured results for other values (orange circles), which can include noise. The algorithm uses the model to predict optimal parameter values (gray diamond), tests those values, then updates the model with the newly measured results before beginning a new iteration with the improved model.

transition. Similarly the evaporation dynamics are manipulated via two controls: P_y and P_z. The cooling and evaporation are divided into stages during which the controls are linearly ramped. We must optimize the endpoints of each ramp for each control, which constitutes a high-dimensional parameter space with tens of parameters.

The optimization problem can be posed as the maximization of a cost function C, which is a mapping from sets of parameter values $X \in \mathbb{R}^M$ to a corresponding cost value $C(X) \in \mathbb{R}$ where M is the number of optimization parameters. The cost is a number which quantifies the quality of the results produced by the set of parameter values and is generally unknown but may be experimentally measured. Bayesian optimization is well-suited for this type of problem as it can tolerate noise in the measured cost values and typically requires testing fewer values of X than other optimization methods [32, 33, 115, 116, 117, 118].

Bayesian optimization begins with collecting a training dataset by measuring $C(X)$ for various values of X, optionally including a user-supplied initial guess. The X are chosen by a training algorithm, which can implement another optimization algorithm or can select X randomly. A model of the cost function is then fit to the training dataset which approximates the unknown $C(X)$. Although Bayesian optimization typically uses a Gaussian process for its model [31], this work uses neural networks [33, 119], chosen for their significantly faster fitting time for our typical number of optimization parameters. Once the model is fit, a standard numerical optimization algorithm is applied to the modeled cost function to determine the X which is predicted to yield the best cost. Optionally this numerical optimization can be constrained to a trust region, which is a smaller volume of parameter space centered around the parameters which have yielded the best cost measured thus far. The actual cost is then measured for the predicted optimal value of X and the model is retrained with the result. Then a new iteration begins by predicting the optimal value for X with the updated model as depicted in Fig. 5-1(b). To encourage parameter-space exploration, additional iterations of the training algorithm can be periodically run and the results incorporated into model as well. The algorithm iterates either until it reaches a set maximum number of runs or until a set number of consecutive

iterations fail to return better results.

5.1.4 Optimization

We constructed the 575 ms cooling sequence as follows. First, we added 100 ms stages of Raman cooling one by one and optimized them individually using the open source packages M-LOOP [32, 33] to implement the Bayesian optimization and Labscript [85] for experimental control. In contrast to Ref [28], no restrictions are made to prohibit evaporation during the Raman cooling stages. After five stages, optimizations tended to turn off the Raman cooling by turning down P_{p} or P_{R} or tuning B_z such that the Raman transition resonant velocity was too large to have much effect. This likely occurred because the cloud temperature was below the effective recoil temperature [28] where Raman cooling, even with optimal parameters, becomes too slow to have much effect and leads to heating for non-optimal parameters. The Bayesian model likely recognized this and converged on parameters values far from those that lead to heating. At this point the cloud was close to condensation, so we then added several shorter 30 ms evaporation stages in which the optical pumping and Raman coupling beams were turned off. We then optimized the evaporation stages simultaneously due to their reduced number of parameters, which gave parameter values that produced a BEC. We then shortened the Raman cooling and evaporation stages with parameter values fixed until only a small and impure BEC was produced, and then reoptimized. All 42 of the Raman cooling and evaporation parameters were reoptimized simultaneously using the previous values as the user-supplied initial guess for X. We repeated this shortening and reoptimizing procedure until the optimization failed to find parameters that could produce sufficiently pure BECs.

The required beam powers generally varied over orders of magnitude so we optimized their controls on a log scale, while B_z was optimized on a linear scale. We included a feedforward adjustment in the B_z control values to account for the light shift ΔE_{p} of the $|2,-1\rangle$ state by the optical pumping beam. We ran five shots and averaged their results for each set of parameter values tested, which took $\sim$10 s accounting for experimental and analysis overhead. The analysis for each iteration

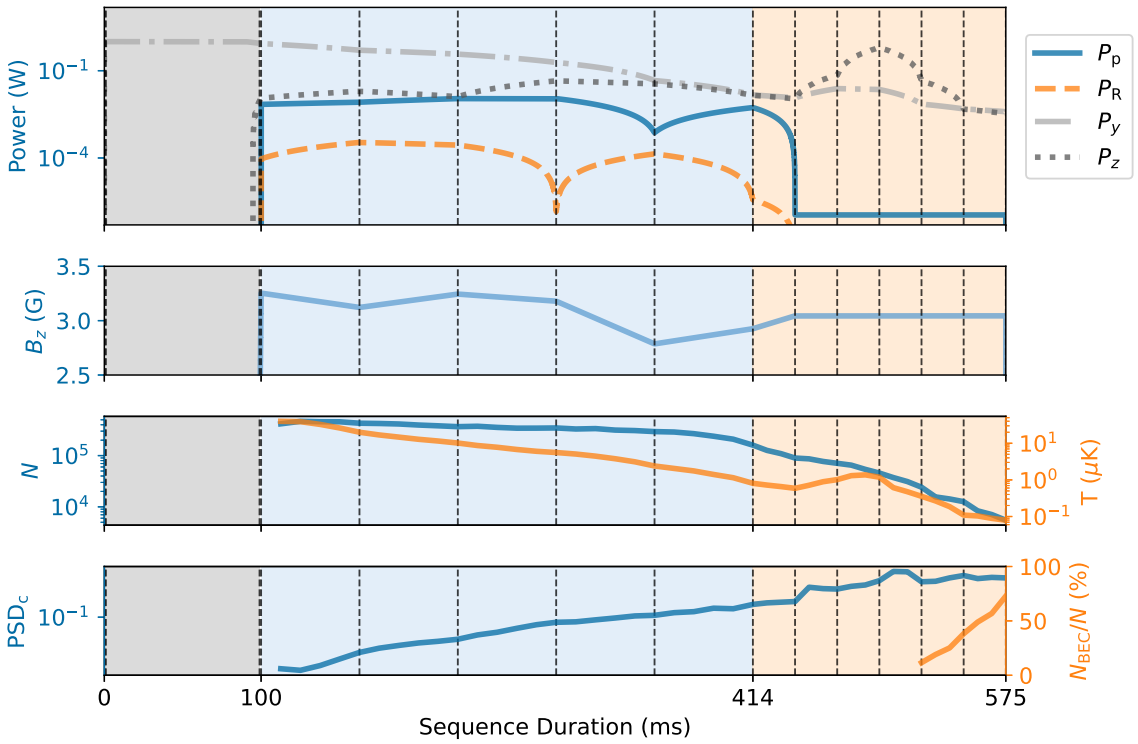
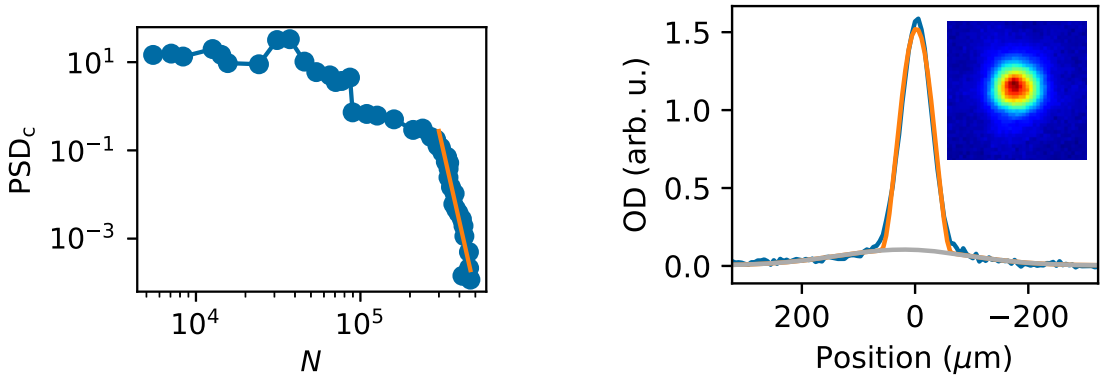


Figure 5-2: The control signals and measured atomic cloud properties of the optimized sequence after removal of the pre-release ramp to fixed trap powers, which is used only during the optimizations. Gray-shaded region marks the MOT loading period, blue marks the Raman cooling period, and orange marks the evaporation-only period.

typically took longer than the time required to perform the experiment, so two optimizations could be efficiently run in parallel on different computers. The number of iterations per optimization varied but typically involved ∼1000 iterations (including the initial training) and ran in several hours, both for the single-stage optimizations and the full-sequence optimizations. A less user-intensive optimization procedure was also attempted which did not involve optimizations of individual stages. Instead the sequence was divided into ten 100 ms stages and optimized from scratch simultaneously. That approach successfully produced a similar BEC, but could only be reduced to 650 ms by the shortening and reoptimizing procedure.



(a) (b)

Figure 5-3: The results of the 575 ms optimized sequence. (a) The atom number N, temperature T, and $\mathrm{PSD_c}$ as a function of time for the 575 ms optimized sequence. (b) Parametric plot of $\mathrm{PSD_c}$ as a function of N which shows the efficiency of the cooling. A fit to the early part of the sequence shows that the initial cooling from $\mathrm{PSD_c} \sim 10^{-4}$ to $\mathrm{PSD_c} \sim 10^{-1}$ occurs with $\gamma \approx 16$, much larger than achievable with evaporative cooling. (c) The horizontal cross section of the resulting atomic cloud after a 24 ms time of flight expansion with an inset showing the 2D absorption image of the cloud, averaged from 200 shots.

5.1.5 Results

The final 575 ms control sequence is depicted in Fig. 5-2, and the corresponding results are depicted in Fig. 5-2 and Fig. 5-3. The sequence begins with a 99 ms MOT loading and compression period which was optimized separately using M-LOOP. Notably, the compression sequence ends with gray molasses [120, 121] which was discovered by the optimization routine and outperforms the bright molasses [17, 18] that was previously used in the manually-optimized compression sequence. The trap beam powers are ramped to their initial Raman cooling values during the last 10 ms of the MOT compression and then the magnetic field is adjusted to its initial Raman cooling value in 1 ms, at which point the horizontal dipole trap holds approximately 5×10^5 atoms. The optical pumping and Raman coupling beams then turn on, initiating the Raman cooling. Five ∼63 ms stages of Raman cooling follow, and then the optical pumping and Raman beams are ramped off as a series of six ∼27 ms evaporation stages begin. As observed in previous work cite(Tranter2018,Nakamura2019,Barker2020), the ramps produced by Bayesian optimization are non-monotonic and unintuitive

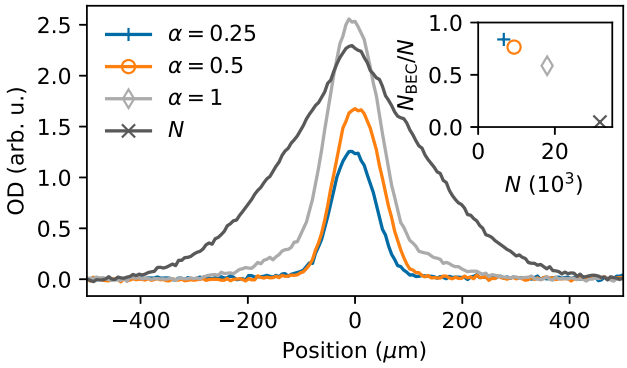


Figure 5-4: Cross sections of atomic clouds optimized for different α (see main text) with a 1 s sequence length, demonstrating the trade off between optimizing for atom number or temperature. Also plotted are the results of optimizing only for total atom number N. For fair comparison, images are taken after release from the fixed trap beam power settings used in the optimizations. The smallest of these clouds has 6.7×10^3 atoms, which is still larger than the value of N_{thresh}, which was set to 2×10^3 atoms. Each cross section is the average of 200 shots, taken after a 24 ms time of flight expansion. (inset) Condensate fraction N_{BEC}/N plotted against N for these cross sections.

which makes it unlikely that a person would discover such solutions in a manual optimization attempt, further indicating the usefulness of Bayesian optimization. The BEC is fully prepared at the end of the evaporation stages, 575 ms after the start of the MOT loading. The final cloud contains 6.2×10^3 total atoms with 4.7×10^3 atoms (76%) in the BEC and is shown in Fig. 5-3(b).

5.1.6 Cost Function

The properties of the optimized atomic cloud depend strongly on how costs are assigned to clouds during the optimization. Generally clouds with larger PSD have a larger atom number N and less expansion energy, which leads to a larger peak optical depth $\mathrm{OD_p}$ in absorption images for a given N. Guided by this, a cost function of the form $C(X) \propto f(N)\mathrm{OD_p^3}N^{\alpha-9/5}$ was used with

$$f(N) = \begin{cases} \left(\frac{2}{e^{N_{\mathrm{thresh}}/N}+1}\right) & N > 0 \\ 0 & N \leq 0 \end{cases} \tag{5.1}$$

where N_{thresh} is a threshold atom number and α is a parameter which tunes the trade-off between optimizing for larger atom number or lower temperature. This cost function uses only directly measurable quantities rather than results from fitting curves to data, which circumvents issues of convergence in fitting routines.

The value of $f(N)$ approaches unity for $N \gg N_{\text{thresh}}$ and exponentially approaches zero for $N \ll N_{\text{thresh}}$. Setting N_{thresh} approximately equal to the noise floor of the N measurement prevents the cost from diverging as N approaches zero and has little effect on the cost for measurable atom numbers. It can also be set to a larger value to target a minimum atom number. Increasing the value of α puts more emphasis on increasing the atom number, while lowering α puts more emphasis on lowering the temperature. For a harmonically-trapped pure BEC after sufficient time of flight expansion, $\text{OD}_{\text{p}}^3 N^{\alpha-9/5}$ scales as $(N_{\text{BEC}})^{\alpha}$ where N_{BEC} is the number of atoms in the BEC. To compare clouds from different sequences fairly, the trap beams are ramped to a fixed power setting before releasing the cloud for imaging. The pre-release trap powers are set to have minimal evaporation and the ramp to those powers can be omitted from the sequence after the optimization is complete. To reduce noise, OD_{p} is set to the average OD of several pixels with the largest OD.

The atomic clouds produced by sequences optimized for different values of α are presented in Fig. 5-4, as well as the results when optimizing for total atom number (N). These optimizations were done with a 1 s sequence duration to ensure ample cooling time and N_{thresh} was set to 2×10^3. Larger values of α lead to clouds with more atoms, but with higher temperatures and lower condensate fractions. Smaller values of α produce purer BECs, but with fewer atoms overall. Setting α to 0.5 was found to make a reasonable compromise so that value was used for the final full-sequence optimization.

5.1.7 Conclusion

In conclusion, we have demonstrated that Raman cooling with far detuned optical pumping light combined with a final evaporation can rapidly produce BECs with a comparatively simple apparatus and without narrow optical transitions. While some

knowledge of atomic physics was used to effectively parametrize the different cooling regimes, Bayesian optimization greatly eased the search for a short sequence to a high purity BEC, quickly discovering high-performing unintuitive sequences that are unlikely to be discovered manually. The algorithm's choice of optimal parameters can be interpreted as identifying physically recognizable tradeoffs between different cooling priorities and can be tuned by the user through the cost function construction. In future applications, user intervention may be further reduced by including the sequence duration as an additional optimization parameter and factoring the sequence length into the assigned cost. The methods demonstrated in this work are generalizable to other species with hyperfine structure, particularly alkalis, and can likely be added to new or existing experiments without extensive effort. Other established techniques, such as Zeeman slowing or dynamic trap shaping, can also be incorporated, enabling further improvements in cloud size and repetition rate at the cost of increased experimental complexity.

The authors would like to thank Martin Zwierlein for loan of a 1064 nm fiber amplifier. They would also like to thank Michael Hush, Harry Slatyer, Philip Starkey, Christopher Billington, and Russell Anderson for stimulating discussions and software assistance. This work was supported by the NSF, NSF CUA, NASA, and MURI through ONR.

5.2 Supplement

5.2.1 Cost Scaling

The peak optical depth $\mathrm{OD_p}$ of a pure BEC after sufficient time of flight expansion scales as $\mathrm{OD_p} \propto N_{\mathrm{BEC}}/A$ where A is the area of the cloud in the image. The area scales in proportion to $\bar{v}^2$ where $\bar{v}$ is the expansion velocity, which is related to the BEC chemical potential via $(1/2)m\bar{v}^2 = (2/7)\mu$ in a harmonic trap [99]. Thus $A \propto \mu$. Furthermore, the chemical potential for a harmonically-trapped BEC scales as $\mu \propto N_{\mathrm{BEC}}^{2/5}$ [99], so $A \propto N_{\mathrm{BEC}}^{2/5}$ and $\mathrm{OD_p} \propto N_{\mathrm{BEC}}^{3/5}$. The expression $\mathrm{OD_p^3} N^{\alpha-9/5}$ then scales

as $(N_{\mathrm{BEC}})^{\alpha}$. Notably this scaling also applies to a harmonically-trapped BEC when imaged in situ. There, the BEC radius R scales as $R \propto N_{\mathrm{BEC}}^{1/5}$ [99]. In that case, $A \propto R^2 \propto N_{\mathrm{BEC}}^{2/5}$ as before. The same arguments then apply again, indicating that $\mathrm{OD}_{\mathrm{p}}^3 N^{\alpha-9/5}$ scales as $(N_{\mathrm{BEC}})^{\alpha}$ for the in situ BEC just as it does for the BEC after long time of flight expansion.

The scaling of $\mathrm{OD}_{\mathrm{p}}^3 N^{\alpha-9/5}$ for a purely thermal cloud is also of note. For a thermal cloud, the RMS size σ for any time of flight is given by

$$\begin{aligned} \sigma &= \sqrt{x^2 + (vt)^2} \\ &= \sqrt{\frac{k_{\mathrm{B}}T}{m\omega_x^2} + \frac{k_{\mathrm{B}}T}{m}t^2} \\ &= \sqrt{\frac{k_{\mathrm{B}}T}{m}}\sqrt{\frac{1}{\omega_x^2} + t^2} \end{aligned} \tag{5.2}$$

Therefore $\sigma \propto T^{1/2}$ and so $A \propto T$. Thus $\mathrm{OD}_{\mathrm{p}} \propto N/T$ and $\mathrm{OD}_{\mathrm{p}}^3 N^{\alpha-9/5}$ scales in proportion to $N^{\alpha+(6/5)}/T^3$. Clouds with smaller temperatures are favored by the cost function, and clouds with larger atom numbers are favored as long as $\alpha > -6/5$. For the case $\alpha = -1/5$, the value of $\mathrm{OD}_{\mathrm{p}}^3 N^{\alpha-9/5}$ scales in proportion to N/T^3, which is proportional to $\mathrm{PSD}_{\mathrm{c}}$. That choice of α was often used when optimizing individual stages before reaching the threshold to BEC. However note that this choice of α leads to the scaling $\mathrm{OD}_{\mathrm{p}}^3 N^{\alpha-9/5} \propto N_{\mathrm{BEC}}^{-1/5}$ for a pure BEC and is thus not a good choice when the cloud reaches condensation.

Chapter 6

1D and 2D Lattice Raman Sideband Cooling

In addition to the work on free-space Raman cooling in optical dipole traps, research was performed on Raman sideband cooling in optical lattices as well. The work along these lines began as an extension to previous work in our lab on Raman cooling in a 2D lattice [24]. A major goal was to increase the number of atoms in the resulting quantum degenerate clouds. The first efforts along these lines simply involved loading more atoms into the 2D lattice before cooling. This approach yielded unusual and interesting results which are detailed in Section 6.1, but it did not lead to larger quantum degenerate gases. Instead larger, but warmer, clouds were produced. These which had the strange feature that their temperatures along the loosely-confined and tightly-confined directions of the trap potential were significantly different.

After the 2D lattice results, efforts were made to perform Raman sideband cooling in a 1D lattice, which had a few potential advantages. Firstly, the dynamics of ultracold atoms in a 1D lattice are very different than those for ultracold atoms in a 2D lattice. Thus switching to a 1D lattice could possibly circumvent the temperature issues observed with the 2D lattice. Furthermore, the volume of our 1D lattice is significantly larger than that of our 2D lattice. That larger volume makes it possible to load significantly more atoms into the lattice from the MOT, thus opening the possibility for larger degenerate clouds. Raman cooling in the 1D lattice did make

it possible to produce larger near-degenerate clouds, reaching a classical phase space density of about $\sim$0.1 with $\sim 10^5$ atoms. However, we were unable to cross the threshold to degeneracy, in part due to the slow thermalization between the loose and tight directions of the trap. These results are presented in Section 6.2. It was only after this work that our focus shifted towards free-space Raman cooling in a crossed optical dipole trap, as presented in Ch. 4 and Ch. 5, which proved to be more fruitful.

Before proceeding, one potential point of confusion is worth clarifying. When Raman sideband cooling in an optical lattice, it is often possible to cool the tightly-confined directions to their ground state. When that occurs, that direction is frozen out of the dynamics and the atoms behave as a lower-dimensional gas. For example, when atoms in a 2D lattice are cooled to the ground state in the tightly-confined directions, their only remaining motional degree of freedom is in the remaining loosely-confined direction. Thus the gas behaves as if it were in 1D. Similarly a gas cooled to the ground state of the tightly-confined direction in a 1D lattice then behaves as if it were in 2D. This is because there are two remaining degrees of freedom, one along each of the loosely-confining directions. Thus cooling in a 2D lattice can yield a 1D gas, and cooling in a 1D lattice can yield a 2D gas. Note that dimensionality of the gas is not the same as the dimensionality of the lattice.

6.1 2D Lattice Raman Sideband Cooling

Our lab showed in Ref. [24] that degenerate Raman sideband cooling in a 2D lattice could be used to produce quantum degenerate gases of ^{87}Rb. In that work, atoms were cooled to the ground state of the two tightly-confined directions of lattice sites, freezing out those degrees of freedom. This causes the gas behave as a 1D system along the loosely-confined direction, leading somewhat unusual behavior. In particular, bosonic atoms can "fermionize", avoiding each other due to their interactions [122, 123, 124, 125, 126, 127]. This has been shown to reduce loss processes that require multiple atoms, such as light-assisted collisions and 3-body loss, which makes it possible to

laser cooling atoms to quantum degeneracy [24, 25]. However, the number of atoms in the cloud was fairly small, about 1400 atoms in total across all of the lattice sites with a peak single lattice site occupation of about 50 atoms. Given that success, a straightforward next step was to try the same procedure but with more atoms. After some upgrades, we attempted this and found some interesting results. Firstly, atom loss was observed, which was not clearly seen in our previous work with fewer atoms. [24]. Additionally, the atoms could not be cooled to the ground state along the tightly confined directions of the trap, even though they are cooled directly. More interestingly, it was found that the temperature along the loose direction of the trap was no longer well-thermalized with the tight direction after cooling. This section details those experimental results. The reader is encouraged to read Ref. [24] before reading this chapter as this chapter is a fairly direct continuation of the results in that paper.

6.1.1 Improved Lattice Loading

Because few atoms were lost during the cooling in Ref. [24], efforts to cool more atoms required increasing the initial atom number by loading more atoms into the lattice from the MOT. That became possible after replacing the nearly-empty rubidium dispensers with a vapor cell rubidium source, as discussed in Section 3.5.3. After that upgrade the rubidium vapor pressure in the chamber increased significantly, greatly increasing the MOT loading rate. The majority of the work presented in this chapter was performed before the control system upgrades discussed in Section 3.4 and thus did not benefit from the enhanced lattice loading achieved with the Bayesian-optimized MOT compression sequence. However, the initial atom number was still about an order of magnitude larger than that in Ref. [24] due to the vapor cell source upgrade alone.

6.1.2 Experiment Sequence

The beam configuration and cooling sequence used in this experiment were similar to that used in Ref. [24]. Atoms were first loaded into a MOT overlapped with the same 2D lattice, then the MOT was extinguished. This loaded atoms along the lengths of the lattices in both the x and y directions, including the wings where the beams were not overlapped. To remove atoms from the wings, the lattice in one direction was turned off then on again, then the same was done for the other beam. This ensured that any atoms which were not trapped by both beams were removed because those atoms would escape when their confining beam was extinguished. This process also provides some release-and-retrap compression of the cloud [26, 24, 25] which increases the number of atoms in lattice sites near the center of the trap. However, in contrast to the sequences in Refs. [26, 24, 25], no cooling was performed before or during the compression. After the retrapping, the cloud was cooled via Raman sideband cooling in the 2D lattice.

6.1.3 Results

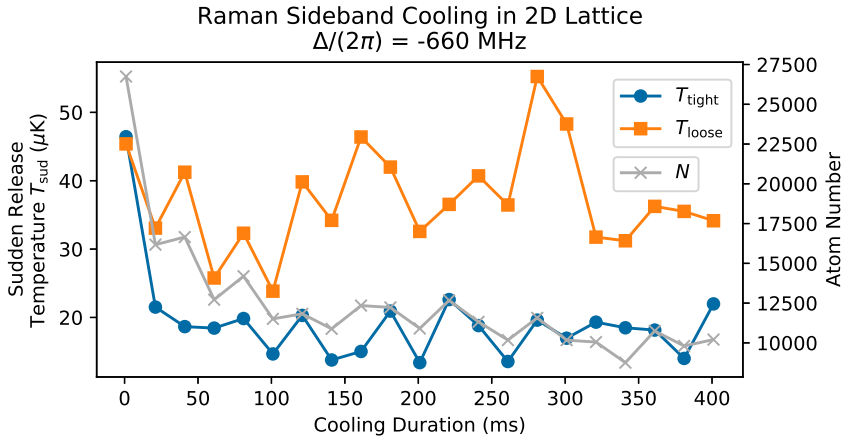
In the 2D lattice, two of the directions are tightly-confined and can be directly cooled by the Raman sideband cooling. The third direction is loosely-confined and has a much smaller trap frequency. That direction is only cooled via thermalization with the tight directions. In other words, atoms with kinetic energy along the loose direction can collide and subsequently reduce their kinetic energy along the loose direction while simultaneously moving one or both of the atoms to more highly-excited harmonic oscillator states in the tight direction. The excitations in the tight direction can then be removed via the Raman sideband cooling.

Cooling of the loose direction via thermalization with the tight directions can be seen to work reasonably well in Fig. 3(a) of Ref. [24]. There the tight direction cools to near its ground state so that its velocity in that direction during its time of flight is dominated by the kinetic energy of the ground state. The temperature of the loose direction follows reasonably well, cooling over time via thermalization with the tight

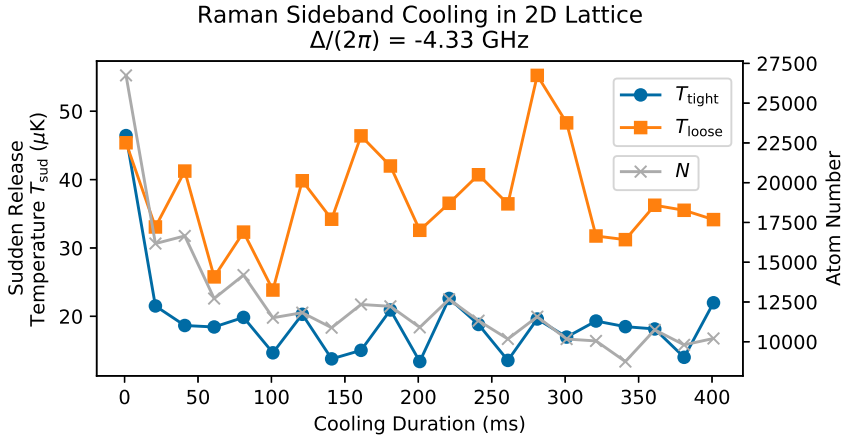
direction. Interestingly the thermalization did not work as well after the atom number was increased. This can be seen in Fig. 6-1 which shows the results after increasing the atom number. In particular, Fig. 6-1 shows the temperatures in the tight and loose directions, as well as the atom number, as the cooling proceeds. The temperatures are measured with sudden release, i.e. the trap light is suddenly extinguished before the cloud expands for its time of flight (see Appendix A). Fig. 6-1(a) shows how the cooling performs when using similar optical pumping detunings as Ref. [24], namely $\Delta/(2\pi) = -660$ MHz with a 6.8 GHz sideband to provide repumping for atoms that decay to the $F = 1$ hyperfine manifold. Fig. 6-1(b) shows a similar curve, but with $\Delta/(2\pi)$ set to -4.33 GHz, which is the value used for most of the work presented in this thesis. No repumping sideband is needed when using a -4.33 GHz detuning as then the carrier scatters photons off of atoms in $F = 1$ and $F = 2$ at similar rates. The optical pumping rates differ somewhat between the two curves in Fig. 6-1, which is why in initial cooling rates are different. However, the trap configurations and Raman coupling Rabi frequencies are the same for both curves.

The exact details of the curves in Fig. 6-1 differ slightly from each other, but they are remarkably similar to each other and very different from those observed in Ref. [24]. Firstly, there is clear atom loss in the curves of Fig. 6-1, which was not observed in Ref. [24]. Here the atom number drops from 2.5×10^4 to 1×10^4 over the course of the cooling while it didn't decrease to any measurable degree in the final cooling stage of Ref. [24]. Furthermore, the tight directions are not cooled to their ground state. The tight trap frequencies are ≈150 kHz and ≈190 kHz in the x and y directions respectively [1] which would give a sudden release temperature of ≈4 μK when the atoms are cooled to their ground state. However, the final sudden release temperatures in the tight direction are ≈10 μK. Additionally, the temperatures of the two directions do not thermalize as well in Fig. 6-1 as they did in Ref. [24]. The data is somewhat noisy but the final loose direction temperature is clearly significantly larger than that of the tight direction. In Fig. 6-1(a) it is not clear that the loose direction

[1] Measurements were also performed with the trap beam powers adjusted to better match the trap frequencies between the two directions. However, similar results were still observed; the loose direction still remained much hotter than the tight direction.



(a)



(b)

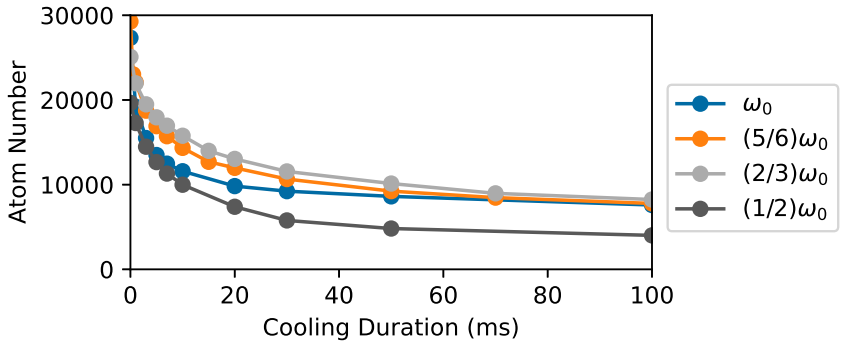
Figure 6-1: Curves of temperature and atom number during Raman sideband cooling in a 2D lattice. The sequences are essentially the same in (a) and (b) except that the optical pumping beam detuning $\Delta/(2\pi)$ is -660 MHz in (a) while it is -4.33 GHz in (b). The optical pumping scattering rates also differ somewhat and the data in (a) requires a repumping sideband to empty the $F = 1$ manifold, but the trap configurations and Raman coupling are the same for both sets of curves. Similar behavior is observed at both detunings. The atom number drops from 2.5×10^4 to 1×10^4 over the course of the cooling. The temperature data is somewhat noisy, but the temperature in the tight direction T_{tight} and the temperature in the loose direction T_{loose} clearly settle to different values. The loose direction remains significantly warmer than the tight direction. Here the temperatures are measured with sudden release (see Appendix A).

cools at all even though the tight direction temperature clearly drops significantly. A careful temperature measurement for the final temperatures in Fig. 6-1(b), which accounted for the initial cloud size before time of flight expansion, indicated that the tight direction sudden release temperature was 11.5 μK and the loose direction temperature was 22.0 μK. Thus the loose direction was roughly twice as warm as the tight direction.

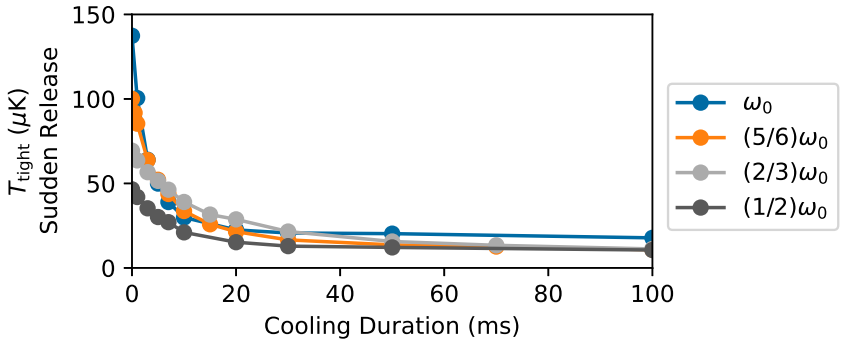
Interestingly the degree to which the temperatures in the tight and loose directions differ depends on the actual value of the trap frequencies, not just the ratios between them. Fig. 6-2 shows cooling curves taken with various trap frequencies. The trap frequencies were varied by attenuating the trap beams, so the tight and loose direction trap frequencies were attenuated by the same factor. Reducing the trap frequencies has a noticeable, but comparatively small, effect on the resulting curves for atom number and sudden release temperature in the tight direction. However, the temperature in the loose direction varies significantly as the trap frequencies are changed. With the smallest trap frequencies (half the maximum values), the loose direction's temperature is about equal to the sudden release temperature of the tight direction. However, with the largest trap frequencies, the loose direction's temperature is about four times larger than the tight direction's sudden release temperature.

One approach for working around the lack of thermalization during Raman sideband cooling in a 2D lattice is to cool the loose direction directly. An experiment was performed along these lines in which, rather than cooling the tight directions with Raman sideband cooling, free-space Raman cooling was performed along the loose direction of the tubes. To achieve this, the Raman coupling was performed using the optical pumping beam and 795 nm Raman coupling beam [2], as was done in the work presented in Ch. 4 and Ch. 5. Notably the two 795 nm beams driving the Raman transition differed in frequency by 2 MHz and the magnetic field was

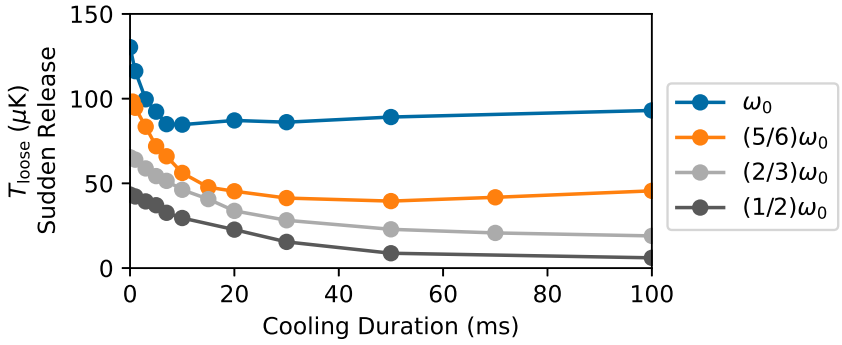
[2]The momentum kick imparted during a Raman transition driven by those beams did have a component in the tight trap directions. Therefore those beams could have driven Raman sideband cooling along the tight directions of the trap simultaneously with the free-space Raman cooling along the loose direction. However, there were no clear sideband resonances when adjusting the Raman transition detuning (by changing the magnetic field), which indicates that any Raman sideband transitions in the tight direction did not significantly affect the cooling.



(a)



(b)



(c)

Figure 6-2: Curves of temperatures and atom number during Raman sideband cooling in a 2D lattice with varied trap frequencies. The trap frequency is varied by attenuating the trap beams. The curves of atom number and tight direction sudden release temperature look fairly similar for the various configurations. However, the loose direction temperature varies significantly between the configurations, becoming much larger when the trap frequency is increased. With the highest trap frequency setting, the final loose direction temperature is roughly four times the tight direction sudden release temperature.

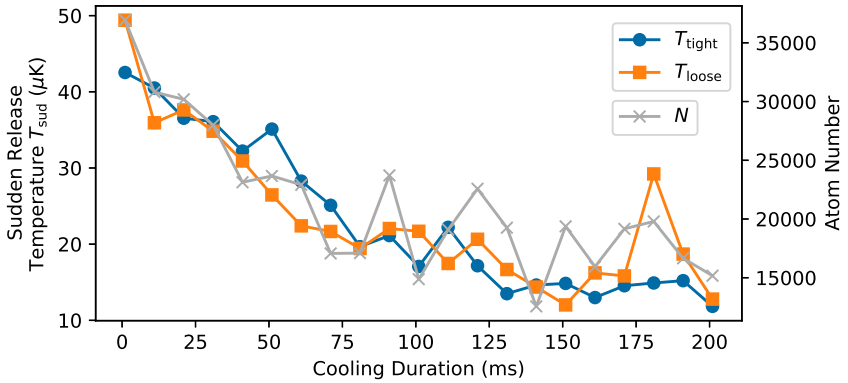


Figure 6-3: Curves of temperatures and atom number during Raman cooling (i.e. with unresolved sidebands) along the loose direction of a 2D lattice. This is similar to the results presented in Fig. 6-1, except that the loose direction is cooled directly and the tight directions are primarily cooled via thermalization with the loose direction. In this configuration the temperatures of the two directions remain well thermalized during the cooling. However, the tight direction is still not cooled to its ground state. The optical pumping beam detuning was $\Delta/(2\pi) = -4.33$ GHz for this measurement.

tuned to compensate for that with the Zeeman shift. Although the trap beams still had polarizations which could drive Raman transitions, they were degenerate and so Raman transitions driven by those beams were detuned by ≈2 MHz from the two-photon resonance. Therefore the trap beams did not drive any Raman coupling or cooling in the tight direction for this measurement and the Raman transitions were instead driven by the 795 nm beams alone. The results when cooling with this configuration are presented in Fig. 6-3. Interestingly the atomic loss is roughly similar to that in Fig. 6-1, but the temperatures in the loose and tight directions remain well-thermalized throughout the cooling. The final temperatures for both directions in Fig. 6-3 are similar to that of the tight direction in Fig. 6-1. Unfortunately, the final temperature in the tight direction is still much larger than its ground state energy. An attempt to add an additional stage of Raman cooling along the loose direction of the 2D lattice did not yield colder temperatures. Therefore, this approach resolved the issue of the lack of thermalization, but it still could not cool the cloud to its tight direction ground state.

6.1.4 Outlook

The cause of the lack of thermalization between the tight and loose directions when Raman sideband cooling in the 2D lattice is not clear. Generally the theory behind the dynamics of 1D Bosonic gases is fairly complicated and can be algebraically intractable. The analysis likely becomes even more difficult when the energies of the system are comparable to the tight direction trap frequency, as then a few higher-energy trap states in the tight direction begin to play a role. Thus it is difficult to determine if the lack of thermalization is due to interesting quantum effects, or if it is simply due to technical issues. More work is required to illuminate the cause of this unusual behavior.

Possible future work could also include performing Raman sideband cooling along the tight directions and free-space Raman cooling along the loose direction simultaneously. That may make it possible to achieve lower temperatures. It may also be beneficial to increase the total atom number by increasing the number of occupied lattice sites (e.g. by making the lattice larger), rather than by increasing the number of atoms per lattice site. Doing so may make it possible to achieve results comparable to those in Ref. [24], but with more atoms in total.

6.2 1D Lattice Raman Sideband Cooling

After it was found that increasing the initial atom number in our 2D lattice did not yield a larger degenerate gas, our focus turned towards Raman sideband cooling in a 1D lattice. As mentioned at the top of this chapter, this yields a 2D gas when the cloud is cooled to the ground state in the tight direction. The 2D Bosonic gas has different physics than a 1D Bosonic gas, and thus the dynamics are different. The new dynamics are interesting in their own right and also had the possibility of working around the temperature issues observed in the 2D lattice. Furthermore, many more atoms could be loaded into the 1D lattice from the MOT due to its much larger volume compared to the 2D lattice. Thus we began investigating Raman sideband cooling in a 1D lattice.

Raman sideband cooling in a 1D lattice had been performed by other groups in the past [128, 34, 35] and more recently [129]. However it had not been done with far-detuned optical pumping light to suppress light-assisted collisions, which were seen to limit the cooling performance in Ref. [35]. Thus we expected that our apparatus, with its far-detuned optical pumping light, could potentially provide improved performance over other experiments.

The apparatus used for the 1D lattice Raman sideband cooling experiments is described in Ch. 3. In particular, the two 1064 nm trap beams propagating along the y-direction were used to form the 1D lattice. Using two separately-controlled beams to generate the lattice made it possible to adjust the radial and axial trap frequencies somewhat independently by adjusting the power in only one beam, which would change the aspect ratio of the trap. For some measurements, the trap beams along the other directions were used to provide additional confinement at certain points in the sequence.

6.2.1 Results

Our most basic sequence for Raman sideband cooling in a 1D lattice is very straightforward. Atoms are captured by the MOT which is overlapped with the 1D lattice. The MOT is then extinguished and the magnetic field is adjusted to the correct magnitude and orientation for the Raman sideband cooling. The optical pumping light is then turned on to initiate the cooling which proceeds for a duration of time. The atoms are subsequently released for a time-of-flight expansion and imaged with absorption imaging.

The evolution of the atom number and loose direction (vertical) temperature during a typical cooling sequence are shown in Fig. 6-4. The cloud extends along the lattice direction, so its size along that (horizontal) tight-confinement direction is dominated by its initial size even after the time-of-flight expansion. Thus the tight direction temperature could not be directly measured with this sequence. A precise temperature measurement was performed after 200 ms of cooling, which measured the loose direction temperature to be 3.47 μK. Because the loose direction only

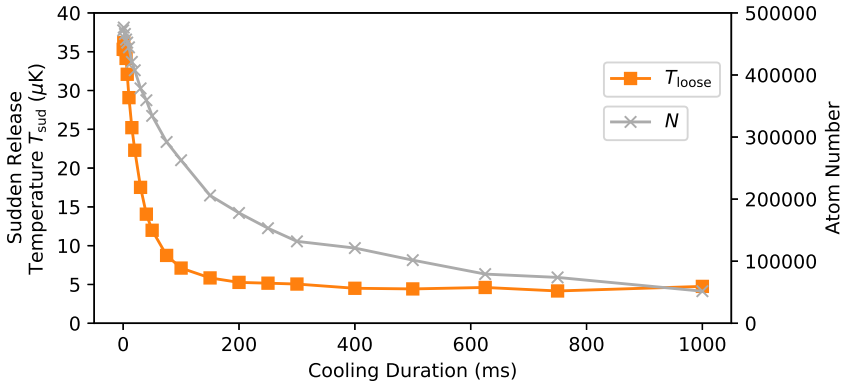


Figure 6-4: A typical curve of the atom number and the temperature in the loosely-confined direction during Raman cooling in a 1D lattice. Atoms are lost and the temperature drops as the cooling proceeds. The optical pumping light was detuned $\Delta/(2\pi) = -660$ MHz and a repumping sideband was used for this measurement.

cools via thermalization with the tight direction, the tight direction should be at the same temperature or even colder. The lattice's tight direction trap frequency was $\omega_{\text{tight}} \approx (2\pi)130$ kHz. With that trap frequency, a 3.47 μK temperature would correspond to $\approx$83% population in the ground state (see Eqn. A.2). Given that the tight direction temperature may be lower, this is a lower bound on the ground state population. Thus the Raman sideband cooling achieved good ground state cooling in the tight direction.

The classical phase space density PSD_c at the 200 ms of cooling mark in Fig. 6-4 can be estimated as follows. At that point in the cooling, there were 1.75×10^5 atoms remaining spread across $\approx$1300 lattice sites, giving $\approx$140 atoms per lattice site [3]. Given that the tight direction is cooled to near its ground state, the PSD_c is given by the formula for a 2D gas:

$$\text{PSD}_\text{c} = N \left(\frac{\hbar \omega_{\text{loose}}}{k_\text{B} T_{\text{loose}}} \right)^2 \tag{6.1}$$

where T_{loose} and ω_{loose} are the temperature and trap frequency in the loose direction,

[3] Due to the repulsive radiation trapping force in the MOT, which leads to a fairly uniform density in the MOT, the distribution of atoms across the lattice sites was fairly uniform. That distribution becomes even more uniform after cooling because sites with more atoms undergo more multi-body loss, bringing their population closer to that of the other lattice sites.

and here N is the number of atoms in a given lattice site (*not* the total number of atoms across all lattice sites). The loose trap frequency was not measured directly, but was calculated to be $\omega_{\text{loose}} \approx (2\pi)2$ kHz for the trap configuration used for Fig. 6-4. Namely this was two 1.8 W 1064 nm counterpropagating beams, with 30 μm waists [4]. Furthermore, the beams were linearly polarized with one polarization rotated 70° relative to the other. Inserting these values in Eqn. 6.1 indicates that $\text{PSD}_c \sim 0.1$ at the 200 ms mark in Fig. 6-4. That is about an order of magnitude shy of the value $\text{PSD}_c \approx 1$ required for quantum degeneracy. However it is still significantly larger than the $\text{PSD}_c = 1.3 \times 10^{-3}$ achieved in Ref. [35] and the $\text{PSD}_c = 1/180$ achieved in Ref. [34].

Cooling the loose direction further would lead to a higher PSD_c, however there is an interesting difficulty with that. The atoms only cool in the loose direction via thermalization with the tight direction, but the thermalization slows down at low temperatures. The final temperature is achieved when the cooling rate is balanced by any heating processes, so the final temperature becomes larger if the cooling via thermalization is slowed. That slowing has a simple explanation. For two atoms to collide and transfer energy from the loose direction to the tight direction, they must have kinetic energy comparable to the tight direction trap frequency $\text{KE} \gtrsim \hbar\omega_{\text{tight}}$. Otherwise the atoms do not have enough energy to move to an excited harmonic oscillator state in the tight direction, as those states are separated in energy by the trap frequency. For identical Bosons, the situation is actually even worse. Symmetry requirements make it impossible to excite just one atom by one level in the tight direction. Therefore two levels of tight direction excitation are required, which increases the initial kinetic energy that the atoms need [130, 35]. The odds that the atoms have sufficient kinetic energy are exponentially suppressed at low temperatures by a factor $p = \exp\left[-2\hbar\omega_{\text{tight}}/(k_B T_{\text{loose}})\right]$, and thus the thermalization rate is suppressed [5] by the same factor [35]. That reduced thermalization rate slows

[4] Unfortunately this data was taken not long after I began running the machine and before we developed ways to systematically document our research. One of the trap beam waists may have actually been 18 μm, which would imply that the PSD_c is actually slightly larger than calculated.

[5] Note that the lack of thermalization seen in Section 6.1.3 often occurred even when $k_B T_{\text{loose}}$ was large compared to $\hbar\omega_{\text{tight}}$. Thus a different mechanism must be behind the lack of thermalization

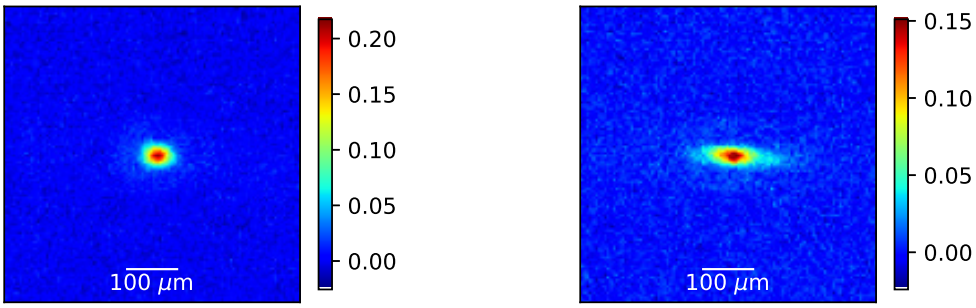
the cooling of the loose directions, which leads to larger temperatures. For the results quoted above, namely $T_{\text{loose}} = 3.47\ \mu$K and $\omega_{\text{tight}} \approx (2\pi)130$ kHz, this factor is 2.7×10^{-3}. Thus the cooling performance and achievable PSD_c is significantly affected by the slowed thermalization at low temperatures.

Two directions were pursued for increasing the PSD_c. The first strategy for increasing the PSD_c was to reduce ω_{tight} to increase the thermalization rate with the loose direction. This approach was successfully used in Ref. [35] where the authors adjusted the interference angle between their lattice beams mid-sequence to tune their tight direction trap frequency. More recently this was also done by phase modulating the lattice [129]. Our approach was to simply reduce the power in one of the beams forming the lattice. This also reduced the value of ω_{loose} somewhat, but the thermalization still improved because the change in ω_{tight} was larger than the change in ω_{loose} [6]. In one measurement, 30 ms of cooling was performed in the same configuration used for the results in Fig. 6-4. Then one of the lattice beam powers was reduced by a factor of 16, which reduced ω_{tight} by a factor of 2. Finally, 300 ms of Raman sideband cooling in the weakened lattice was performed. This produced a cloud with $T_{\text{loose}} \approx 2.2\ \mu$K and ≈200 atoms per lattice site, giving about twice the PSD_c achieved in Fig. 6-4. Thus some improvement was seen, but not enough to reach degeneracy. Additional improvement is likely possible with other configurations, e.g. different trap powers or interference angles, as the parameter space was not fully explored.

The other strategy used to increase PSD_c was to simply increase the number of atoms per lattice site, which increases the critical temperature and the PSD_c at a given temperature. This was done with a release-and-retrap compression [26, 24, 25] using the trap beam along the x-direction to compress the central region of the 1D lattice. This retrapping procedure unfortunately only compressed the central region of the cloud and allowed most of the atoms (the ones in the rest of the 1D lattice) to escape. However, it did increase the number of atoms in each occupied lattice site

seen there.

[6] In the extreme limit that one beam was turned off completely, ω_{tight} would go to zero while ω_{loose} would only drop by a factor of $\sqrt{2}$ to 2 (depending on the angle between the polarizations of the beams). The beam power was never actually turned down to that extent, but this demonstrates that ω_{tight} could be tuned relative to ω_{loose} by adjusting the power in one beam.



(a) (b)

Figure 6-5: Images of clouds after undergoing Raman sideband cooling in a 1D lattice, taken with a short time-of-flight expansion. Atoms were only loaded into the central region of the 1D lattice; lattice sites further along the length of the 1D lattice were intentionally left empty. In (a) a small cloud in the central region of the lattice is visible. In (b) the cloud extends further along the lattice because atoms have spilled out of the central lattice sites and into other nearby lattice sites during the Raman sideband cooling.

even though the total number of atoms was significantly reduced.

An unexpected behavior was observed when cooling after compressing the cloud: the atoms would spill out of their lattice sites. This can be seen in the images in Fig. 6-5. In that data, a small cloud was prepared in the central region of the 1D lattice by first loading atoms into a separate orthogonal 1D lattice, then transferring them to the 1D lattice used for the cooling. This produced a small cloud of atoms which did not extend far along the 1D lattice (which extends horizontally in the images), as can be seen in Fig. 6-5(a). However, when as the cloud was cooled, atoms tended to spill out of the central lattice sites and end up spread out further along the 1D lattice, as can be seen in Fig. 6-5(b). Strangely this spilling only occurs when the Raman sideband cooling is applied; if the atoms are simply held in the lattice without cooling, they do not spill.

The spilling seems to be worse when the lattice interference angle is large or when ω_{tight} is reduced by reducing the power in one of the lattice beams. This makes sense because in order to spill, atoms must end up with an amount of energy that is greater than the depth of a lattice site but less than the total trap depth [7]. That way the

[7]Note that because the polarization of the two lattice beams are different (which is necessary to

atoms can move between lattice sites, but do not escape the trap completely. The range of energies which satisfy that criteria is larger when the interference angle between the trap beams is larger or when the power in one of the trap beams is reduced. Unfortunately this makes it difficult to combine both techniques for increasing the $\mathrm{PSD_c}$. If the cloud is compressed to increase lattice site occupation and ω_{tight} is reduced to improve thermalization, then spilling occurs which reduces the lattice site occupation. The spilling likely also occurs when the cloud isn't compressed and instead extends along the length of the 1D lattice beams. However, in that situation most occupied lattice sites are near other comparably-occupied lattice sites. Then the spilling out of one of these sites is roughly cancelled out by the spilling into that site from other nearby sites. Thus spilling is less of an issue for long/uncompressed clouds.

A few efforts were made to reduce the effects of spilling for the compressed clouds. One was to perform the Raman sideband cooling in stages. Initially ω_{tight} was set to a large value to reduce spilling, then once the cloud was cooled sufficiently, ω_{tight} was reduced to improve thermalization. This staged approach reduced the degree of spilling. Another technique was to provide additional confinement in the direction of the 1D lattice, such as by turning on a dipole trap beam orthogonal to the 1D lattice. This prevented atoms from spilling far along the 1D lattice. It is likely that atoms still spilled between lattice sites within the dipole trap's confinement, though this would have not been observable with our imaging resolution. Interestingly that may provide further compression of the cloud, though again it would be nontrivial to detect with our system.

The mechanisms behind the spilling are not fully understood. One possibility is that although atoms are cooled on average during a Raman cooling cycle, they are occasionally heated. For example, on occasion it will take more scatterings than usual before an atom is successfully optically pumped to its initial ground state. If the heating during this process is unusually large, then the atom can gain enough

drive the Raman transitions), they do not fully interfere. Thus the lattice depth, i.e. the amplitude of the potential energy spatial modulations with period $\lambda/2$, is less than the full trap depth.

energy to escape its lattice site before being cooled and trapped into another lattice site. It is also possible that other dynamics are to blame for the spilling.

6.2.2 Outlook

Future work on Raman sideband cooling in a 1D lattice could involve a more systematic optimization over the parameter space to improve the $\mathrm{PSD_c}$ that can be achieved. For example, improvement may be possible via further reduction of ω_{tight}, possibly after previous stages of cooling with larger ω_{tight} to avoid spilling and to remove more energy with fewer scatterings. However, if ω_{tight} is reduced too far, the cooling will leave the Lamb-Dicke regime which may be detrimental (see Section 2.7.3). Adjustments of the interference angle between the beams may also yield better results. Use of an automated optimization procedure, such as Bayesian optimization implemented by M-LOOP (see Ch. 5 and Refs. [31, 32, 33]), would make the task of optimization more tractable. Adjusting the beam waists of the trap beams may also yield further improvements. For example, decreasing the beam waist increases the ratio $\omega_{\mathrm{loose}}/\omega_{\mathrm{tight}}$ which should help thermalization between the tight and loose directions. Alternatively, the issue of thermalization between the tight and loose directions of the trap may be circumvented by (free-space) Raman cooling the loose direction directly, as was done in the 2D lattice in Section 6.1.3. Lastly it may also be interesting to combine Raman sideband cooling in a 1D lattice with evaporation to bridge the gap to degeneracy. That approach would take advantage of the fast and efficient cooling provided by Raman sideband cooling in a 1D lattice with the ability of evaporation to make pure degenerate gases, akin to the work presented in Ch. 5.

Chapter 7

3D Lattice Raman Sideband Cooling

In addition to the Raman cooling work performed in the free space limit, some research on Raman sideband cooling in a 3D lattice was performed. This had a few possible benefits and goals, discussed below. The work is still ongoing as of this writing, but preliminary results are presented. Section 7.1 discusses the final temperatures that were achieved with this technique, measuring the fraction of atoms in the ground state of their lattice site. It is also shown that increased laser power is expected to yield better performance. Attempts to cool multiply-occupied lattice sites without loss, presented in Section 7.2, were not promising. As an alternative approach to increase the final lattice site occupation after cooling, efforts to increase the fraction of singly-occupied lattices sites are presented in Section 7.3. A well-prepared and well-cooled cloud could be released from the 3D lattice into a crossed optical dipole trap (cODT) to create a BEC. That could make it possible to use Raman sideband cooling in a 3D lattice to produce BECs and it could be used as a diagnostic to measure the performance of the Raman sideband cooling. Efforts along these lines are presented in Section 7.4.

The apparatus used for this work is described in Ch. 3. The beam waists of the various trap beams used in this work are presented in Table 7.1. See Section 3.1.3 for a description of the configuration and orientation of the trap beams and see Section 3.2.3 for a description of how the 1064 nm light used for the trap beams is generated. The optical pumping light was detuned by $\Delta = -(2\pi)4.33$ GHz, which is

Trap Beam	Beam Waist
X Westward	18 μm
X Westward Retroreflection	18 μm
Y Northward	30 μm
Y Southward	18 μm
Z Downward	14 μm

Table 7.1: Beam waists of the various 1064 nm trapping beams used in our 3D lattice Raman cooling work.

the same detuning used for the work presented in Ch. 4 and Ch. 5.

7.1 Ground State Cooling

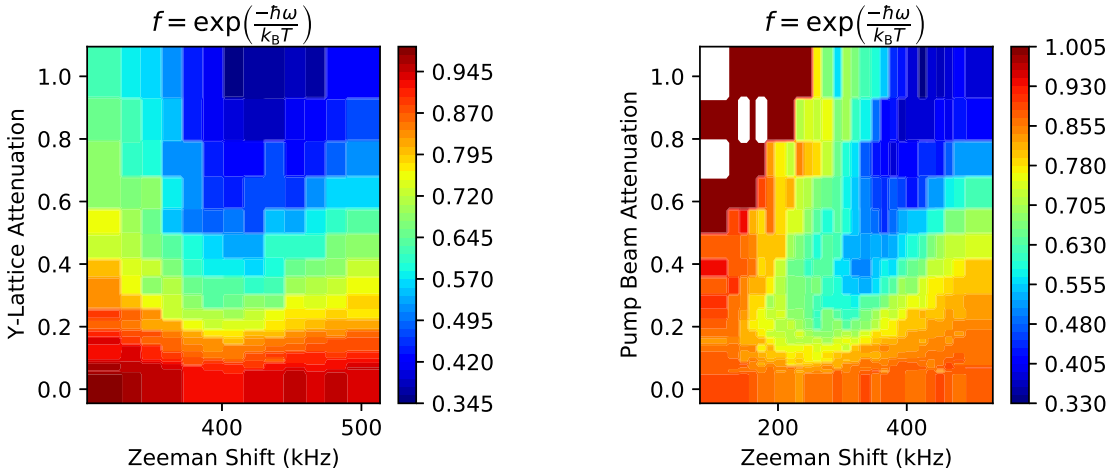
One possible benefit of 3D sideband cooling is that it could make it possible to cool atoms to the ground state of their trap without requiring any thermalizing collisions. This is possible because Raman sideband cooling is capable of ground state cooling and because each direction can be cooled by Raman sideband cooling directly in a 3D lattice. That could potentially allow for much faster cooling than achievable with the free space Raman cooling work presented in Ch. 4 and Ch. 5 as the cooling there is slowed by the finite thermalization rate, particularly when the temperature drops below the effective recoil temperature. Raman sideband cooling in a 3D lattice could also be faster than Raman sideband cooling in lower-dimensional lattices as thermalization is required to cool the weakly confined directions in those lattices.

The basic goal of laser cooling is to make the atomic cloud as cold as possible. Therefore the parameters of the system were tuned to minimize the temperature and maximize the ground state population. The tunable parameters include the powers of each of the trap beams, the power in the optical pumping beam, the magnetic field (which brings the Raman sideband transition into and out of resonance), and the polarization angles of three of the trap beams (Y Northward, the X beam, and the X beam retroreflection). The trap beam powers and polarization angles tune both the trap frequencies and the Raman coupling Rabi frequencies for the various directions. The magnetic field must be optimized to bring the Raman sideband

transition into resonance. Although Bayesian optimization could be applied to the cooling optimization, we have not yet used it for that purpose.

The temperature of the clouds produced via Raman sideband cooling was generally measured in one of two different ways: sudden release or adiabatic release. In each approach the cloud's kinetic energy during a time of flight expansion is measured then a temperature is assigned based on that kinetic energy. In sudden release, the trap is quickly extinguished and atoms retain the kinetic that they had in the trap, including the ground state's kinetic energy. In adiabatic release, the lattice is adiabatically turned off, which removes much of the kinetic energy that atoms had in the trap before they are released to expand freely. In either case a "temperature" T_{sud} or T_{ad} (for sudden or adiabatic release respectively) is then assigned based on the kinetic energy after release KE to be $2\mathrm{KE}/k_{\mathrm{B}}$. Notably, neither of these "temperatures" are exactly equal to the true thermodynamic temperature T of the atoms in the lattice, which is the value that appears for the temperature in the Boltzmann factor. However, the true thermodynamic temperature can be calculated from either T_{sud} or T_{ad} using the relations discussed in Appendix A. Sudden release has the advantage that it is simple to implement, but has the disadvantage that it must be measured very accurately in order to calculate the true thermodynamic temperature when the temperature is small compared to the trap frequency. Adiabatic release is more sensitive in that regime, but is more difficult to implement precisely. For these reasons, we generally used adiabatic release to compare the performance between different sequences, and it is particularly well-suited for fairly comparing results between sequences with different trap frequencies. However, because of the difficulty of implementing adiabatic release precisely, we often used sudden release when precise quantitative results were needed. See Appendix A for more details about these temperature measurement methods.

Because the temperature could differ along different trap eigenaxes, it was measured separately for the two directions which could be imaged. The imaging axis was in the horizontal plane, rotated about 17° from the x-axis. Therefore the temperature was measured in a horizontal direction T_{h} and the vertical direction T_{v}. The temperature along the imaging axis could not be measured directly.



(a) (b)

Figure 7-1: Cooling performance for varied beam powers after 1 second of 3D Raman sideband cooling. To fairly compare results from trap configurations with different trap frequencies, the value of f is plotted rather than T_{sud}. (a) Results as the powers of the trap beams in the y-direction are attenuated to various degrees. The Zeeman shift is also scanned to account for the change in trap frequency. The best performance (lowest f) occurs with the largest achievable beam powers. (b) Similar to (a) except that the optical pumping beam's power is varied rather than the power of any trap beams. White regions represent points where the atomic distribution fitting routines did not converge. The optimal Zeeman shift moves towards larger values for larger pump powers in part due to the light shift of the $|F = 2, m_f = -1\rangle$ state by the optical pumping beam. As with the trap beams, optimal performance is achieved with the maximum power setting.

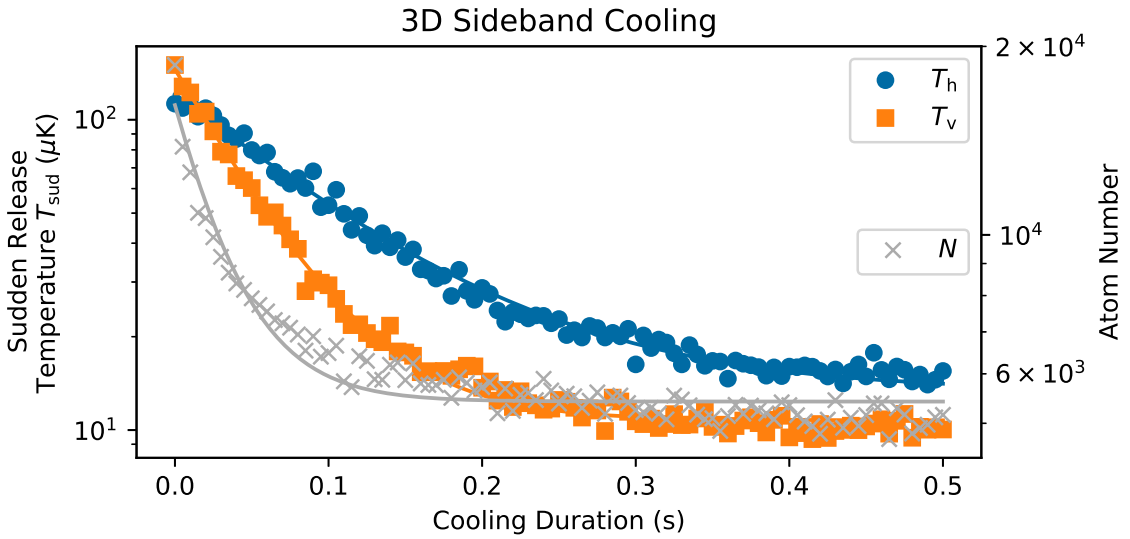


Figure 7-2: Typical curves of temperature and atom number N during Raman sideband cooling in a 3D lattice plotted with exponential fits. In this measurement, the horizontal temperature T_h and vertical temperature T_v start at different values and cool with different exponential timescales (102 ms and 49 ms respectively). This occurs because the cooling in each direction happens independently and some of the cooling parameters, such as the trap frequencies and Raman coupling Rabi frequency, are different for different directions. Here the "temperatures" are measured with sudden release, i.e. the lattice is quickly extinguished when the atoms are released for time of flight expansion.

When optimizing the various parameters of the sequence, it was generally found that performance improved with larger powers for the optical pumping beam and for the trap beams. This is shown for the optical pumping beam and for the lattice in the y-direction in Fig. 7-1. Thus all of the beams were set to their maximum available power, and performance would likely improve if larger powers were achievable. A typical measurement of the atom number and temperatures over the course of the cooling is presented in Fig. 7-2. Atom loss occurs at a rate comparable to the cooling rate. Although the horizontal and vertical temperature curves are well-fit by a single exponential, the atom number curve is not as well-fit by a single exponential. This may indicate that there are multiple loss processes with different timescales playing a role. For example, there may be a loss rate of light-assisted collisions in initially multiply-occupied lattice sites and a separate loss rate for unbound atoms binding to previously singly-occupied sites then undergoing a light assisted collision. These loss processes were seen to occur with different timescales in Ref [26].

The ground state population after cooling was also measured. To make this measurement, the cloud was cooled for 1 s to ensure that the minimal temperatures were reached. Then the lattice was suddenly extinguished and the temperature was measured by fitting the atomic distribution with a Gaussian at varied time of flight, then inferring T_{sud} from the rate of expansion of the cloud. The measured temperatures were 9.0 μK in the horizontal direction and 7.3 μK in the vertical direction. The lattice frequency in the y-direction was separately characterized [1] to be to be 240 kHz for the 45° angle between the polarization of the y-direction lattice beams used in this measurement. Thus the value of f in the y-direction was calculated to be 0.21 and so the fraction of atoms in the ground state in that direction is 79%. The population in the 3D ground state is lower due to possible excitations in the other directions. Calculating the ground state population for those directions would require careful

[1] The trap frequency was measured by varying the Zeeman shift to see various Raman sideband resonances, both on the heating and cooling sides. The trap frequency was then determined from the spacing between the resonances. This measurement was performed with a small interference angle between the lattice beams. That reduced the Raman coupling Rabi frequency and yielded narrower resonances which provided more precise results. The trap frequency then scales as $\sqrt{\cos\theta}$ where θ is the angle between the lattice beam polarizations.

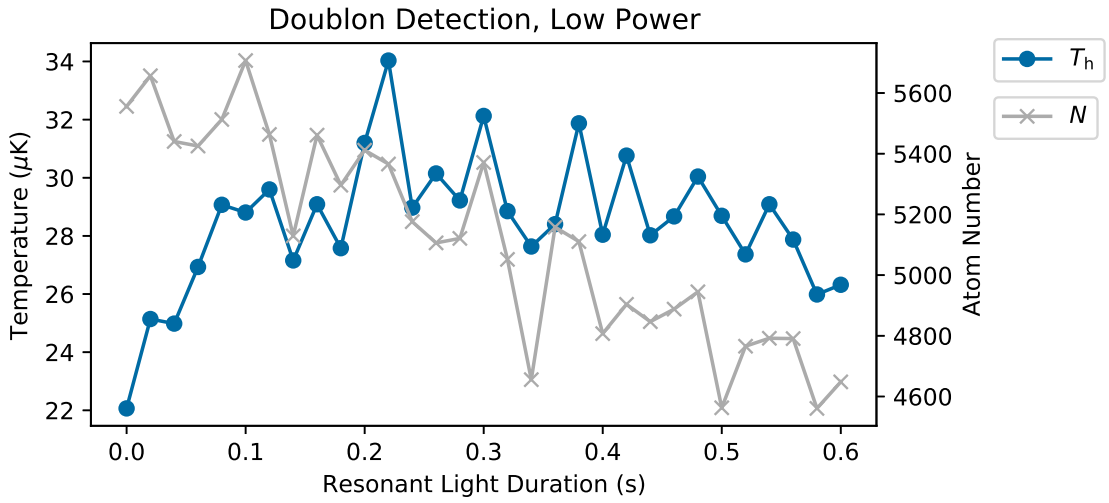
calibration of the trap frequencies in those directions.

7.2 Attempts to Cool Multiply Occupied Sites

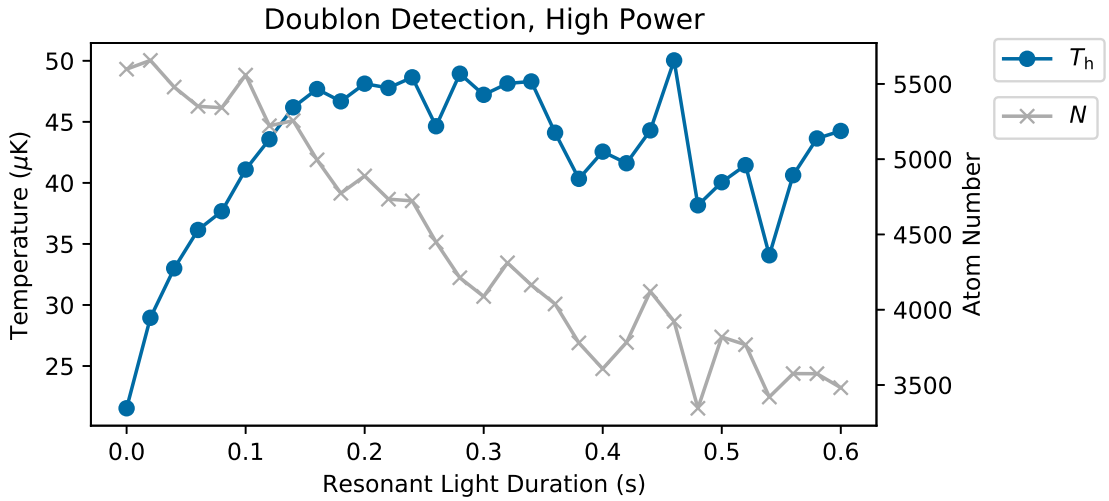
Another goal for 3D Raman sideband cooling would be to cool multiple atoms in the same lattice site. This could make it easier to produce a BEC when releasing the 3D lattice to a cODT, as a larger temperature in the 3D lattice could then be tolerated and more atoms would remain. Cooling three atoms in the same lattice site is unlikely to be practical due to 3-body loss. This is the case because 3-body loss occurs at a rate that scales strongly with density, and the density in a 3D lattice is typically quite large compared to that of a bulk gas. However, 3-body loss cannot occur in lattice sites with only two atoms, thus it may be possible to cool two atoms in the same lattice site. Preliminary work performed previously in our lab seemed to indicate that this may be possible [79].

Although atom loss is clearly visible in Fig. 7-2, it is possible that some multiply-occupied lattice sites survive the cooling. One straightforward way to detect multiply-occupied sites is to illuminate the cloud with resonant light and check for excess loss. If the cooling does not eject atoms from multiply-occupied sites, then the loss should suddenly increase with the introduction of resonant light as that light will then induce light-assisted collisions. Results of such a measurement are presented in Fig. 7-3. In that dataset the 3D Raman sideband cooling duration is fixed at 1 s and the total sequence length is held constant to avoid changes in vacuum loss. Resonant light is shone on the atoms for the last section of the cooling, but the duration of the resonant light is varied. For example, the 0.1 s data points in Fig. 7-3 correspond to a sequences in which atoms undergo 3D Raman sideband cooling for 1 s and resonant light is shone on the atoms for 0.1 s starting 0.9 s after the start of the cooling.

If the 3D Raman sideband cooling could successfully cool two atoms in the same lattice site, then there should be a sudden drop in atom number when the resonant light is turned on. That is not observed in Fig. 7-3, which may indicate that there are no doubly-occupied sites after the cooling. The loss observed occurs slowly and speeds



(a)



(b)

Figure 7-3: Temperature and atom number after sideband cooling with resonant light shone on the atoms for a varied amount of time. The cooling duration is fixed at 1 s and the resonant light is shone on the atoms for the last section of the cooling. The intensity of the resonant light is smaller in (a) and larger in (b), but the intensity was not measured for either dataset. The data is somewhat noisy but the atom loss is slower than the heating. In particular in (b) the loss appears to be slower initially then speed up after the cloud heats. This suggests that the loss is associated with heating the atoms rather than being due to light-assisted collisions in multiply-occupied lattice sites which survived the cooling. This indicates that the 3D Raman sideband cooling likely does not successfully cool any measurable fraction multiply-occupied sites. However, these results are preliminary and more checks are necessary. In particular, due to technical issues even the sequence with nominally 0 s of resonant light does actually shine some resonant light on the atoms due to the finite AOM extinction ratio.

after the cloud is heated, indicating that the loss is associated with recoils heating the atoms out of their lattice sites rather than being due to atoms in multiply-occupied sites undergoing light-assisted collisions. However, the data in Fig. 7-3 is preliminary and more checks are necessary. In particular, for technical reasons the shutter blocking the resonant light was opened even during the sequence with nominally 0 s of resonant light. The AOM power control was set to zero during that time, but some resonant light was still sent to the atoms due to some offsets in the system and the AOM's finite extinction ratio. Therefore it is possible that doubly-occupied sites were present in the cloud but were removed by the resonant light even for the data points with nominally 0 s of resonant light. This is somewhat plausible given that, as shown in Ref. [26], only a few resonant photon scatterings are necessary to drive a light-assisted collision in a 3D lattice. A more careful check in which a shutter is used to provide complete extinction is necessary to conclusively determine whether or not doubly-occupied sites are present after 3D Raman sideband cooling.

Although the existence of doubly-occupied sites after cooling is not fully excluded by the experimental measurements made thus far, there are theoretical reasons to doubt that doubly-occupied sites would survive cooling. In particular, hyperfine-changing collisions are likely an issue at the high densities in a 3D lattice. The RMS cloud size in each direction for an atom in the ground state of a harmonic trap is $x_{\mathrm{RMS}} \sim \sqrt{\hbar/(m\omega)}$. For ^{87}Rb in a trap with $\sim$100 kHz trap frequencies, this corresponds to $x_{\mathrm{RMS}} \sim 30$ nm. The atomic number density is then $\sim 10^{22}$ m^{-3} = 10^{16} cm^{-3}. The hyperfine-changing collision rates for ^{87}Rb atoms in the $F = 2$ manifold with various spin mixtures are measured in Ref. [65]. Although some spin mixtures are somewhat stable, most exhibit loss rate coefficients $G \sim 10^{-12}$ cm^3/s. With that coefficient and the densities achieved with two atoms in a 3D lattice site, the loss rate is $\sim 10^4$ s^{-1} and the typical lifetime is then $\sim$100 μs. That is very short compared to the $\sim$1 s cooling time and thus it is unlikely that doubly-occupied sites would survive. Notably although the loss rate for spin-polarized clouds was too low to be measured in Ref. [65], most mixtures show loss. During the 3D Raman sideband cooling, the atoms pass through the states $|F = 2, m_F = -2\rangle$, $|2, -1\rangle$, $|1, -1\rangle$ and

possibly others as well. If any combination of those states exhibits this loss, then the doubly-occupied sites are unlikely to survive cooling.

Thus prospects for cooling doubly-occupied sites are not promising, at least for the choice of states used for the Raman cooling in this work. It may be possible to perform cooling in the lower hyperfine manifold, which would not exhibit hyperfine-changing collisions. However, that would limit the extent to which the optical pumping light could be detuned while maintaining a dark state.

7.3 Singly Occupied Lattice Site Preparation

Given that cooling multiply-occupied sites is a formidable task, strategies to increase the number of singly-occupied sites are worth investigating. It may then be possible to laser cool directly into a Mott-insulator-like state with one atom per site. The array of singly-occupied lattice sites would then qualitatively resemble a Mott insulator, though it may differ from a true Mott insulator state due to the more uniform lattice site occupation. For a true thermally populated lattice, a band insulator would occur near the bottom of the 3D lattice due to its overall harmonic confinement [2]. The Mott-insulator-like state may not show that same behavior, depending on the details of how it is populated.

Release-and-retrap cloud compression techniques, as shown in Refs [26, 24, 25], could be used in increase the lattice site occupation to ≈1 atom per site. If the compression is performed in 2D lattices, then Tonks gas physics may make it possible to preferentially load lattice sites with individual atoms. Adding a weak lattice in the third direction (along the compression direction) could be used to increase the effective interaction strength [126] and may increase the probability of single atom occupation. Alternatively, carefully controlled light-assisted collisions using slightly blue-detuned light could be used to preferentially remove single atoms from multiply-occupied sites. This has been shown to increase the efficiency of loading single atoms into optical

[2] This assumes that the 3D lattice is generated by Gaussian beams as is typically the case for AMO experiments. A flat 3D lattice would not have a bottom and so would not show that behavior.

tweezers to over 50% [131, 132]. We have not yet investigated the former approach, but some preliminary experiments with the latter approach have been performed.

The working principle behind the blue-detuned loading is as follows. First note that, as can be seen from Fig. 2-7, atoms are resonantly excited to the repulsive excited state potential when driven by blue-detuned light. Atoms are generally excited near their Condon point where the molecular interaction shifts the energy levels by an amount equal to the light's detuning, then they are repelled away from each other. In the limit that the atoms become well-separated before the excited atom decays, this releases an amount of kinetic energy equal to molecular level shift at the excitation point, which is equal to the light's detuning. Therefore, by illuminating the atoms with blue-detuned light, the amount of energy released by a light-assisted collision can be controlled to some extent. Now suppose two atoms are in a lattice site have a temperature small compared to the lattice site depth U_{site}. If the detuning is set to a value of $2U_{\text{site}}$ or greater, then both atoms will typically be ejected by a light-assisted collision. However, if the detuning is set slightly lower, then enough energy will be released for only one atom to escape and one atom can be left behind in the lattice site. Thus carefully controlled light-assisted collisions can be used to preferentially populate lattice sites with single atoms. This process isn't perfect. It is possible that neither atom escapes, or if the atoms have some initial energy it is possible that both atoms escape. However, with well-chosen values the probability of leaving a site singly-occupied can be significantly increased over 50%. In Ref. [132] the authors were able to achieve 80% singly-occupied loading in an array of 100 optical tweezers.

In order to characterize the performance of the blue-detuned loading in our system, an unresolved-sideband Raman cooling sequence (akin to those used in Ch. 4 and Ch. 5) was used to load a cold and dense cloud into the cODT. The 3D lattice was then ramped on over the cloud, leading to the loss of most of the atoms due to the high resulting densities. Further loss was then incurred when the cloud underwent 3D Raman sideband cooling. As shown in Fig. 7-4, the number of atoms which survived the cooling could be increased by applying an 0.5 ms pulse of blue-detuned light to the cloud at the start of the sideband cooling. This is indicative of the blue-detuned

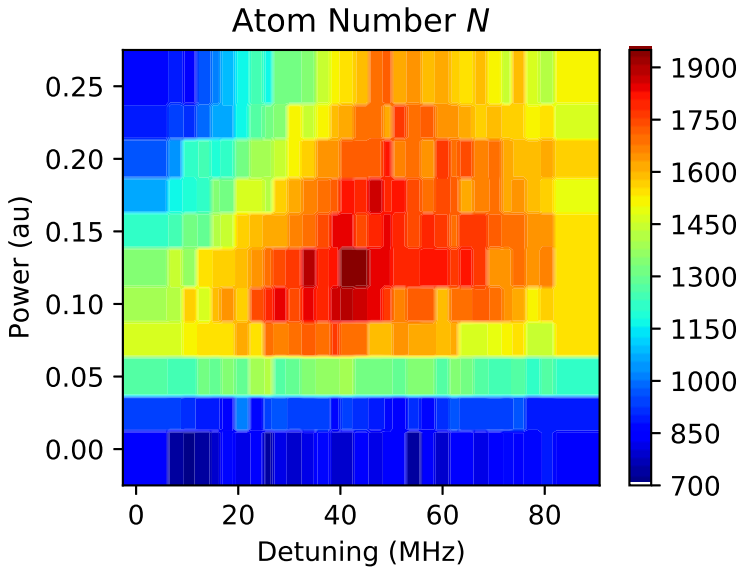


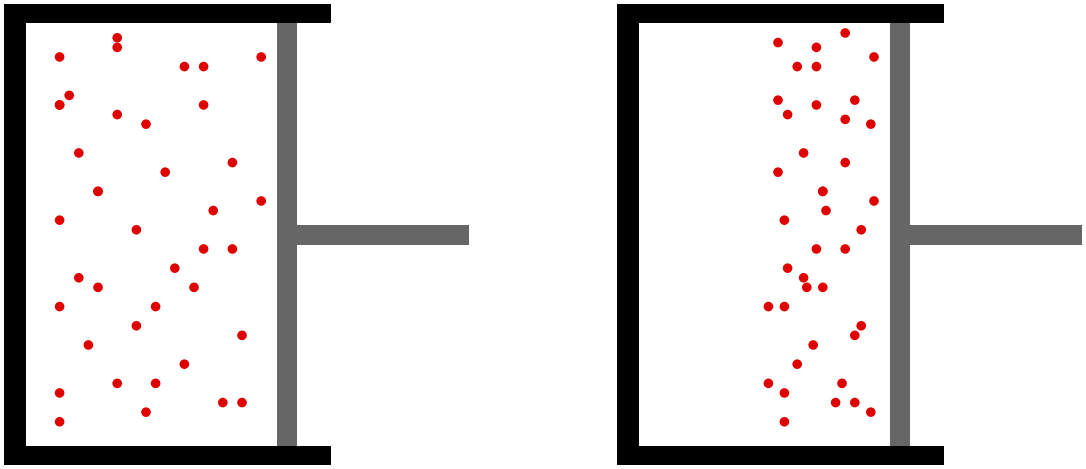
Figure 7-4: Remaining atom number after 3D Raman sideband cooling when an 0.5 ms blue-detuned light pulse with varied detuning and power is applied to the cloud at the start of the sideband cooling. The power is plotted in control units, which have a nonlinear relation to the actual beam power. The conversion has not been calibrated. With carefully chosen parameter values, the blue-detuned pulse can preferentially load lattice sites with single atoms, which can then survive the sideband cooling. The number of atoms remaining after the cooling is maximized when the detuning from the $F = 2 \rightarrow F' = 3$ resonance is $\approx +40$ MHz. The detuning accounts for the $\approx$35 MHz light shift of the $2 \rightarrow 3'$ resonance due to the trap light, so the optimal detuning is $\approx$75 MHz from the free atom $2 \rightarrow 3'$ resonance.

loading trick preferentially populating lattice sites with single atoms. The singly-occupied lattice sites then cannot undergo an undesired light-assisted collision during the sideband cooling and thus survive to the end of the sequence. Thus the blue-detuned loading technique for loading optical tweezers with single atoms also shows promise for preferentially loading 3D lattice sites with single atoms.

7.4 Release into cODT

After atoms have been cooled via Raman sideband cooling in a 3D lattice, it can be interesting to subsequently release the atoms into a cODT. This could make it possible to create a single 3D BEC from the cloud of atoms spread across lattice sites. That could be useful as a BEC production technique and it could be useful as a diagnostic to measure the performance of the Raman sideband cooling in the 3D lattice. That technique, combined with 3D Raman sideband cooling's ability to cool atoms to their ground state without collisions, might also make it possible to create BECs of species with unfavorable collisional properties.

Ref. [133] presents theoretical results for releasing atoms into a single trap, such as a cODT, after laser cooling them in a 3D lattice. However there are some subtleties in the assumptions in that work which imply that the results do not necessarily apply to our system. That work derives criteria for the required lattice site occupation probability and ground state population in the 3D lattice in order for a BEC to appear after releasing the atoms into the single trap. An outline of the calculation is provided below, and more details are available in Ref. [133]. It is assumed that multiply-occupied sites undergo loss until only zero or one atoms remain in each lattice site and atoms are laser cooled to a temperature which is the same for all lattice sites. The entropy per particle is calculated for the cloud in the 3D lattice as a function of the lattice site occupation probability and the probability that the atom (if present in the lattice site) is in the 3D ground state of the lattice site. That latter quantity is a function of the cloud temperature. The entropy per particle is also calculated for the final trap. It is then assumed that the atoms are released from the lattice



(a) (b)

Figure 7-5: Examples of a classical gas in a cylinder with a piston which demonstrate that a system must start in a thermal state in order for entropy to be conserved when making slow changes to the system. (a) The gas starts in a thermal state, filling all of its available space. If the piston is moved slowly, the entropy will remain constant. (b) Suppose the gas is artificially set up to occupy only half of the available space. The gas will immediately expand to fill the chamber (i.e. thermalize), which will increase the entropy. That entropy increase will occur if the piston is moved slowly, or even if the piston isn't moved at all. Thus the entropy is increased by thermalization if the system doesn't start in a thermal state, even if the system is nominally changed slow enough to be adiabatic.

adiabatically so that the atom number and entropy per atom are the same after the transfer to the single trap. The entropy per particle and atom number in the final trap can be used to calculate the final temperature, and thus the condensate fraction. In particular a critical entropy per particle at the BEC threshold can be calculated. No BEC forms when the entropy per particle is above the threshold, and a nonzero condensate fraction appears if the entropy per particle is below the threshold. If the entropy per particle is the same before the release, criteria can be placed on the site occupation probability and ground state probability achieved in the lattice in order to observe a BEC after release into the single trap.

However there are some subtleties about the requirements for adiabaticity. There are two criteria for the entropy to be conserved when the lattice is extinguished.

Firstly, the Hamiltonian must be changed slowly and adiabatically from the initial Hamiltonian to the final Hamiltonian. Secondly, the cloud must start in a thermal distribution, otherwise the entropy will increase no matter how slowly the Hamiltonian is changed. That second criteria often isn't explicitly stated in thermodynamics classes, but it is required in general. To see that this is a requirement, consider the simple example of a classical gas in a cylinder with a piston, depicted in Fig. 7-5. The gas can be adiabatically expanded by slowly moving the piston, which conserves the entropy if the gas starts in a thermal state. However, imagine a contrived scenario in which the gas is artificially prepared to initially occupy only half of the available space in the cylinder rather than filling the chamber all the way to the piston. Then the entropy will suddenly increase as the gas expands to fill all of its available space. That entropy increase occurs because the initial state (gas in only half of the cylinder) is not a thermal state, and the entropy increases as the state thermalizes by expanding to fill the chamber. In that case it doesn't matter if you move the piston slowly or not, the cloud's entropy will increase as it suddenly fills the chamber. This shows that the cloud must start in a thermal state in order for the entropy to be conserved, even if the Hamiltonian is changed very slowly.

Ref. [133] attempts to deal with that second criteria, namely that the cloud must start in a thermal state, in an interesting way. The initial distribution of atoms across the lattice sites does not follow a thermal distribution for the simple lattice Hamiltonian [133]

$$\mathcal{H}_{\text{lat}} = \sum_{\mathbf{i}} \sum_{\{j\}} \left(\epsilon_{\mathbf{i}} + \hbar\omega \sum_{\alpha=x,y,z} j_\alpha \right) \hat{\mathbf{n}}_{\mathbf{i},\{j\}} \tag{7.1}$$

where for simplicity it is assumed that the lattice frequency ω is the same for all directions in all lattice sites. Although Eqn. 7.1 looks complicated, its interpretation is relatively straightforward. The operator $\hat{\mathbf{n}}_{\mathbf{i},\{j\}}$ corresponds to the number of atoms in the 3D harmonic oscillator state $|j_x, j_y, j_z\rangle$ of lattice site $\mathbf{i}$. The $\epsilon_{\mathbf{i}}$ term is the energy offset of lattice site $\mathbf{i}$, which accounts for the fact that the bottom of different lattice sites can have different potential energies, e.g. due to the overall Gaussian

beam shape of the lattice beams. The energy of an atom in state $|j_x, j_Y, j_z\rangle$ of lattice site $\mathbf{i}$ is then $\epsilon_{\mathbf{i}} + \hbar\omega \sum_{\alpha=x,y,z} j_\alpha$. The energy of the whole system is then the sum of the energies of all of those $\{j\}$ states in each lattice site times the population of each of those states, summed overall all lattice sites $\mathbf{i}$. Generally thermal states for $\mathcal{H}_{\text{lat}}$ will have multiply-occupied lattice sites. However, if multiply-occupied sites are lost after laser cooling, then the resulting atom distribution will not have any multiply-occupied sites. Therefore the distribution is not a thermal distribution for the Hamiltonian $\mathcal{H}_{\text{lat}}$. Thus the entropy would not be conserved when releasing the lattice into a single trap.

Ref. [133] argues that although the distribution of atoms after laser cooling is not a thermal state for the Hamiltonian $\mathcal{H}_{\text{lat}}$, it is a thermal state for the Hamiltonian

$$\mathcal{H} = \mathcal{H}_{\text{lat}} + \mathbf{\Theta}\left(\hat{\mathbf{n}}_{\mathbf{i}}\right) \tag{7.2}$$

Here $\hat{\mathbf{n}}_{\mathbf{i}} = \sum_{\{j\}} \hat{\mathbf{n}}_{\mathbf{i},\{j\}}$ is the operator giving the number of atoms in lattice site $\mathbf{i}$. The effective interaction term $\mathbf{\Theta}\left(\hat{\mathbf{n}}_{\mathbf{i}}\right)$ is defined to be 0 when lattice site $\mathbf{i}$ has zero or one atoms and is defined to be ∞ when the site has two or more atoms. That effective interaction term makes it so that all states with multiple atoms in the same lattice site are energetically inaccessible with that Hamiltonian. Although that interaction is not real, it ensures that the thermal distribution of atoms under that Hamiltonian qualitatively matches the distribution of atoms after laser cooling in the lattice. Ref. [133] suggests that the ensemble of laser-cooled atoms in the lattice may be described as a thermal state of $\mathcal{H}$ and the entropy would then be conserved as the atoms are adiabatically released from the lattice into the single trap by slowly turning off that Hamiltonian.

One subtlety with the above argument has to do with the way lattice sites are populated. A thermal cloud under Hamiltonian $\mathcal{H}$ would populate each state in each lattice site with a probability proportional to the Boltzmann factor for the state, which is set by the state's total energy and the system's temperature T. Once consequence of this is that the distribution of the population of states within a lattice site would be

set by the temperature, and the temperature would be the same for all of the lattice sites. Given that the laser cooling likely performs similarly for all lattice sites for many systems, this is a reasonable assumption. However, in order for the atom distribution to be thermal it is also necessary that the probability of a lattice site having an atom is also set by its Boltzmann factor with the same temperature. In other words, for a thermal state of $\mathcal{H}$, the likelihood that a lattice site with energy $\epsilon_{\mathbf{i}}$ has an atom should be suppressed due to a factor of $\exp\left[-\epsilon_{\mathbf{i}}/(k_{\mathrm{B}}T)\right]$ in the Boltzmann factor. The actual ratio of populations between lattice sites will of course also be affected by the contribution from $\mathbf{\Theta}\left(\hat{\mathbf{n}}_{\mathbf{i}}\right)$ to the Boltzmann factor as well. The important point here is that the relative occupation between different lattice sites must follow a specific distribution set by $\mathcal{H}$, T, and N in order for the atomic distribution to be a thermal distribution for $\mathcal{H}$. That isn't necessarily the case for the laser cooled cloud. Typically the atoms follow some initial distribution across lattice sites determined by the dynamics before the laser cooling [3], followed by loss of atoms in multiply-occupied sites. Thus the distribution of atoms across sites doesn't necessarily follow a thermal distribution for $\mathcal{H}$. This implies that the entropy should generally be expected to increase when the cloud is allowed to thermalize (e.g. atoms are allowed to move between lattice sites or are combined into one trap), even if the Hamiltonian is changed very slowly.

The difference between the populations of lattice sites for a thermal distribution with Hamiltonian $\mathcal{H}$ compared to the distribution after laser cooling can be seen with a relatively simple example. Suppose that lattice sites are loaded randomly before the cooling and some random lattice sites undergo light-assisted collisions and lose their atoms during the laser cooling. Further suppose that the laser cooling cools all of the atoms to zero temperature locally, i.e. to the ground state of their lattice site. This produces an array of atoms in their ground states with random holes interspersed

[3] For example, if the lattice is ramped up over a MOT then the density distribution will likely be fairly uniform due to the repulsive radiation trapping force in a MOT. As another example, if the lattice is ramped on over a trapped thermal cloud, then the distribution across lattice sites will be determined by the initial cloud's temperature rather than the temperature of atoms in the 3D lattice after cooling.

throughout the array where atoms were lost or were never there to begin with. Now consider what the thermal state of $\mathcal{H}$ at zero temperature is. Because the temperature is zero, the system must be in its absolute ground state. Due to the $\mathbf{\Theta}\left(\hat{\mathbf{n}}_\mathbf{i}\right)$ term, multiply-occupied sites are not allowed. Therefore the ground state would be to have the N atoms occupying the N lowest-lying lattice sites (i.e. the N lattice sites with lowest $\epsilon_\mathbf{i}$), and each atom would be in the 3D ground state of its lattice site. Note that this thermal state of $\mathcal{H}$ is different than the distribution of atoms after laser cooling them locally to zero temperature. This is because the laser-cooled array of atoms contains defects while the thermal state is defect-free. Increasing the temperature of the thermal distribution would lead to holes in the array which could make it more qualitatively resemble the laser-cooled cloud. However, that temperature would also lead to atoms populating higher-energy states in lattice site rather than just the 3D ground state. In other words there is no temperature for which the thermal state of $\mathcal{H}$ is the same as the distribution produced by the laser cooling. This implies that the laser cooled cloud is not in a thermal state of $\mathcal{H}$ and therefore its entropy will increase if it is allowed to thermalize. In particular, thermalization here under Hamiltonian $\mathcal{H}$ would involve allowing the atoms in the laser-cooled cloud to move between lattice sites. The atoms would then be able to exchange energy between excitations in their local lattice site and potential energy due to the energy offset of their lattice site. If the atoms were allowed to do that then the distribution of population across lattice sites would change and some atoms would occupy lattice site states other than the 3D ground state. The system could then be described by a thermal distribution, but the entropy would have been increased during this thermalization.

Although not pointed out in Ref. [133], there is a subtlety associated with the adiabatic transition from $\mathcal{H}$ to that Hamiltonian of a single trap as well. Simply slowly extinguishing the lattice does not adiabatically transition $\mathcal{H}$ to the single trap Hamiltonian. The issue is that the $\mathbf{\Theta}\left(\hat{\mathbf{n}}_\mathbf{i}\right)$ term must also be adiabatically turned off. However, that term was just added in by hand to make the thermal distribution for the Hamiltonian look more like the clouds resulting from laser cooling in a 3D lattice. There is no actual infinite on-site interaction and so there is no $\mathbf{\Theta}\left(\hat{\mathbf{n}}_\mathbf{i}\right)$ contribution

that can be adiabatically turned off. Thus turning off the lattice will generally result in an increase in entropy no matter how slowly it is done. It is almost as if there were a $\boldsymbol{\Theta}\left(\hat{\mathbf{n}}_{\mathbf{i}}\right)$ term present when the initial distribution was made, but it was suddenly turned off at the end of the cooling. The distribution is not a thermal distribution for the actual Hamiltonian $\mathcal{H}_{\text{lat}}$, so the entropy will generally increase if the cloud is allowed to thermalize.

One example that demonstrates this issue is the following thought experiment which takes advantage of the reversibility of adiabatic processes. Imagine a lattice with an overall Gaussian profile which is released into a harmonic trap, as in Fig. [3] of Ref. [133]. Suppose that every lattice site has exactly one atom and every atom is perfectly cooled to the ground state of its lattice site. That state has zero entropy and thus should become a pure BEC after release to the cODT if the entropy is conserved. If the release to the cODT is adiabatic, then it should be reversible. Thus this would imply that adiabatically turning on a lattice over a pure BEC would create a lattice with one atom in each site. However, it is well known that turning on a lattice with a Gaussian profile over a BEC creates singly occupied sites in some locations, but generally creates multiply-occupied sites nearer the bottom of the lattice. This contradicts the assumption that the singly-occupied lattice is adiabatically connected to the pure BEC, and thus it must be that entropy is not conserved when releasing that lattice into the cODT.

In order to adiabatically connect the singly-occupied lattice to the pure BEC, a real $\boldsymbol{\Theta}\left(\hat{\mathbf{n}}_{\mathbf{i}}\right)$ term would have to be implemented. Suppose that it could be implemented and that the initial distribution was truly a thermal distribution for $\mathcal{H}$ (i.e. the issue of relative population in different lattice sites was also resolved). In that case, the release from the cODT would look like the following. The lattice could be ramped down until tunneling between lattice sites could occur. Then $\boldsymbol{\Theta}\left(\hat{\mathbf{n}}_{\mathbf{i}}\right)$ could be slowly turned off, allowing atoms to adiabatically tunnel towards the bottom of the lattice. This could possibly create multiply occupied sites, which would be allowed due to the finite on-site interaction energy as $\boldsymbol{\Theta}\left(\hat{\mathbf{n}}_{\mathbf{i}}\right)$ is turned down. Then the lattice can be adiabatically turned off the rest of the way. This gently and adiabatically releases

the atoms into the cODT, turning off both $\mathcal{H}_{\text{lat}}$ and $\boldsymbol{\Theta}(\hat{\mathbf{n}}_\mathbf{i})$ adiabatically and leaving behind a pure BEC. The reverse direction also makes sense. To create a singly-occupied lattice from a pure BEC one could slowly turn on a lattice and slowly turn up the on-site interaction energy until atoms avoid each other so much that each lattice site is singly occupied.

The non-adiabaticity due to the lack of a real $\boldsymbol{\Theta}(\hat{\mathbf{n}}_\mathbf{i})$ term and non-thermal initial state is demonstrated experimentally in Fig. 7-6. Atoms are Raman sideband cooled in a 3D lattice, then trap beam powers are ramped down. Most of the lattice beams are fully turned off, but two are left on to a varied power to create a cODT. The two cODT beams are attenuated from their maximum powers by the same factor. If the entropy is always the same before and after the release, then the temperature after release should scale with the square root of the cODT attenuation [4]. However, in Fig. 7-6 the temperature scales linearly with the cODT attenuation. That indicates that the entropy is not the same for the various final trap depths, and thus the entropy must not be conserved during release to the 3D lattice. This is further confirmed by calculating the phase space density in the cODT. That value should be independent of the cODT attenuation if the final entropy is constant, but it clearly increases at lower cODT powers.

If the lattice were turned off too quickly, then the release to the cODT would be non-adiabatic and the final entropy may not be constant. However, that is not the issue here. For the lattice ramp to be nominally adiabatic (neglecting the $\boldsymbol{\Theta}(\hat{\mathbf{n}}_\mathbf{i})$ term) until the atoms are released from their lattice sites, the rate of change in trap frequency must be small compared to the trap frequency squared $\dot{\omega}/\omega^2 = \epsilon \ll 1$ until the lattice site trap frequency approximately equals that trap photon recoil energy $\hbar\omega \approx E_{\text{r,trap}}$. At that point tunneling begins to take over. The beams are exponentially ramped off with a 0.5 ms timescale for the power, which corresponds to a 1 ms timescale for the trap frequency because the trap frequency scales with the

[4] If the entropies are the same in the cODT with different trap beam powers, then the factor $\exp[-\hbar\omega/(k_{\text{B}}T)]$ should be the same so that the populations in each state are the same. For that to be true, the ratio ω/T must be fixed, so T scales linearly with ω. The trap frequency scales with the square root of the cODT power attenuation, so T must scale with the square root of the cODT power attenuation as well in order for the entropy to remain constant.

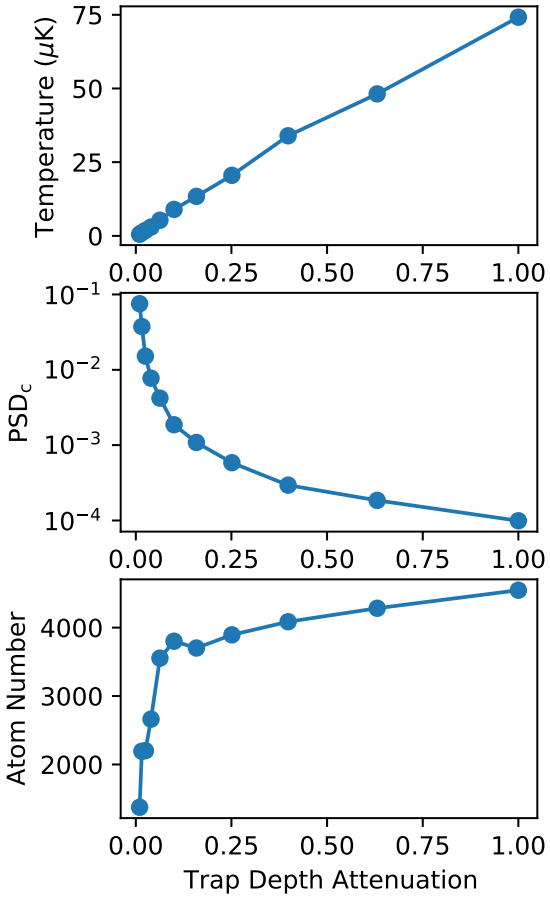


Figure 7-6: Cloud parameters after performing Raman sideband cooling in a 3D lattice then "adiabatically" turning off the lattice to release the cloud into a cODT. That sequence is repeated for many different final trap depths and results are plotted as a function of the attenuation of the cODT depth. Both cODT beams are at attenuated by the same factor. If entropy is conserved, then the final temperature should scale in proportion to the square root of the trap depth and the PSD should be independent of the trap depth. However, this is not the case. The final temperature scales linearly with trap depth, likely indicating that atoms roll down the sides of the cODT when they are released, picking up an amount of kinetic energy proportional to the trap depth in the process. This occurs despite the fact that the lattice release is very slow (see text). At very large attenuations there is significant atom loss and the PSD_c increases, indicating evaporation.

square root of the beam powers. With that timescale, $\dot{\omega}/\omega^2 \approx 0.1$ when $\hbar\omega = E_{\mathrm{r,trap}}$ and so $\mathcal{H}_{\mathrm{lat}}$ is turned off adiabatically.

The actual explanation for the temperature behavior seen in Fig. 7-6 is much simpler. The lattice is turned off slowly, so atoms are gently released from their lattice location. The atoms then freely roll down the sides of the cODT, releasing their potential energy in the process. The amount of energy released depends on the location of the atom but is proportional to the trap depth of the cODT. That released potential energy is the dominant source of energy in the final cloud, so the final temperature is proportional to the trap depth. This gives the linear relation between cODT attenuation and temperature seen in Fig. 7-6.

At this point it is worth noting that care must be taken if the blue-detuned loading suggested in Section 7.3 is used in conjunction with the release to a single trap. Generally the lattice depth is less than the trap depth. This is particularly true if the lattice beams are configured to drive the Raman transitions. That requires their polarizations to differ, in which case the beams do not fully interfere. When the lattice depth is less than the trap depth, it is possible to drive blue-detuned light-assisted collisions which give an atom enough energy to escape its lattice site but not enough energy to escape the trap completely. The atom could then settle into a high-lying lattice site, and then could release significant energy when it rolls down the confining potential in the cODT. On a related note, it is also possible for the atom to travel up one of the beams to a region outside of the 3D lattice where only one or two beams have non-negligible intensity, which is also undesirable.

As mentioned above, one possible way to resolve the issue with the lack of a real $\boldsymbol{\Theta}(\hat{\mathbf{n}}_\mathbf{i})$ term would be to intentionally implement a real $\boldsymbol{\Theta}(\hat{\mathbf{n}}_\mathbf{i})$ term. This could possibly be done with a Feshbach resonance. Using the Feshbach resonance to tune the scattering length from zero to infinity can tune the on-site interaction energy from zero to infinity. However, the effects of atom loss in multiply-occupied sites must be considered when attempting this approach, and the loss coefficients can be affected by the Feshbach resonance. This also does not resolve the issue of whether or not the initial distribution of atoms across lattice sites is thermal.

Another solution is to carefully tailor the trap potential to circumvent these issues altogether. If the trap bottom is flat enough such that ramping up a lattice over a BEC would yield a pure Mott insulator with no multiply-occupied sites, then releasing a Mott-insulator like state can likely yield a BEC. In particular for a trap with a perfectly flat bottom, there is no overall potential for the atoms to roll down after being released from the lattice, thus circumventing the issue completely. Furthermore, the issue of relative lattice site populations is also moot if the lattice is flat and each site is occupied with the same probability. For that reason, the results presented in Fig. [2](a) of Ref. [133] should be accurate even without an actual implementation of the $\boldsymbol{\Theta}(\hat{\mathbf{n}}_\mathbf{i})$ term.

7.5 Outlook

Future work could benefit from larger trap frequencies and higher optical pumping scattering rates. These improvements would likely yield lower temperatures and higher populations in the 3D ground state of each lattice site, as suggested by the results in Section 7.1. Faster optical pumping could be achieved with more power or smaller detuning. In particular if multiply-occupied lattice sites cannot be cooled then near-resonant light may perform comparably to far-detuned light, which allow for a larger optical pumping rate for a given amount of available optical power. Future work on releasing the laser-cooled cloud from a lattice into a single trap could also benefit from the use of a flat-bottom trap, as this could work around some of the issues discussed in Section 7.4. Given the minimal amount of collisions in this approach, it may also be interesting to apply these techniques to species which have unfavorable collisional properties.

Chapter 8

Rydberg Cavity QED Apparatus

Before I started working on Raman cooling, I spent the first few years of my PhD developing an optical cavity experiment. Hao Zhang made the first draft of the design which consisted of two optical cavities which overlapped at their mode waists. I worked with him on the project during my first year, then took over on it after he graduated. Over the following years additional design requirements were added to the project at the same rate at which they were implemented. The project slowly evolved to target experiments on cavity quantum electrodynamics (cavity QED or cQED) with Rydberg atoms. After a few years, I eventually had to stop working on the project to take over running the Raman cooling experiment when Jiazhong Hu graduated. Although I was not able to see the cavity project through to completion, I did add several elements to it, and this chapter documents those contributions. In particular, this chapter documents the design as it stood when I last worked on the project. Since then, further work on it has been performed by Josh Ramette then Alyssa Rudelis, so the final version installed into our vacuum chamber may differ.

8.1 Design Requirements

The design requirements for the cavity apparatus changed over the years, but the list when I last worked on it was as follows:

1. Two optical cavities which overlap at their mode waists.

2. Microwave antenna for 6.8 GHz and RF wires.

3. One cavity with high-finesse and one asymmetric cavity with a "leaky" mirror.

4. Compatible with Rydberg atoms.

5. Optics for high resolution imaging.

6. Optics for tweezers.

7. Free spectral range approximately equal to half the ^{87}Rb ground state hyperfine splitting.

8. Improved optical access through the large viewports.

The implementation of each of these criteria is discussed in the following sections.

8.2 Overall Structure

Many details about components of the Rydberg cQED apparatus will be easier for the reader to understand if they first have seen a high-level overall description of the apparatus. Therefore this section includes images of the apparatus to provide context for the reader. The apparatus is somewhat complex, so images of subassemblies of the apparatus are also presented.

The heart of the Rydberg cQED apparatus is a pair of near-orthogonal optical cavities, depicted in Fig. 8-1, which have modes that overlap at their waists. Each cavity consists of two concave mirrors pointed towards each other. The structure is designed to allow two diagonal MOT beams and one horizontal MOT beam through as shown in Fig. 8-2, though the beams may clip on the structure to some extent. Two in-vacuum lenses are mounted around the cavities, one above and one below, as shown in Fig. 8-3. These provide high-resolution imaging paths to/from the atoms and will be discussed further in Section 8.8.

As shown in Fig. 8-4, the cavity ring and lens tubes are connected via support rods with platforms at the ends. The entire structure mounts to the bottom of a large

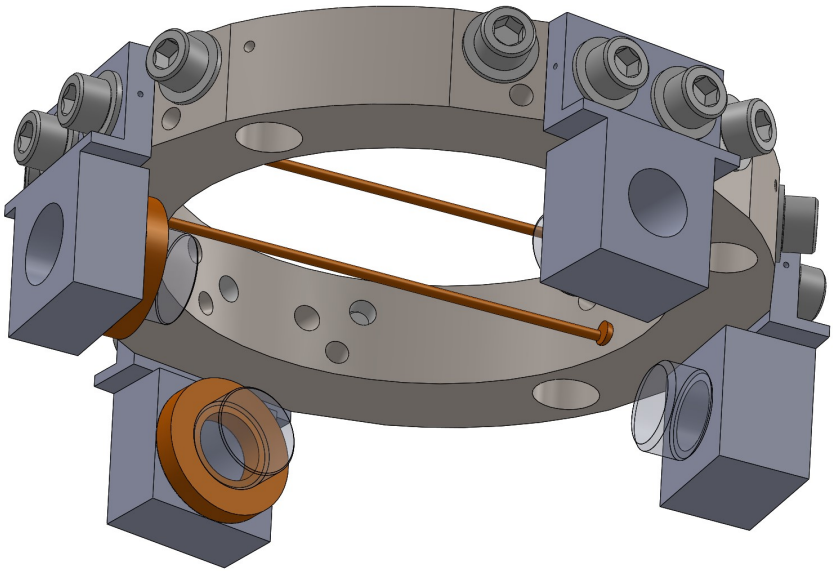
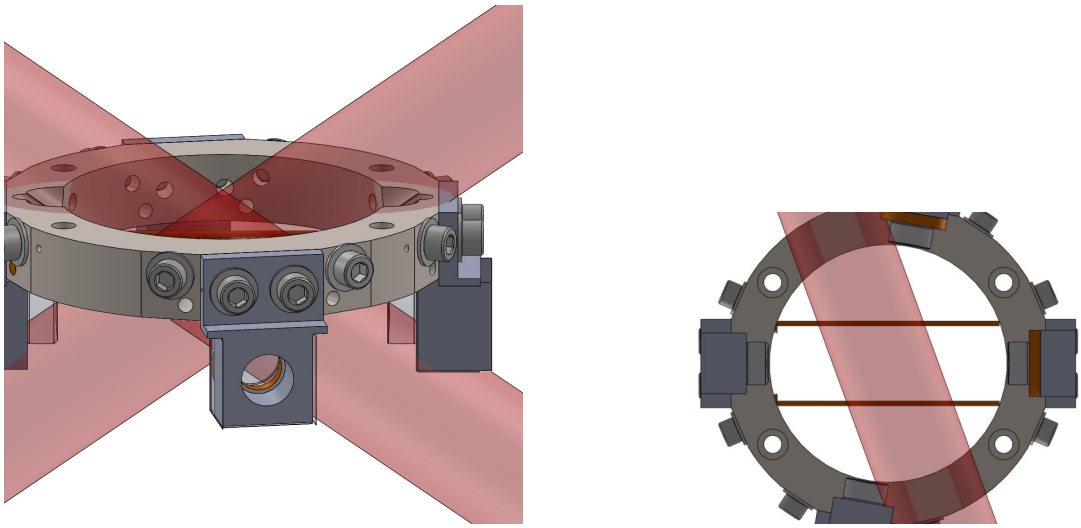
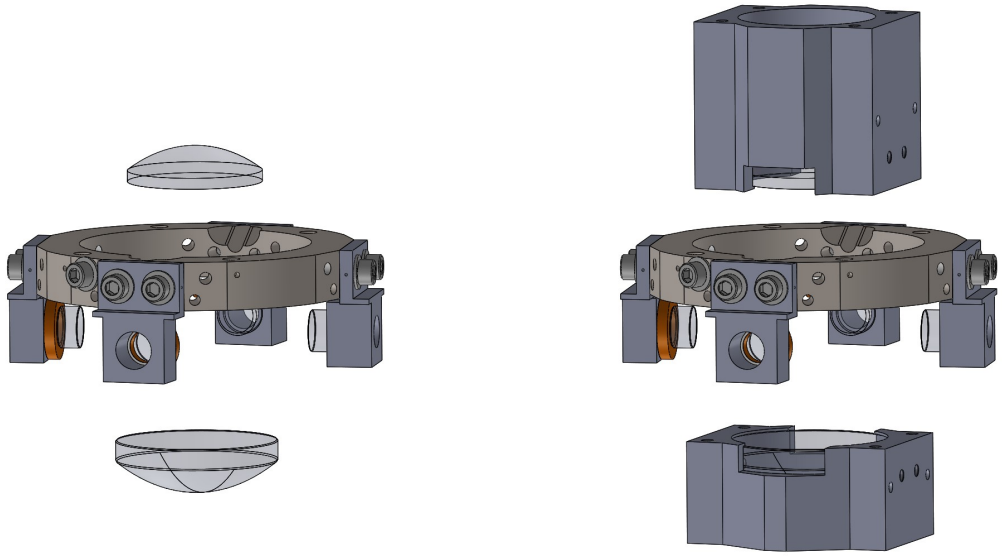


Figure 8-1: CAD rendering of the cavity ring structure with two optical cavities, each made from two mirrors. One mirror in each cavity has a piezo (brown-orange ring) for tuning the cavity length. Each mirror is mounted to a mirror holder which extends down below the ring. The mirror holders are attached via screws through clearance holes to the ring structure which allows some freedom in positioning the mirrors. The RF wires, discussed in Section 8.5, are visible along one cavity axis. One cavity axis is slightly rotated to make room for a MOT beam, so the cavity axes intersect at an 80° angle.



(a) (b)

Figure 8-2: The paths of the MOT beams through the cavity structure. (a) The diagonal MOT beams. Some material is removed from the ring to reduce clipping of those beams. (b) The horizontal MOT beam. The two cavities are not quite orthogonal to make room for this beam. One cavity is rotated 10°, so the cavity modes intersect at 80°.



(a) (b)

Figure 8-3: CAD rendering of the cavity ring and lenses (a) without and (b) with their lens tubes. The supporting structure for the cavity ring and lens tube is hidden for clarity.

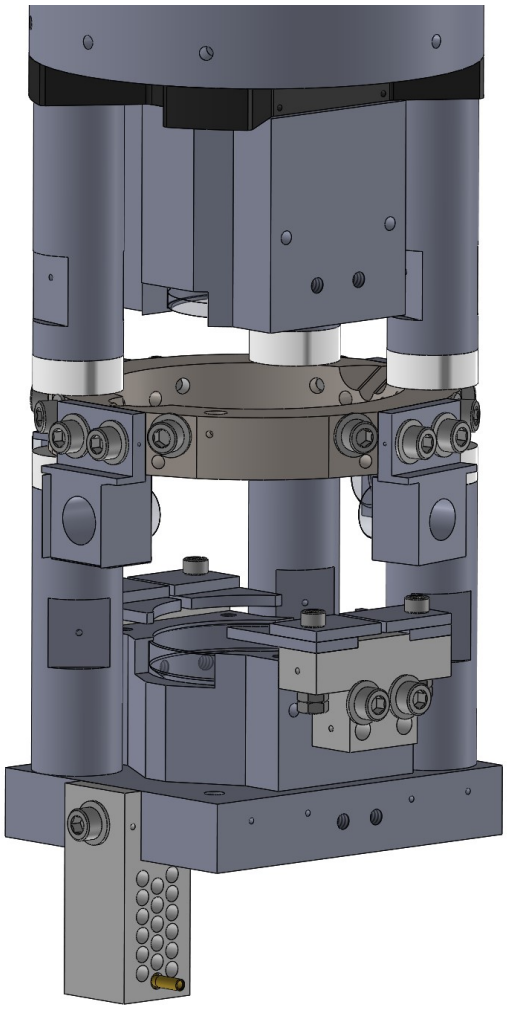


Figure 8-4: Rydberg cQED apparatus structure around the cavity ring. Support rods connect the rings to the surrounding components and platforms for the lens tubes. Vespel spacers are placed between the rods and the ring to provide some mechanical damping. The Macor cable connection organizer is visible at the bottom. Its design is similar to that of D-sub connectors and it is designed to house pins intended for D-sub connectors.

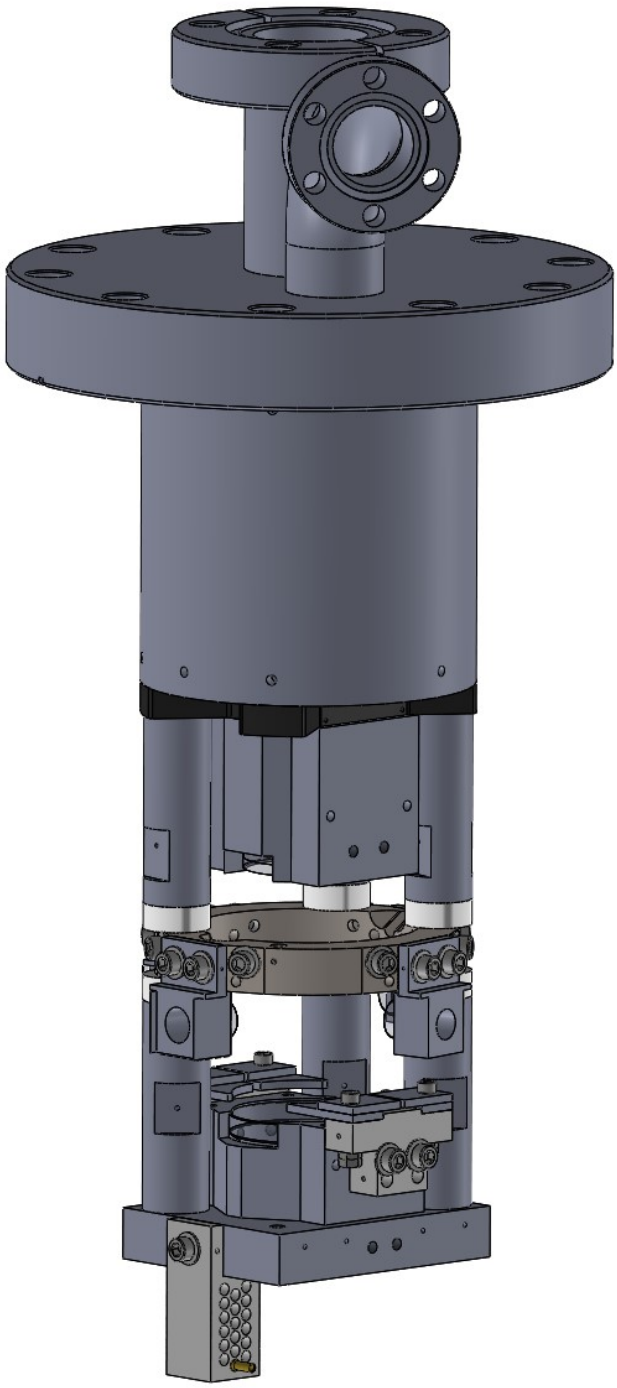


Figure 8-5: The full Rydberg cQED apparatus. The components visible in Fig. 8-4 are present in the bottom half of the image. The upper half of the image shows the vacuum flange to which the structure mounts, as well as cylindrical support piece that connects the rest of the apparatus to the flange.

Property	Value
Mirror Radius of Curvature	25 mm
Cavity Length	43.83 mm
Resonator g parameter	-0.75
780 nm Mode Waist	45 μm
1560 nm Mode Waist	64 μm

Table 8.1: Geometric properties of the cavities, which are the same for both cavities. These are the nominal values from the current version of the design; the actual values may differ due to tolerances or due to future changes to the design.

cylindrical piece, which in turn mounts to the top flange of the vacuum chamber as can be seen in Fig. 8-5. The flange has three CF joints: one 4.625" CF to connect it to the top port of the vacuum chamber, one 2.125" CF for a top viewport [1] and one angled 1.33" CF for wiring.

A photo taken during an assembly test run is presented in Fig. 8-6. Non-plated screws were used in this assembly test, which led to galling. The final assembly will be performed with silver-plated screws to avoid that issue. Silver-plated screws are available commercially and were purchased from UC components. By contrast, the other plated components were gold-pated at MTL, as will be discussed in Section 8.7.

8.3 Cavity Specifications

This section discusses the Rydberg cQED apparatus cavity parameters as of my last edits to the design. The design may change before it is installed into the vacuum chamber, so later theses and publications from our group should be consulted in order to determine that actual values. Furthermore, actual cavity performance can depend on additional details, such as the position of dust particles on the mirror surface. Therefore some values cannot be known with certainty until the apparatus is installed into the vacuum chamber and characterized. However, preliminary tests which were performed in atmosphere indicate what results should be achievable.

The two cavities have the same geometry, and the geometric features are summa-

[1] A 2.75" CF viewport will likely be used. It will be attached to the 2.125" CF viewport via a zero-length reducer.

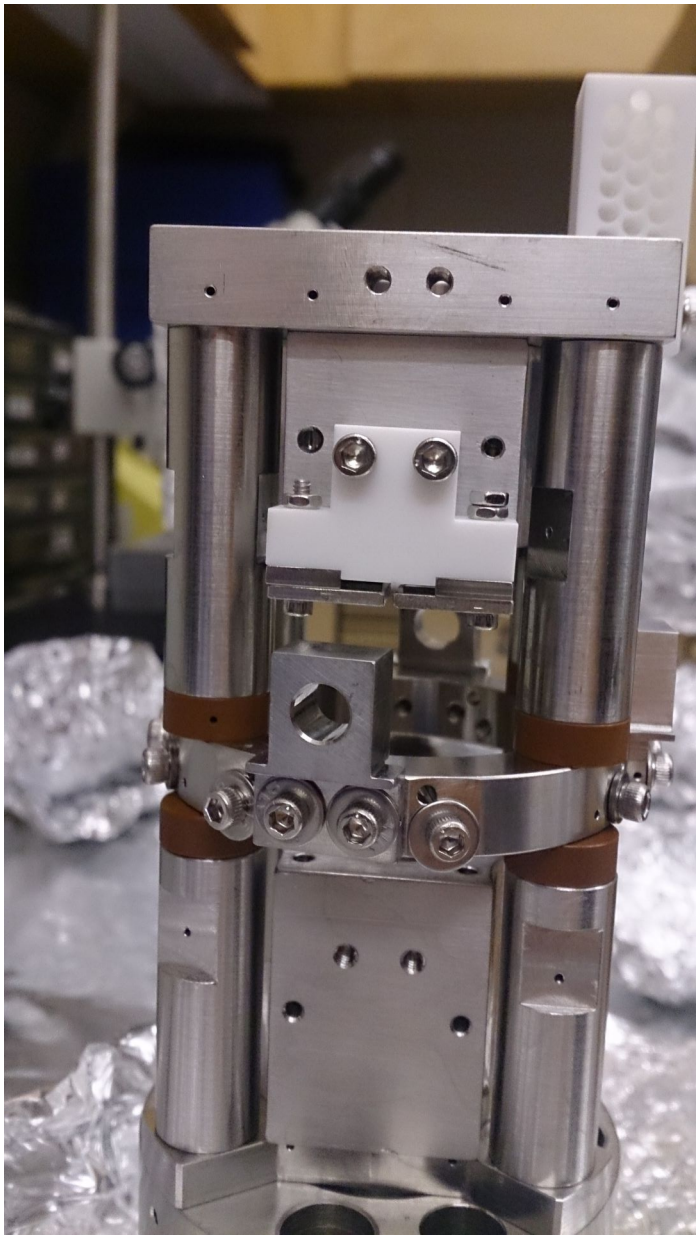


Figure 8-6: A photo of the Rydberg cQED apparatus during an assembly test to ensure that the pieces fit together properly. The apparatus is sitting upside down in the photo. The cavity mirrors and piezos have not yet been glued to their mounts, and none of the components have been gold plated.

Cavity	Wavelength	Transmission	Loss
HR	780 nm	5 ppm	7 ppm
	1560 nm	1.8 ppm	1.8 ppm
Leaky	780 nm	50 ppm	N.D.
	1560 nm	350 ppm	N.D.

Table 8.2: Approximate mirror transmission and loss for the high-reflectivity mirrors (HR) and the lower-reflectivity mirrors ("Leaky"). Mirror transmissions were measured directly by measuring the attenuation of light sent through a mirror. Losses were calculated after measuring the ringdown time of a test cavity of known length which was constructed with two mirrors of the given type. "N.D." indicates "not detectable" because the cavity light leakage due to mirror transmission was so large that the leakage due to scattering and absorption loss was too small to be accurately measured.

rized in Table 8.1.

Mirrors from different coating runs with different reflectivities have been considered for the apparatus. The mirrors with higher reflectivity were kindly given to us courtesy of Monika Schleier-Smith's group at Stanford. The "leaky" mirrors with lower reflectivity were purchased directly from Advanced Thin Films. The approximate transmission and loss coefficients for these mirrors are presented in Table 8.2. The values are approximate because they can vary somewhat from mirror-to-mirror. In particular the loss coefficients can vary significantly even for one mirror, increasing significantly if the light hits the mirror surface near a defect or piece of dust.

The mirror transmission coefficients were measured directly by measuring the attenuation of a beam directed at one mirror. Care was taken to clean the beam profile to avoid allowing light around the edges of the mirror and into the detector. In particular an iris was added to the beam path and the transmitted power was monitored as the iris was slowly closed. Initially the power would reduce as the iris closed because it would block some stray light in the wings of the beam. After the iris closed further, the transmitted power would flatten out and become more or less constant for a range. This corresponded to the accurate region where the iris blocked stray light but did not clip the beam. Finally, after the iris closed even further, the power would drop as the iris began to clip the beam. Often there would be a small increase in transmitted power when the iris initially clipped the beam, likely due to

light scattering off of the edges of the iris then propagating around the mirror and into the detector. The iris was set a position in the flat region, then the power incident on a photodiode was measured with/without the mirror in place. The transmission coefficient was then calculated by taking the ratio of those two powers.

The loss could not be measured in the same manner. Instead a cavity was constructed from two mirrors and the cavity ringdown time was measured. The ringdown time τ is equal to the time it takes light to make a round trip in the cavity (i.e. from one mirror to the other and back) divided by the fractional amount of light removed from the cavity per round trip. The former of those two quantities is simply $2L/c$ where L is the cavity length. The latter of those two quantities can be written as $\sum_i(T_i + l_i)$, which is the transmission coefficient T_i and loss coefficient l_i for each mirror i summed over both mirrors. Thus

$$\tau = \frac{2L/c}{\sum_i(T_i + l_i)} = \frac{1}{\mathrm{FSR}\sum_i(T_i + l_i)} \tag{8.1}$$

where $\mathrm{FSR} = c/(2L)$ is the cavity's free spectral range [2]. That can be rearranged to give

$$\sum_i(T_i + l_i) = \frac{2L}{c\tau} \tag{8.2}$$

If it is assumed that the mirrors have the same transmission and loss coefficients [3] T and l then that equation implies

$$l = \frac{L}{c\tau} - T \tag{8.3}$$

The loss values included in Table 8.2 were calculated using Eqn. 8.3. For the leaky mirror cavity, the dominant way light leaves the cavity is through mirror transmission

[2]Note that the FSR as defined by $c/(2L)$ is given in Hertz, not an angular frequency. Throughout this thesis frequency parameters are generally defined as angular frequencies, but the FSR is an exception because it is traditionally defined in this manner. With this definition, the cavity finesse is then $\mathcal{F} = (2\pi)\mathrm{FSR}/\kappa$ where κ is the cavity linewidth as an angular frequency.

[3]The transmission coefficients for different mirrors from the same coating run are typically fairly similar. However, losses can vary significantly because small defects and dust can create significant loss. However, if the light hits both mirrors on a clean and defect-free region then the loss is likely similar for the two mirrors.

Cavity	Wavelength	Finesse	Cooperativity	Linewidth	Ringdown Time
HF	780 nm	260,000	15	(2π)13 kHz	12 μs
	1560 nm	860,000	100	(2π)4 kHz	40 μs
Leaky	780 nm	91,000	5.3	(2π)38 kHz	4.2 μs
	1560 nm	18,000	2.0	(2π)190 kHz	0.82 μs

Table 8.3: Expected optical parameters for the Rydberg cQED apparatus cavities at the probe (780 nm) and trap (1560 nm) wavelengths. "HF" denotes the high-finesse cavity made with the higher-reflectivity mirrors, while "Leaky" denotes the asymmetric cavity made with one leaky mirror. The exact values can vary depending on e.g. how close the cavity mode is to dust or defects on the mirror surface. These are the expected values if the mirrors are clean and the cavity mode does not strike any significant defects on the mirror surfaces.

rather than loss. In that case, the two terms on the right side of Eqn. 8.3 are nearly equal. The measurement of T is then not accurate enough to provide an accurate value for l. This is why the loss for those mirrors is marked "not detectable" in Table 8.2.

As of this writing, then plan is to make one cavity with two high-reflectivity mirrors, then the second cavity with one high-reflectivity mirror and one "leaky" mirror with higher transmission. The expected optical parameters for these cavities at the trap and probe wavelengths are summarized in Table 8.3. The values for the high-finesse cavity are calculated from the measured ringdown time of a test cavity formed with two high-reflectivity mirrors. The values for the leaky cavity were calculated based on results from two separate test cavities, one being the test high-finesse cavity and the other being a test cavity formed with two leaky mirrors. A test asymmetric cavity with one leaky mirror and one high-reflectivity mirror was also setup, but it slightly underperformed compared to the expected results ($\sim$10% lower finesse). That likely occurred because the cavity mode hit one of the mirrors at a bad spot. If that occurs for the actual asymmetric leaky cavity used in the experiment, then the mirrors will likely be repositioned so that the cavity mode strikes them in a better spot. For that reason, the table instead shows the more optimistic expected performance based on the measurements of the other two test cavities.

In Table 8.3, the finesse $\mathcal{F}$ is calculated as

$$\mathcal{F} = \frac{2\pi}{\sum_i (T_i + l_i)} \tag{8.4}$$

and the linewidth (in angular frequency) is the inverse of the cavity ringdown time calculated from Eqn. 8.1. The cooperativity η here is defined as

$$\eta = \frac{24\mathcal{F}/\pi}{k^2 w_0^2} \tag{8.5}$$

where k is the wavevector and w_0 is the cavity mode waist. Calculating the values for the leaky cavity in Table 8.3 requires knowing the loss coefficients for the leaky mirror, which were not detectable. As an approximation they were assumed to be the same as the corresponding loss values of the high-reflectivity mirrors. This has little effect on the calculation results though because the dominant way light leaves the cavity is via transmission through the leaky mirror.

8.4 Overlapping Cavity Modes

The first design criteria for the cavity apparatus was that there should be two cavities which overlap at their mode waists. Hao created a design for this which formed the basis of the apparatus. It consisted of a ring structure with four mirror mounts arranged along the perimeter as depicted in Fig. 8-1. Much of the design and the procedures presented in this section are largely his ideas, however they are not presented in his thesis [78] and so are recorded here for completeness.

The mirror mounts do not have any actuators, but there is some tolerance around their clearance holes for the screws that attach them to the ring. The tolerance makes it possible to move the mirrors in a plane perpendicular to their cavity axis by loosening the screws then shifting the mirror holder. This freedom is necessary for overlapping the cavity modes. The alignment procedure consists of aligning light into each cavity separately, then adjusting one cavity's position to align with the other. Each of these steps will be discussed in its own subsection below.

8.4.1 Single Cavity Alignment Procedure

Aligning light into an existing cavity is somewhat difficult because most beam positions and directions produce no measurable signal. This is particularly true when the cavity is somewhat near concentric. Fortunately the mirrors are easily added or removed from the ring, which makes an easier alignment procedure possible. The cavity mode's axis passes through the center of curvature of each mirror [4]. Based on that a procedure was developed to align the center of curvatures of the mirrors with the beam one at a time.

First, the downstream mirror is installed onto the cavity ring structure and the upstream mirror is removed. The mirror holder's clearance holes are roughly centered around the screw to allow for future movement in any direction if necessary. The beam path is then oriented towards the center of the mirror and aligned so that the reflected beam overlaps with the incoming beam, which ensures that the beam passes through the center of curvature of the mirror. The reflected beam converges then diverges due to the curved surface of the cavity mirror, so it is necessary to check the beam overlap somewhat near the mirror itself before the beam has diverged too much. The beam may need to be walked to ensure that it both strikes the center of the mirror and reflects back on itself simultaneously. For precise alignment, a trick can be used to ensure that the beam hits the center of the mirror. If the light for the alignment comes from an optical fiber, then light at a wavelength for which the cavity mirror high-reflection (HR) and anti-reflection (AR) coatings do not work well can be sent through the fiber. Visual fault inspectors produce red light that works well for this approach for our cavities. The fiber for 780 nm light is still fairly single mode for the red light but the cavity coatings do not perform well for red light. The visual fault inspector's light follows the same path as the 780 nm light, but is partially transmitted and partially reflected at each mirror surface. That implies that some light is reflected off of the back side of the cavity mirror which has the flat AR-coated surface. If the beam is properly aligned, then both the reflection off of the HR-coated side and the reflection off of the AR-coated side will overlap with the incoming beam. The beam's

[4] For an excellent introduction to optical cavities, see Ref. [134].

alignment can be walked until this is the case. The reflection off of the AR-coated side remains collimated because the back surface of the mirror is flat, which makes it easy to distinguish from the reflection off of the curved HR-side of the mirror.

Once the first mirror and beam are aligned to each other, the mode matching lens can be inserted into the beam path. It can be carefully positioned in the plane perpendicular to the beam axis by ensuring that the beam is not deflected by the lens. In other words, when the lens is properly centered on the beam it should focus the beam but not change its direction. The position along the cavity axis is less critical and can be roughly measured out with a ruler. When picking a mode matching lens, the lensing due to the cavity mirror should ideally be accounted for. In practice I've found for our cavity geometry that neglecting that effect when choosing a mode matching lens can still give sufficient coupling for the alignment procedure. After the mode matching lens is inserted it may be worth repeating the alignment between the first mirror and beam if it has been perturbed by the presence of the lens. Ideally the lens should not affect it, but adjusting the beam path with kinematic mirror mounts is often more precise than the lens alignment so it may be helpful to do it again. This is particularly true if the lens is in a fixed mount positioned by hand, which is how we have typically mounted them.

Once the lens, beam, and first mirror are aligned, the second cavity mirror is inserted. It can be approximately positioned by ensuring that the reflection off of its HR-coated side overlaps with the incoming beam. Once that is done, light at the cavity wavelength should be sent into the fiber. The frequency of the light should be scanned back and forth over at least one FSR and the cavity transmission should be monitored. It can be helpful to use an APD to monitor the cavity transmission as they can be very sensitive and detect light even when the transmitted power is fairly low. It is also helpful to have a camera to view the transmitted light, as the camera's spatial resolution makes it easy to see which TEM mode(s) the light is coupling to. We often install a beam splitter after the cavity in order to use both an APD and camera at the same time. Additionally, lasers generally emit some amount of spontaneously emitted light which can propagate with the light from the lasing

Figure 8-7: Image of optical cavity transmission taken with a monitoring camera. A TEM mode is visible. The mode appears to be a $TEM_{1,3}$ mode, though it may actually be a higher order mode with more lobes that are outside the camera's field of view.

mode. The spontaneously emitted light can have a different wavelength than the lasing mode, often different enough to pass through the cavity mirror coatings. This can lead to the appearance of a dim spot in the camera image which marks were the beam would hit the camera if not for the cavity's mode structure. The beam most strongly couples to cavity modes which have a strong signal at the same position on the camera as that dim spot. Therefore that spot of spontaneous emission can be a helpful alignment tool as well.

Once the cavity transmission APD and/or camera are setup, the second mirror is roughly aligned, and the laser is scanning, then the second mirror's position can be more carefully adjusted. To do so, the mounting screws on the second mirror are loosened and the mirror's position is adjusted. This can be done coarsely by pushing it around directly with one's hands, and smaller adjustments can be made by tapping

on it. When the alignment is sufficiently accurate, cavity transmission should start to occur. A list of possible reasons why this may not occur is included near the end of this subsection. The cavity transmission should appear as spikes in the APD signal as the laser frequency scans past cavity resonances. On the camera there should be flashes of TEM mode patterns, such as the one shown in Fig. 8-7. When the alignment is somewhat bad, typically very high order TEM modes appear on the camera and the APD shows many resonances. This is because the light then couples to many different modes rather than just the Gaussian ($TEM_{0,0}$) mode. When the alignment is better, primarily low order modes dominate the transmission spectrum. This can be used as a signal to improve the alignment; simply adjust the second mirror's position so that primarily lower order TEM modes are seen on the camera.

Once the second mirror is reasonably well-aligned, a final set of adjustments can be made to finely align the beam to the cavity. The beam is typically steered using kinematic mounts so it is easier to precisely adjust the beam alignment than it is to precisely position the cavity mirrors. The signal for adjusting the beam alignment is the same as that used to position the second mirror: adjust the beam so that the light mainly couples to low-order TEM modes, particularly the Gaussian mode. Typically the horizontal position and angle are adjusted together, then the vertical direction is done, and back and forth until the alignment converges. For this stage it can be helpful to use the spontaneous emission spot as a guide. One horizontal steering knob is adjusted to move the spontaneous emission spot slightly towards the cavity axis, which is where the center of the TEM modes appears. Then the other horizontal steering knob is adjusted to improve the cavity coupling, then those two steps are repeated several times. The light should start coupling to different order modes. If the movement is toward lower-order modes then the process is continued in that direction. If the movement is towards higher-order modes, then the roles of the two adjustment knobs are switched. Once the alignment in the horizontal direction ceases to improve, the same procedure is performed on the vertical steering knobs as well. It can be necessary to switch back and forth between the two directions a few times. Eventually the light should couple well to the Gaussian cavity mode, at which

point the single cavity alignment is essentially complete. If maximal cavity coupling efficiency is desired, then the lens position can be adjusted and then the coupling reoptimized. That can then be repeated until the lens position has been optimized to maximize the coupling into the Gaussian mode, which should also minimize the coupling into the other modes.

It may be the case that no cavity modes are observed after the coarse alignment, which can have a few possible causes. The simplest is that the alignment may not be sufficiently accurate. In that case the light will not couple strongly to any of the cavity modes and the little transmission that does occur may be below the noise floor of the detection equipment. Another possible cause is that the cavity mirrors can get dirty. It can be helpful to examine the mirror surfaces under a microscope to look for dust or other contamination. If some contamination is present, then the mirrors will need to be cleaned.

8.4.2 Two Cavity Alignment Procedure

Once light has been coupled into each cavity, the two cavity modes must be overlapped. The cavity modes have a Rayleigh range large compared to their waist, so the cavity mode positions only need to be adjusted along their narrow direction in order to overlap well. The first step in that process is determining the displacement between the cavity waists. To do that, a slit (Thorlabs S200K) is mounted on a vertical translation stage with a high-precision positioning micrometer, as shown in Fig 8-8. A Thorlabs DM10 differential micrometer is well-suited for this purpose due to its $\approx$1 μm resolution. The slit is angled so that its wide direction is in the plane defined by the two cavity modes and its narrow direction is orthogonal to the cavity axes. The slit is also rotated about an axis orthogonal to the cavity plane to an angle such that both cavity modes can pass through it. The cavity transmission is then monitored as the slit is translated vertically.

Originally we determined the cavity positions by finding the position of the slit which maximized the transmission, However Muqing Xu pointed out that there is a more accurate way to find the cavity's center. The idea is that the cavity transmission

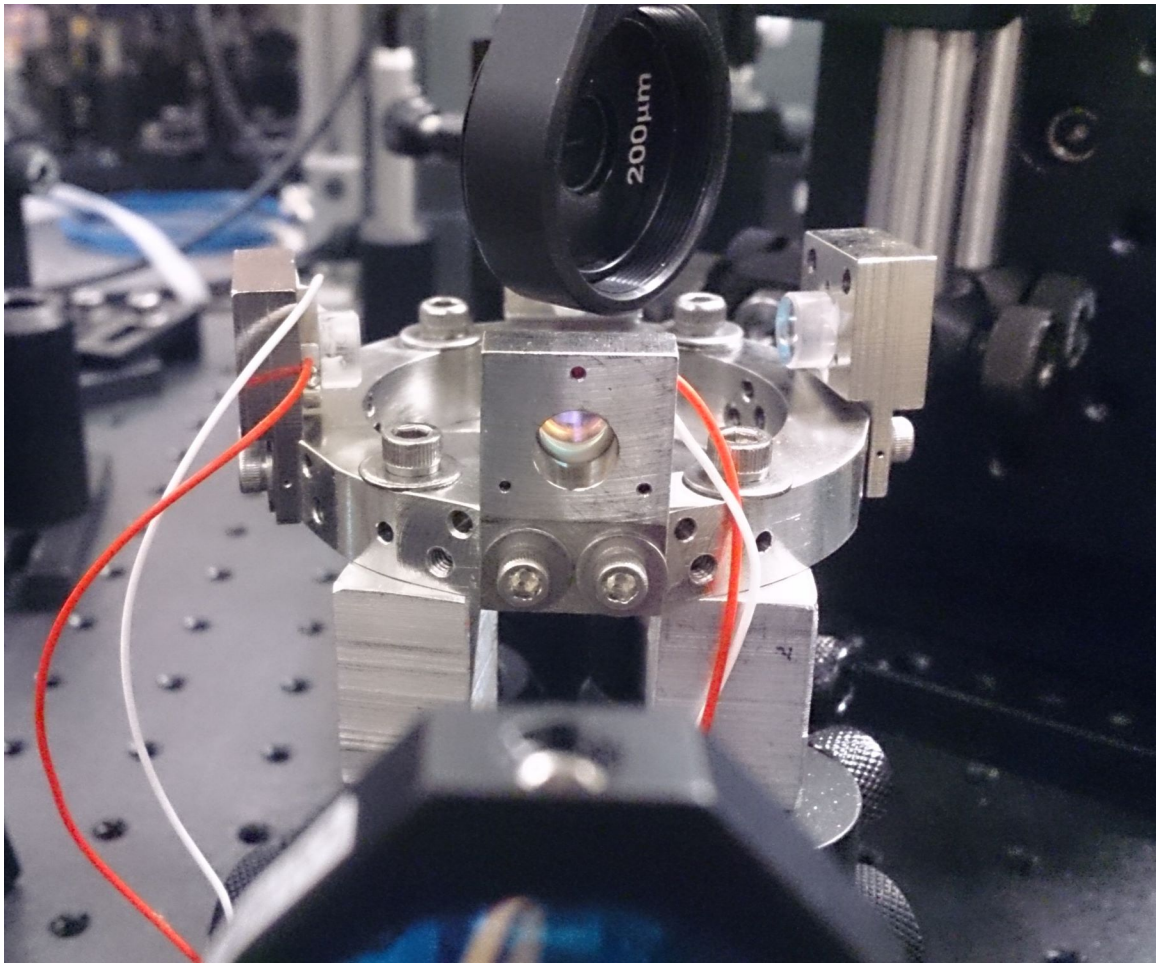


Figure 8-8: Setup for aligning the mode waists of two cavities. A slit is mounted on a translation stage such that it can be moved orthogonally to the plane of the cavities. The slit itself can be seen just to the left of the "200 μm" label. Cavity transmission appears when the slit is aligned with the cavity mode and disappears when the material around the slit blocks the cavity mode. The displacement between the cavity modes can be determined from how their transmission varies as the slit is translated. The cavities shown here are from an early test version of the apparatus, not the final version.

is insensitive to first order in perturbations to the slit's position around the maximum transmission point. That makes it somewhat inaccurate to find the cavity mode center by maximizing the transmission. One the other hand, the cavity transmission changes much more sensitively with position away from the peak transmission point. Therefore the cavity mode center can be found much more accurately by finding the two positions which cut the transmitted power by a given factor, then take the point halfway between those positions as the center. To reduce hysteresis effects, the slit is translated only in one direction while making this measurement. This consideration applies to optimizing symmetric signals in general, and when it is applied here it allows us to determine the cavity center position with $\approx$1 μm precision. That is the case despite the fact that the cavity Gaussian mode waists are $\approx$45 μm.

Using the above procedure, the centers of the two cavity axes can be found and their displacement determined. One of the cavities then needs to be adjusted to overlap with the other. Notably, adjusting the beam direction does *not* change the mode positions; the mode position is determined entirely by the cavity mirror positioning. Therefore it is necessary to move at least one of the mirrors of one cavity. This is most easily done by repositioning the upstream mirror of one of the cavities, as it requires the least effort to realign if it accidentally becomes very misaligned. Positioning the mirrors precisely is generally very difficult. Although many ideas were tried, the procedure that worked best was surprisingly crude. The best approach found was to loosen the screws on a mirror slightly, then tap on it with a screwdriver. The screws should be left tight enough so that the mirror stays in place on its own but slides slightly when tapped. The cavity modes can then be aligned by loosening the screws on a mirror, tapping on the mirror, retightening the screws, then measuring the cavity mode position. The tapping procedure does not provide well-controlled step sizes, but it can provide fairly small step sizes. That process is repeated many times until the cavity mode alignment is sufficiently good. It may be necessary to adjust the beam steering mirrors to keep the light well-coupled into the cavity as its position is adjusted. Often the displacement from the tapping will overshoot the target position. However, the tapping procedure can be repeated many times, moving

the cavity position back and forth across the other cavity's position until they happen to overlap very well.

In the end the accuracy of the overlap is not set by how precisely the position can be adjusted via the tapping; it is set by how accurately the overlap can be measured. The precision of the tapping just sets how many positioning attempts are required before the target accuracy is achieved. This is akin to sub-recoil laser cooling. The scattering does not necessarily bring the atom closer to zero velocity, but the atom will end up close to zero velocity by random chance after enough attempts. With this procedure the cavity mode positions can be aligned to $\approx$1 μm, which is the precision in measuring the cavity mode center.

8.5 Microwave and RF

The cavity apparatus is designed to facilitate manipulation of the atomic states within the ground state hyperfine manifolds. This generally involves radiation in one of two frequency bands, either microwave signals around 6.8 GHz to drive transitions between the hyperfine manifolds or RF signals $\lesssim$10 MHz to drive transitions within hyperfine manifolds. The microwave band has wavelengths on the order of the scale of the apparatus, while the RF signals have wavelengths large compared to the apparatus. For this reason, two different systems are design for applying those signals: one for microwaves and one for RF.

The microwaves are best transmitted along a coaxial cable to have a controlled transmission line impedance and to minimize radiation from the transmission line. For this reason a UHV-compatible SMA cable is used to carry the signal. Once the signal reaches the end of the transmission line, it must be coupled into free space via an antenna. Different antenna ideas were considered, but we settled on the simplest one. The idea is to simply take the end of a pigtailed SMA cable and split the inner and outer conductors apart to arrange them into a half-wave dipole antenna. Preliminary tests in air on a standard SMA cable have shown that this can couple microwave radiation into the air. The actual antenna will be mounted near the cavities

by attaching it to the cavity mounting structure.

RF will be applied to the atoms via a pair of wires that run along one cavity axis, which are visible in Fig. 8-1. Each lead of each wire will be connected to the vacuum feedthrough independently. This makes it possible to control the amplitude and phase of current through each separately. It also makes it possible to float the wires to use them as additional electrodes by applying a bias voltage to them, which will be discussed in Section 8.7. The control over amplitude and phase makes it possible to generate an arbitrary RF polarization in the plane perpendicular to the wires. Mounting holes for a second set of RF wires along the other cavity axis are included in the design which would allow for arbitrary RF polarization in any direction. However, these RF wires would clip the MOT beams and their shadows would lie very close to the cavity mode intersection region. For this reason the current plan is to omit the second pair of RF wires.

8.6 Lattice Registry in the Asymmetric Cavity

The optical cavities are designed for two wavelengths: one for trapping and one for probing. The probe wavelength is 780 nm, set by the ^{87}Rb D_2 line wavelength. The trap wavelength was chosen to be 1560 nm. This is because the lattice spacing for a lattice made with 1560 nm light is 780 nm. That makes it so that the antinodes of the trapping lattice (which are the trap minima for the atoms) line up with the antinodes of the probe wavelength standing wave, which maximizes the atom-cavity coupling.

The mirrors used in the cavity introduce phase shifts to the light as it is reflected, and that affects where the nodes of the standing wave lie. The size of the phase shift depends on the mirror coating and can generally be different for 780 nm light and 1560 nm light. If the mirrors on the opposite ends of the cavity introduce different phase shifts on the light, then the lattice can be shifted towards one side or the other. Furthermore, if the differential phase shift (i.e. the phase shift of 780 nm light upon reflection minus the phase shift of 1560 nm light upon reflection) is different for the

two mirrors, then one lattice may be shifted further than the other. This is not ideal as it can ruin the lattice registry, moving atoms trapped by the 1560 nm light away from the antinodes of the probe light standing wave.

For the symmetric cavity with two identical mirrors, this does not cause any issue because the phase shifts from the mirrors for a given wavelength are equal. This means that the 780 nm standing wave and the 1560 nm standing wave may be slightly stretched by the phase shifts, but they remain centered around the cavity's center. However, in the asymmetric cavity the situation is different. It is quite possible that the "leaky" mirror with higher transmission introduces a different phase shift than the higher-reflectivity mirror for a given wavelength, and the difference may not be the same for the two wavelengths. That can adversely affect the lattice registry as described above. Thus it is necessary to measure the displacement between the two standing waves for the asymmetric cavity.

Much of the previous efforts to map out the position of standing waves in an optical cavity relied on perturbing the cavity modes with sub-wavelength mechanical structures. For example in Ref. [135], a scanning near-field optical microscopy (SNOM) fiber tip was placed near the cavity modes and slowly moved while the cavity losses were monitored. The cavity losses increased when the SNOM tip was near an antinode and decreased when the SNOM tip was near a node. Monitoring the transmission of both the trap and probe light made it possible to determine the displacement between those two standing waves with ≈30 nm precision [135]. The same approach should be possible using nanowires/nanorods [136, 137, 138] or micromechanical membranes [139, 140]. In Ref. [141], a single trapped ion was used in a similar manner to map out cavity mode structure which provided 60 nm resolution, limited by the thermal motion of the ion. In this work an all-optical approach [142] was used instead, which yields much better precision. It also made better use of the equipment already available in our lab and did not require the purchase of expensive sub-wavelength mechanical devices. Of historical note, this work was completed in 2016, before Ref. [142] was published.

The all-optical procedure for determining the offset between the standing waves

of the asymmetric cavity is performed by characterizing two different test cavities: one with two leaky mirrors and one with two high-reflection mirrors. The displacement between the standing waves at the trap and probe wavelengths can then be determined from measurements made on those cavities. The process of making the required measurements is eased by the fact that some of the 1560 nm light is frequency doubled to 780 nm, which makes it possible to scan light at both wavelengths in sync. The procedure assumes that mirrors from the same coating run give the same phase shift and it assumes that the mirror phase shifts do not change significantly between adjacent cavity modes. That second assumption is valid for long cavities with length $L \gg \lambda$. In that limit, the fractional change in frequency when scanning over one FSR is small and so the change in mirror phase shift between adjacent cavity modes is small. The procedure has the benefit that all of the required measurements are frequency measurements. This is beneficial because frequency measurements generally can be very accurate. Some derivations are first required in order to determine which measurements are necessary, and these are done in the following subsections.

8.6.1 Cavity Resonant Frequencies

Light is resonant with the cavity if it accumulates a phase equal to a multiple of 2π during a round trip in that cavity. Using $k = 2\pi/\lambda = 2\pi\nu/c$ where ν is the frequency of the light and c is the speed of light, that condition can be written as[5]

$$2\pi(q+1) = \frac{2\pi\nu}{c}(2L) + 2(n+m+1)\phi_{\mathrm{d}} + 2(\pi + \phi_{\mathrm{m}}) \tag{8.6}$$

Here q is an integer related to the number of nodes in the standing wave along the cavity axis [6], while m and n are integers specifying the TEM mode of the resonance. In other words m and n specify the number of nodes along the transverse axes. L is the distance between the surfaces of the cavity mirrors along the cavity axis and ϕ_{m} is the phase shift due to reflection off of the mirror relative to a perfect reflector's π

[5] See e.g. Ref. [134], Ch. 19 Eqn.(22). Here that equation has been generalized to include phase shifts due to the mirrors.

[6] There is a node at each end of the cavity and $(q-1)$ nodes in between.

phase shift [7]. This assumes that the ϕ_{m} is the same for both mirrors; for the more general case $2\phi_{\mathrm{m}}$ can be replaced with the sum of the individual mirror phase shifts $\phi_{\mathrm{m},1}+\phi_{\mathrm{m},1}$. Generally ϕ_{m} will be different for different wavelengths, even for the same mirrors. Lastly ϕ_{d} is the diffractive phase shift due to the Guoy phase shift along the cavity. Its value is given by

$$\phi_{\mathrm{d}} = -\cos^{-1}\left(\pm\sqrt{g_1 g_2}\right) \tag{8.7}$$

where the positive sign is taken when $g_1, g_2 > 0$ and the negative sign is taken when $g_1, g_2 < 0$. The values of the "resonator g parameters" are defined as [134] $g_i = 1-(L/R_i)$ where R_i is the radius of curvature of mirror i. Note that ϕ_{d} depends only on the cavity geometry and not on the wavelength or cavity mode. Eqn. 8.6 can be rewritten as

$$\nu = \left[q-(n+m+1)\,\frac{\phi_{\mathrm{d}}}{\pi}-\frac{\phi_{\mathrm{m}}}{\pi}\right]\mathrm{FSR} \tag{8.8}$$

Where $\mathrm{FSR} = c/(2L)$ is the free spectral range.

8.6.2 Displacement Due to Mirror Phase Shift

A conversion between ϕ_{m} and standing wave position will be needed to determine how far the standing waves are shifted. To derive that relation, consider a cavity consisting of two mirrors: one idealized mirror which imparts no phase shift and one realistic mirror which imparts phase shift ϕ_{m}. Because only one mirror imparts a phase shift, the resonant frequency is

$$\nu = \left[q-(n+m+1)\,\frac{\phi_{\mathrm{d}}}{\pi}-\frac{\phi_{\mathrm{m}}}{2\pi}\right]\mathrm{FSR} \tag{8.9}$$

Taking the z to be the cavity axis and setting position of the idealized mirror to be at $z=0$ gives the following form for the electric field amplitude as a function of position

[7] Assuming that the index of refraction of the mirror's top layer is larger than the index of refraction of the intracavity region, which is typically air or vacuum. If the reverse is true, then the π phase shift does not occur. That factor of π does not affect the results of this section though as it can effectively be absorbed into q.

for the Gaussian ($m = n = 0$) mode

$$\begin{aligned} E(z) &= E_0 f(x,y) \sin\left[\frac{2\pi}{c}\left(q - \frac{\phi_\mathrm{d}}{\pi} - \frac{\phi_\mathrm{m}}{2\pi}\right)(\mathrm{FSR})\, z + \psi(z)\right] \\ &= E_0 f(x,y) \sin\left[\pi\left(q - \frac{\phi_\mathrm{d}}{\pi} - \frac{\phi_\mathrm{m}}{2\pi}\right)\frac{z}{L} + \psi(z)\right] \end{aligned} \tag{8.10}$$

Where $\psi(z)$ is the Guoy phase term and $f(x,y)$ is the transverse profile of the mode. The experiment's cavities use two mirrors with the same radius of curvature and we are interested in how the lattice is shifted near the center of the cavity as that is where the atoms will be trapped. The mode's waist is at the center of the cavity, so the Guoy phase is approximately zero there. In that region, the positions of the nodes z_p are approximately given by

$$z_p = \frac{pL}{q - \frac{\phi_\mathrm{d}}{\pi} - \frac{\phi_\mathrm{m}}{2\pi}} \tag{8.11}$$

where p is an integer which indexes the nodes in the cavity. In the limit of large q, i.e. the limit that the cavity length is large compared to a wavelength, this can be approximated as

$$\begin{aligned} z_p &= \left(\frac{pL}{q}\right)\frac{1}{1 - \frac{\left(\frac{\phi_\mathrm{d}}{\pi} + \frac{\phi_\mathrm{m}}{2\pi}\right)}{q}} \\ &\approx \left(\frac{pL}{q}\right)\left[1 + \frac{\left(\frac{\phi_\mathrm{d}}{\pi} + \frac{\phi_\mathrm{m}}{2\pi}\right)}{q}\right] \end{aligned} \tag{8.12}$$

From that, it is clear that the displacement of the node due to ϕ_m is

$$d = \left(\frac{pL}{2\pi q^2}\right)\phi_\mathrm{m} \tag{8.13}$$

In particular this is valid near the center of the cavity where $p \approx q/2$, thus

$$d = \left(\frac{L}{4\pi q}\right)\phi_\mathrm{m} \tag{8.14}$$

In short, introducing a phase shift ϕ_{m} to one mirror displaces nodes near the center of the cavity by an amount d towards that mirror, where d is given by Eqn. 8.14. This was derived assuming large q, which is valid when the cavity length L is much larger than a wavelength.

Eqn. 8.14 can also be derived in a simpler manner with some hand-waving. Consider the same cavity as above in which one mirror is idealized with no extra phase shift and the other imparts a ϕ_{m} phase shift. Near the realistic mirror, the electric field looks something like the following sum of incoming and reflected waves

$$\begin{aligned} E(z) &= E_0 \sin(kz - \omega t) + E_0 \sin(-kz - \omega t + \pi + \phi_{\mathrm{m}}) \\ &= 2E_0 \cos\left(\omega t - \frac{\phi_{\mathrm{m}}}{2}\right) \sin\left(kz - \frac{\phi_{\mathrm{m}}}{2}\right) \end{aligned} \tag{8.15}$$

From the sine factor it is clear that the mirror phase shift displaces the nodes of the standing wave by $\phi_{\mathrm{m}}/(2k) = \phi_{\mathrm{m}}\lambda/(4\pi)$ near the mirror. One can imaging that increasing ϕ_{m} stretches out the cavity mode like an accordion since the node at the idealized mirror stays fixed at its surface while the other end is moved outwards by $\phi_{\mathrm{m}}\lambda/(4\pi)$. Due to the accordion nature of the stretching, the nodes near the center of the cavity move half as far as those on the end, i.e they shift by $\phi_{\mathrm{m}}\lambda/(8\pi)$. For $q \gg 1$, q is approximately equal to the number of nodes in the cavity, which are spaced by $\approx\lambda/2$. This implies $q \approx 2L/\lambda$ and so $\lambda \approx 2L/q$. Inserting that into the expression for the displacement of nodes near the middle of the mode gives $L\phi_{\mathrm{m}}/(4\pi q)$, which agrees exactly with Eqn. 8.14.

Of particular interest is the difference in standing wave displacement for wavelengths which differ in q by a factor of two. This is the case [8] for the commensurate trap and probe standing waves. Applying Eqn. 8.14 to each wavelength then taking the difference yields

$$d_{q,2q} = \frac{L}{8\pi q}\left(2\phi_{\mathrm{m},q} - \phi_{\mathrm{m},2q}\right) \tag{8.16}$$

[8] If the trap wavelength corresponds to q, then the probe wavelength used should actually correspond to $2q \pm 1$. The antinodes of the q mode line up with the nodes of the $2q$ mode (neglecting shifts from the mirror and Guoy phase shifts), in which case the atoms would be trapped where they don't couple to the probe light at all. However for this characterization procedure we will be interested in measuring the the $2q$ mode.

where $\phi_{\mathrm{m},q}$ and $\phi_{\mathrm{m},2q}$ are the mirror phase shifts for the modes q and $2q$ respectively. The mirror phase shifts for those modes may be significantly different because they differ in wavelength by approximately a factor of two. The value of $d_{q,2q}$ specifies how much further mode q is moved towards the realistic mirror relative to mode $2q$.

The results in this subsection thus far have been derived for the case that one mirror has no phase shift. This was done so that the effects from one mirror's phase shift can be studied independently. The generalization to two real mirrors with phase shifts is straightforward. The second mirror shifts the standing wave in the same manner but in the opposite direction. Therefore the total shift of the q standing wave relative to the $2q$ standing wave is

$$\begin{aligned}\Delta d_{q,2q} &= d_{1,q,2q} - d_{2,q,2q} \\ &= \frac{L}{8\pi q}\left(2\phi_{\mathrm{m1},q} - \phi_{\mathrm{m1},2q}\right) - \frac{L}{8\pi q}\left(2\phi_{\mathrm{m2},q} - \phi_{\mathrm{m2},2q}\right)\end{aligned} \tag{8.17}$$

Where $d_{i,q,2q}$ is the value of $d_{q,2q}$ for mirror i as defined by Eqn. 8.16. Similarly $\phi_{\mathrm{m}i,q}$ and $\phi_{\mathrm{m}i,2q}$ are the phase shifts due to mirror i for the q and $2q$ modes respectively.

8.6.3 Straightforward Measurement Approach

Before discussing the approach actually used to measure the lattice registry, it is worth discussing a more straightforward approach. This approach is more intuitive but requires measuring two large and nearly-equal numbers then taking the difference between them, which makes it difficult to get precise results. The principle behind the measurement is that the values of the terms in Eqn. 8.8 can be determined by measuring the frequency difference between a few cavity modes and the absolute frequency of one cavity mode.

Consider a cavity with two identical mirrors. The frequency difference between the Gaussian $\mathrm{TEM}_{q+1,0,0}$ and $\mathrm{TEM}_{q,0,0}$ modes (i.e. the FSR) is $c/(2L)$ where the subscripts on TEM are q, m, and n respectively. Note that this is only true if the mirror phase shift is the same for the resonant frequencies $\nu_{q,0,0}$ and $\nu_{q+1,0,0}$ of those modes. That should be approximately true if $\nu_{q,0,0}$ and $\nu_{q+1,0,0}$ are very close, i.e.

$(\nu_{q,0,0} - \nu_{q+1,0,0})/\nu_{q,0,0} \ll 1$. This is valid for $q \gg 1$ as then the q contribution dominates the sum in the brackets in Eqn. 8.8. In that case $\nu \sim q\text{FSR}$ and so $(\nu_{q,0,0} - \nu_{q+1,0,0})/\nu_{q,0,0} \sim \text{FSR}/(q\text{FSR}) \sim 1/q$. Another way to think of it is that tuning light from mode q to $q+1$ requires shrinking the wavelength by enough to fit one half of an extra wavelength into the cavity. Doing so makes it such that the standing wave nodes line up with the ends of the cavity again (thus fulfilling the resonance condition), but with one more node in the middle than before. Making room for that extra half wave in the cavity requires shrinking the $q/2$ waves in the cavity by a combined total of $\lambda/2$, which shrinks each wave by $\approx\lambda/(q)$. The fractional change in frequency is then $\sim 1/q$, and so the change in frequency between adjacent modes is small for $q \gg 1$. Techniques for measuring the frequency difference between the $\text{TEM}_{q+1,0,0}$ and $\text{TEM}_{q,0,0}$ modes will be discussed in Section 8.6.5.

Once the FSR is known, the value of ϕ_d can be determined by measuring the frequency difference between the $\text{TEM}_{q,1,0}$ and $\text{TEM}_{q,0,0}$ modes (either m or n can be varied). Techniques for performing that measurement will also be discussed in Section 8.6.5. From Eqn. 8.8, that frequency difference is equal to $-\text{FSR}(\phi_\text{d}/\pi)$, so

$$\phi_\text{d} = -\pi\left(\nu_{q,1,0} - \nu_{q,0,0}\right)/\text{FSR} \tag{8.18}$$

Care must be taken to ensure that the q is the same for the two modes. The value of ϕ_d should be approximately determined from the mirror radii of curvature and the approximate cavity length, which can be used to estimate where the $\text{TEM}_{q,1,0}$ mode should be relative to the $\text{TEM}_{q,0,0}$ mode. That should be sufficient to identify the $\text{TEM}_{q,1,0}$ in the cavity transmission spectrum, then the actual frequency difference can be measured precisely to obtain a more accurate value for ϕ_d.

Once FSR and ϕ_d are known, the absolute frequency $\nu_{q,0,0}$ of the $\text{TEM}_{q,0,0}$ mode must then be measured. Generally the value of q is not known a priori, but it can be determined from Eqn. 8.8 by taking advantage of the fact that q must be an integer.

Rearranging that equation gives

$$q = \frac{\nu_{q,0,0}}{\mathrm{FSR}} + \frac{\phi_\mathrm{d}}{\pi} + \frac{\phi_\mathrm{m}}{\pi} \tag{8.19}$$

Although ϕ_m is not known, it is generally small. Therefore the value of q can be determined by inserting the measured values into Eqn. 8.19 and setting $\phi_\mathrm{m} = 0$, then rounding the value of q to the nearest integer.

Once the value of q has been determined, the only unknown parameter in Eqn. 8.8 is ϕ_m. Rearranging that equation one more time gives

$$\phi_\mathrm{m} = \pi \left(q - \frac{\nu_{q,0,0}}{\mathrm{FSR}} \right) - \phi_\mathrm{d} \tag{8.20}$$

Thus the mirror phase shift can be determined from the measurements listed above performed on a cavity with two identical mirrors.

The above measurements must be performed once for the trap wavelength and once for the probe wavelength to determine the mirror phase shifts for each. That process is performed twice, once for a cavity with two leaky mirrors and once for a cavity with two high-reflection mirrors. After that, the phase shift for each type of mirror at each wavelength is known and the relative standing wave shift can be determined from Eqn. 8.17, giving the final result. Note that the measurements performed on each test cavity are separate, so it is not necessary for them to have the exact same L or q. It is only necessary that L and q are fixed during all the measurements made on one test cavity at a given wavelength. There is also no requirement that L is the same for these test cavities as it is for the actual asymmetric cavity that will be used in the experiment, so the test cavities can be made longer to satisfy $q \gg 1$ if necessary. Of course the maximum cavity length is twice the mirrors' radius of curvature, as no stable cavity modes exist for separations larger than that [134].

It is worth noting that the terms in parentheses in Eqn. 8.20 are large and nearly equal. Therefore much precision is lost when the difference is taken, and so the absolute frequency $\nu_{q,0,0}$ must be measured very precisely. In particular it must be known with an error small compared to FSR in order to give a meaningful result. Further-

more the FSR must be known with similar fractional precision. In our experiment the FSR is ≈3.4 GHz, which is comparable to the accuracy of our wavemeter. Therefore the wavemeter cannot a give sufficiently precise measurement for $\nu_{q,0,0}$. It may have been possible to instead measure the laser frequency by measuring its beat note with a laser locked to an atomic transition, but a simpler approach was taken instead. That alternative approach does not require any absolute frequency measurements and is described in the following subsection.

8.6.4 Alternative Measurement Approach

When determining the absolute frequency of a cavity resonance with sufficient precision is inconvenient or impractical, an alternative approach can be used. This approach is based on measuring the difference between the $2q$ mode's resonant frequency and twice the q mode's resonant frequency. As will be discussed in Section 8.6.5, this was particularly easy for our experiment because some of the light that can drive mode q is frequency doubled to light near the $2q$ resonance.

As with the straightforward approach, this approach begins by setting up a test cavity with two identical mirrors. The FSR and ϕ_{d} are measured as before in the straightforward approach. However, instead of measuring the absolute frequency of a cavity mode, the frequency difference $\nu_{2q,0,0} - 2\nu_{q,0,0}$ is measured. Using Eqn. 8.8, that difference can be written as

$$\begin{aligned}\nu_{2q,0,0} - 2\nu_{q,0,0} &= \left[2q - \frac{\phi_{\mathrm{d}}}{\pi} - \frac{\phi_{\mathrm{m},2q}}{\pi}\right]\mathrm{FSR} - 2\left[q - \frac{\phi_{\mathrm{d}}}{\pi} - \frac{\phi_{\mathrm{m},q}}{\pi}\right]\mathrm{FSR} \\ &= \left(2\phi_{\mathrm{m},q} - \phi_{\mathrm{m},2q} + \phi_{\mathrm{d}}\right)\frac{\mathrm{FSR}}{\pi}\end{aligned} \tag{8.21}$$

This shows that the resonant frequency for the $2q$ mode isn't exactly twice the frequency of the q mode; it is offset due to both the difference in mirror phase shifts and due to the diffractive phase shift. That equation can be rearranged slightly to yield

$$2\phi_{\mathrm{m},q} - \phi_{\mathrm{m},2q} = \frac{\pi\left(\nu_{2q,0,0} - 2\nu_{q,0,0}\right)}{\mathrm{FSR}} - \phi_{\mathrm{d}} \tag{8.22}$$

This can be combined with Eqn. 8.16 to yield

$$\begin{aligned} d_{q,2q} &= \frac{L}{8\pi q}\left[\frac{\pi\left(\nu_{2q,0,0}-2\nu_{q,0,0}\right)}{\mathrm{FSR}}-\phi_{\mathrm{d}}\right] \\ &= \frac{L}{8q}\left[\frac{\left(\nu_{2q,0,0}-2\nu_{q,0,0}\right)}{\mathrm{FSR}}-\frac{\phi_{\mathrm{d}}}{\pi}\right] \end{aligned} \tag{8.23}$$

In the limit $q \gg 1$, the number of nodes in a cavity mode along the axis is approximately equal to q and each node is spaced by approximately $\lambda/2$. In that case $L/q \approx \lambda/2$, and so the above can be rewritten as

$$d_{q,2q} = \frac{\lambda}{16}\left[\frac{\left(\nu_{2q,0,0}-2\nu_{q,0,0}\right)}{\mathrm{FSR}}-\frac{\phi_{\mathrm{d}}}{\pi}\right] \tag{8.24}$$

Inserting Eqn. 8.18 makes it possible to express $d_{q,2q}$ in terms of only directly measurable quantities

$$d_{q,2q} = \frac{\lambda}{16}\left[\frac{\left(\nu_{2q,0,0}-2\nu_{q,0,0}\right)+\left(\nu_{q,1,0}-\nu_{q,0,0}\right)}{\mathrm{FSR}}\right] \tag{8.25}$$

Here λ is the wavelength for the q mode, which is approximately twice as large as that of the $2q$ mode.

At first glance this may seem no better than Eqn. 8.20 because taking the difference $(\nu_{2q,0,0} - 2\nu_{q,0,0})$ seems to require subtracting two large and nearly equal numbers. However, as will be discussed in Section 8.6.5, it is not necessary to measure those two values independently then subtract them. Instead that difference can be measured directly. Also note that a 1% error in measuring any of the quantities $(\nu_{2q,0,0}-2\nu_{q,0,0})$, $(\nu_{q,1,0} - \nu_{q,0,0})$, FSR, or λ only changes the value calculated for $d_{q,2q}$ by $\sim\lambda/10^3$. Therefore this approach can yield a value for $d_{q,2q}$ with an accuracy much better than a wavelength even from experimental measurements with modest precision.

To determine the final relative standing wave displacement in the asymmetric cavity, the above measurements are first performed on a test cavity with two leaky mirrors to determine $d_{q,2q}$ for that type of mirror. That process is then done again on a test cavity with two high reflection mirrors to get $d_{q,2q}$ for that type of mirror.

The displacement between the standing waves is then given by the difference of those two values, as in Eqn. 8.17. Using the results from this section, that equation can be rewritten as

$$\Delta d_{q,2q} = \frac{\lambda}{16}\left[\frac{(\nu_{1,2q,0,0} - 2\nu_{1,q,0,0}) + (\nu_{1,q,1,0} - \nu_{1,q,0,0})}{\mathrm{FSR}_1} + \right. \\ \left. - \frac{(\nu_{2,2q,0,0} - 2\nu_{2,q,0,0}) + (\nu_{2,q,1,0} - \nu_{2,q,0,0})}{\mathrm{FSR}_2}\right] \tag{8.26}$$

where $\nu_{i,q,m,n}$ denotes the $\mathrm{TEM}_{q,m,n}$ resonant frequency for the test cavity made with mirrors i. Similarly FSR_i denote FSR measured for those test cavities, which may be different as they may have different lengths.

8.6.5 Measuring Mode Frequency Differences

The above approaches for determining the difference in standing wave displacement require measuring the frequency difference between various cavity resonances. The most straightforward way to do this is to simply lock a laser to one of the resonances, then measure its frequency with a wavemeter or other device. That process can then be repeated for each resonance, then the differences between them can be taken. However, this has the same downside as the approach mentioned in Section 8.6.3, namely that it involves subtracting large and nearly-equal numbers.

Fortunately there is a simpler approach to measuring the frequency difference between resonances which can be done using a fiber EOM. This approach does not require a high-precision absolute frequency measurement of the light and does not require locking the laser. To measure the FSR, the light is run through a fiber EOM on its way to the test cavity. The laser frequency is repeatedly ramped and the cavity transmission signal is monitored on an oscilloscope, which is set to trigger on the cavity transmission peak as the laser passes a $\mathrm{TEM}_{q,0,0}$ resonance. When the fiber EOM is driven with a frequency close the the FSR, then additional transmission peaks appear next to the central one as the sidebands scan by other Gaussian modes with different q. The fiber EOM drive frequency can then be adjusted to align the transmission peaks of the sidebands with the transmission peak of the carrier. When

the peaks line up, the frequency of the fiber EOM drive is equal to the FSR. At that point the FSR can be determined by reading off the frequency of the microwave source if it has a display, or the frequency of the source can be measured with a spectrum analyzer.

When the light is scanned past a cavity resonance, the transmitted power shows ringing as it decays, giving a "ring-down" curve [143, 144]. A very precise alignment of the sidebands to the carrier can be achieved by ensuring that the ringing in their transmission signals are in phase. The light from the two modes typically is overlapped on the same photodetector, so the signals are not measured separately. However, when the ringing is not in phase between the carrier and sideband then it will tend to average out, and so they can be brought in phase by adjusting the fiber EOM drive frequency to maximize the amplitude of the ringing. The typical measurement precision for this approach is set by the linewidth of the cavity resonance, though the actual measurement can be more accurate by a factor of the SNR. Additional accuracy can also be achieved by tuning the sideband frequency to a multiple of the FSR. For example measuring 10(FSR) with a precision on the order of the cavity linewidth κ gives a value for the FSR with a precision on the order of $\kappa/10$. A similar benefit should apply to overdriving the fiber EOM so that higher-order sidebands are visible as well, however that can make triggering the oscilloscope more difficult.

The frequency difference between the $\mathrm{TEM}_{q,1,0}$ and $\mathrm{TEM}_{q,0,0}$ modes, needed to determine ϕ_d, can be measured in a similar way with a fiber EOM. Note that the beam may need to be misaligned slightly from the cavity axis in order to couple light into the $\mathrm{TEM}_{q,1,0}$ mode. Technically a trick analogous to measuring a multiple of the FSR is also possible for measuring ϕ_d. Instead of measuring the frequency difference between the $\mathrm{TEM}_{q,1,0}$ and $\mathrm{TEM}_{q,0,0}$ modes, one can measure the frequency difference between the $\mathrm{TEM}_{q,m,0}$ and $\mathrm{TEM}_{q,0,0}$ modes with $m > 1$. That frequency difference is m times more sensitive to the value of ϕ_d and so should provide a result m times more accurate for a given precision in aligning transmission peaks in the oscilloscope trace. However this is typically not so practical because a Gaussian beam does not couple well to those higher order modes, which reduces the SNR of their transmission peaks.

For that reason, the work presented here used the $\mathrm{TEM}_{q,1,0}$ and $\mathrm{TEM}_{q,0,0}$ modes to determine ϕ_{d}.

Interestingly the frequency difference $(\nu_{2q,0,0} - 2\nu_{q,0,0})$ can also be measured in a similar manner in our apparatus due to the presence of a frequency doubler for the light. The trap light, which is at the frequency near the q resonance, is split into two paths. One path goes to the cavity as before, while the other goes to the frequency doubler. The doubled light is near the $2q$ resonance and is at exactly twice the frequency of the original light; it even tracks the trap light's frequency as it scans. This is clear from energy conservation; the doubler simply combines two trap light photons into one photon with twice the energy. As before note that the $2q$ resonance is not at exactly twice the frequency of the q resonance due to the diffractive phase shift and mirror phase shifts.

Suppose a fiber EOM is placed in the doubled light's beam path and both the trap light and the doubled light are coupled into the cavity. The procedure is then the same as before; the trap light's frequency is ramped and an oscilloscope is triggered on the transmission signal. In this case the transmission signal can be easily measured separately for the trap light and for the doubled light because they significantly differ in wavelength. The fiber EOM drive frequency is then adjusted until the sideband of the doubled light hits the cavity resonance at the same point in the ramp as the trap light. The frequency difference $(\nu_{2q,0,0} - 2\nu_{q,0,0})$ is then equal to the fiber EOM drive frequency at that point. As advertised, the frequency difference is measured directly rather than by measuring each resonant frequency separately. No absolute optical frequency measurement is needed. The same procedure can also be performed by placing a fiber EOM on the trap beam's path rather than the doubled light's path if desired. This can be helpful if a fiber EOM is only available for that wavelength, or if the increased optical power handling of fiber EOMs at longer wavelengths increases the SNR of the transmission signals.

Measurement	Test Cavity 1	Test Cavity 2
λ_{trap}	1560.622 nm	1560.600 nm
FSR	3313.48 ± 0.1 MHz	3341.1 ± 0.3 MHz
$\nu_{q,1,0} - \nu_{q,0,0}$	2645.72 ± 0.1 MHz	2644.08 ± 0.15 MHz
$\nu_{2q,0,0} - 2\nu_{q,0,0}$	-1181.68 ± 0.1 MHz	-1134 ± 2 MHz

Table 8.4: Measured results of the test cavities set up to determine the relative standing wave shift for the trap and probe wavelengths in the apparatus's asymmetric cavity. Test cavity 1 was composed of two "leaky" mirrors and test cavity 2 was composed of two high-reflection mirrors. The wavelength of the frequency doubled light was measured with a wavemeter, then the trap wavelength was calculated by doubling that result. The frequency difference measurements were made with the techniques described in Section 8.6.5.

8.6.6 Results For Our Cavity

Given the resources available in our lab, we opted to measure the relative standing wave displacement for our trap and probe standing waves using the approach presented in Section 8.6.4. That approach is performed by setting up two test cavities: one using two leaky mirrors and one using two high-reflection mirrors. A series of measurements on the frequency differences between selected cavity modes is then performed on each test cavity using the techniques outlined in Section 8.6.5.

The test cavities for these measurements were approximately $L = 4.4$ cm long and all of the mirrors have a 2.5 cm nominal radius of curvature. The resonator g parameters are then approximately -0.76 and the FSR is approximately 3.4 GHz for both test cavities. The value of q for the 1560 nm trap mode is then $q \approx 2L/\lambda \approx 5.6 \times 10^4$, so errors from the theory approximations on the order of $\sim 1/q$ are in the 10^{-5} to 10^{-4} range. This is comparable to the errors for most of the measurements.

From Eqn. 8.7, the value of ϕ_{d} is approximately -2.4 rad. Eqn. 8.8 then implies that the $\text{TEM}_{q,1,0}$ mode is higher in frequency than the $\text{TEM}_{q,0,0}$ mode by about $(0.77)\text{FSR} \approx 2.6$ GHz. Assuming that the mirror phase shift contribution is smaller than ϕ_{d}, Eqn. 8.21 shows that the frequency of the $\text{TEM}_{2q,0,0}$ mode is lower than twice the frequency of the $\text{TEM}_{q,0,0}$ mode. These rough results can be used to identify the correct modes in the transmission spectrum, then the actual frequency difference can be measured precisely.

The measurements made on each test cavity are summarized in Table 8.4. The uncertainties for the frequency difference measurements were determined by tuning the fiber EOM drive frequency until the peaks in the transmission signal were visibly misaligned. Therefore the errors quoted are bounds, which makes them larger and more pessimistic than standard deviation error estimates. The precision of the frequency difference measurements for cavity 1, made with the "leaky" mirrors, was actually ∼0.01 MHz. However, the values tended to drift by ≈0.1 MHz over the timescale needed to make all of the required measurements so that value was used for the uncertainty. The ratio of cavity mirror transmission to loss for test cavity 2 at the trap wavelength was only ∼1, which reduced the amount of transmitted power. The trap light was detected with an unamplified InGaAs photodiode (Hamamatsu G12180-005) used in photovoltaic mode with a 60 kΩ load resistor. For these reasons, the trap light transmission peaks were only ≈20 mV tall on the oscilloscope, and so the SNR was not as high as for the other measurements. This is why the uncertainty of the $\nu_{2q,0,0} - 2\nu_{q,0,0}$ measurement for cavity 2 is much larger than the other uncertainties. The other measurements on cavity 2 could be made using the doubled light, which was detected with an amplified APD, thus giving a better SNR. The overall uncertainty in the final result for $\Delta d_{q,2q}$ will then limited by that of the $\nu_{2q,0,0} - 2\nu_{q,0,0}$ measurement to be $\sim(\lambda/16)(1/500) \approx 0.2$ nm. A future attempt at this measurement could narrow down the uncertainty by using more power at the trap beam wavelength or by using an amplified detector for the transmitted light at the trap wavelength to improve the SNR. However that precision is more than adequate for our purposes.

The resulting parameter values inferred from the measurements in Table 8.4 are summarized in Table 8.5. The value of ϕ_{d} was calculated using Eqn. 8.18 and the value for $d_{q,2q}$ was calculated using Eqn. 8.25. Finally, the value of $\Delta d_{q,2q}$ was calculated by taking the difference between the two values for $d_{q,2q}$, which is equivalent to using Eqn. 8.26. The final estimate is that the trap wavelength standing wave is shifted towards the high-reflectivity mirror more than the probe wavelength standing wave by 1.0 ± 0.2 nm. Again, the dominate source of error is the measurement of $\nu_{2q,0,0} - 2\nu_{q,0,0}$ for test cavity 2 due to the poor cavity transmission signal at the trap wavelength for

Measurement	Test Cavity 1	Test Cavity 2
ϕ_{d}	-2.5085 rad $\pm$ 0.1 mrad	-2.4862 rad $\pm$ 0.3 mrad
$d_{q,2q}$	43.10 ± 0.02 nm	44.1 ± 0.2 nm
$\Delta d_{q,2q}$	1.0 ± 0.2 nm	

Table 8.5: Resulting parameter values for the test cavities calculated from the measurements in Table 8.4. ϕ_{d} is the diffractive phase shift of the cavity mode due to the Guoy phase. The value of $d_{q,2q}$ specifies how far the trap wavelength standing wave is shifted relative to the probe wavelength standing wave by a mirror from the corresponding test cavity. A positive value indicates that using that mirror in a cavity shifts the trap standing wave further towards that mirror than the probe standing wave. The value of $\Delta d_{q,2q}$ shows the calculated displacement between the trap and probe standing waves for the asymmetric cavity which will be made using one mirror from each test cavity. A positive value there indicates that the trap wavelength standing wave is displaced further than the probe wavelength standing wave towards the (higher-reflectivity) mirror from test cavity 2. The result is that the standing waves will be offset by only $\approx$1 nm, which is very small compared to the 780 nm length scale of the standing waves. Therefore the lattice registry is excellent.

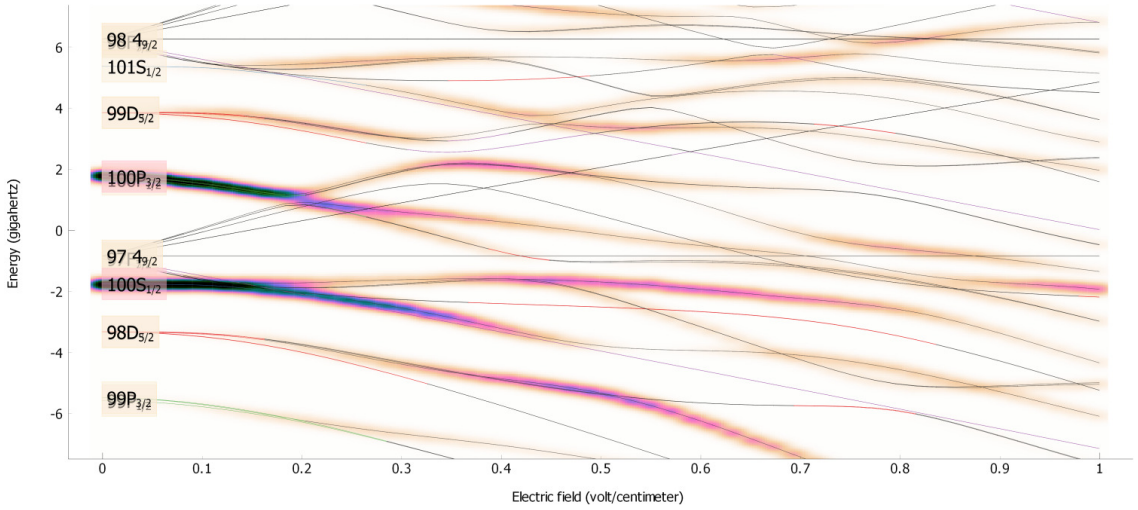
that cavity. However, this measurement precision is still more than sufficient. The 1 nm shift is negligible compared to the 780 nm length scale of the standing waves, so the lattice registry is excellent and should have little effect on the atom-cavity coupling.

Notably the $\pm$0.2 nm precision of this all-optical method is significantly better than the $\pm$30 nm precision of the mechanical SNOM-based approach in Ref. [135] and the $\pm$60 nm precision of the trapped-ion-based approach in Ref. [141]. More recently Ref. [142] demonstrated a complimentary all-optical approach targeted at much shorter cavity lengths. That method yielded a precision of $\pm$0.6 nm [9], which is comparable to the uncertainty of the results presented here.

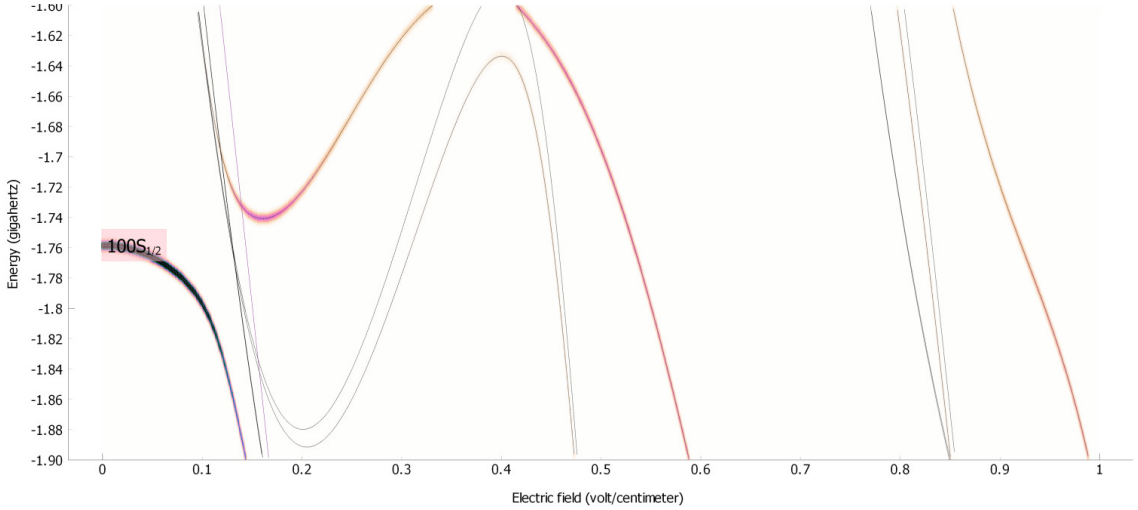
8.7 Rydberg Compatibility

Rydberg atoms barely hold on to their valence electron, which makes them very sensitive to electrostatic fields. Thus care must be taken to protect Rydberg atoms

[9] That work specifies that the $\pm$30 nm uncertainty from [135] corresponds to a mirror phase shift uncertainty of $\pm 25^\circ$. I've used that to convert their $\pm 0.5^\circ$ uncertainty from their all-optical method to a lattice shift uncertainty of $\pm$0.6 nm.



(a)



(b)

Figure 8-9: Stark shift plots of Rydberg state energy levels for rubidium as a function of electrostatic field, generated by Pairinteraction [145]. (a) The state energies on the GHz-scale. Significant state mixing occurs for fields $\gtrsim$0.3 V/cm, indicated by the fade in the shading around the $100S_{1/2}$ curve. (b) Same as in (a), except zoomed in vertically to show changes on the MHz-scale. Even fields of 0.1 V/cm are sufficient to cause shifts by $\sim$10 MHz.

from stray electrostatic fields when designing a Rydberg apparatus. This is demonstrated in Fig. 8-9, which shows the energy levels of rubidium Rydberg states around $100S_{1/2}$ as a function of electric field, as calculated by Pairinteraction [145]. To use the $100S_{1/2}$ state, the electric field likely needs to be reduced to a value small compared to 0.3 V/cm (30 V/m) to avoid effects from mixing in other states. Depending on the experiment, it may even be necessary to make the field small compared to 0.1 V/cm (10 V/m) in order to sufficiently reduce level shifts due to the Stark effect. This section discusses the design features related to containment and mitigation of electrostatic fields.

8.7.1 Piezos

One particular possible source of trouble in Rydberg cQED systems is the piezoelectric transducer (piezo) required to tune the cavity length. While Rydbergs are exquisitely sensitive to electrostatic fields, the piezoelectric effect is generally somewhat weak. For that reason, piezos and typically require large electric fields in order to drive sufficient changes to its dimensions. Typically the piezos are fairly close to the atoms (about 2 cm in our apparatus), so effort must be made to keep those large electric fields well-contained in order to protect the Rydbergs.

A test setup was developed to measure the electrostatic fields produced by the piezos used for the cavities (Noliac NAC2123-C06 with Kapton-insulated leads). The piezo under test was placed on a small metal pedestal inside a Faraday cage, as shown in Fig 8-10 and Fig 8-11. The Faraday cage was made from a die cast aluminum box lined with UHV foil to keep the piezos and their leads clean. The lid of the Faraday cage was made from a piece of cardboard wrapped in UHV foil with a circular hole cut in the center of it. The sensor of an electrostatic fieldmeter was fit snuggly in the hole. The electrostatic fieldmeter consisted of a Monroe Electronics 257D driver custom-configured to work with a model 1036F probe. The sensor is somewhat large and lacking in spatial resolution, however no better commercial options with sufficient sensitivity were found [10]. The sensor is capable of performing order-of-magnitude

[10] Many electrostatic fieldmeters are produced and sold with the intention of using them to hunt

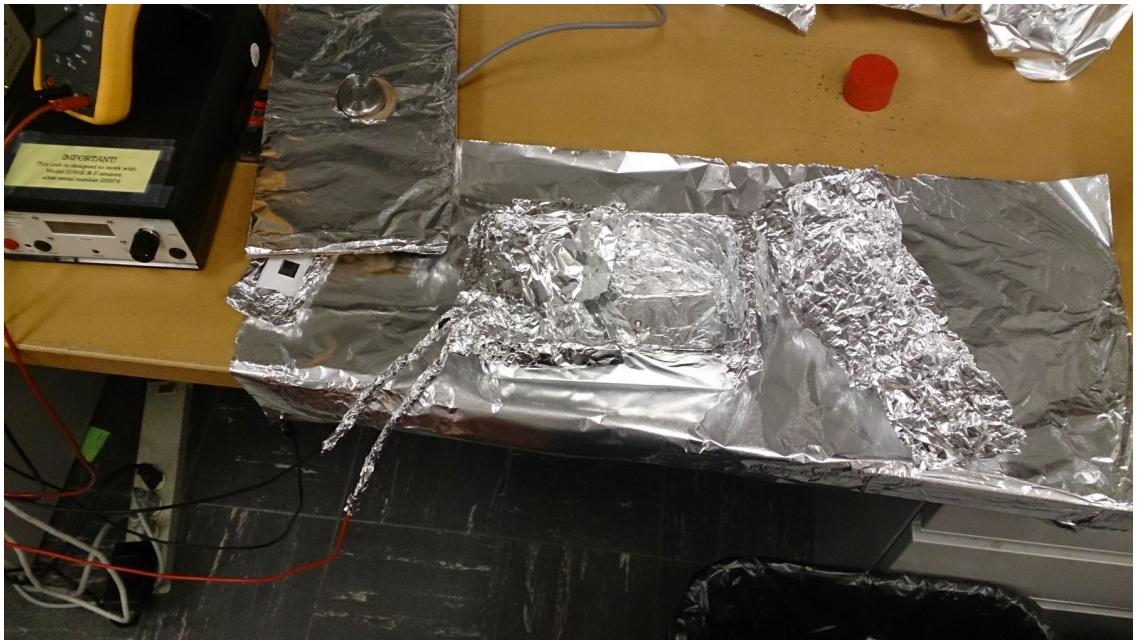


Figure 8-10: Photo of the Faraday cage test setup for measuring electrostatic fields produced by a piezo. The lid with the round hole and electrostatic fieldmeter sensor inserted is visible near the upper-left of the image.

measurements of the piezo's electric field a few centimeters from its surface though and thus is sufficient for our purposes. Notably the sensor is primarily sensitive to electric fields along its axis, but fields orthogonal to that direction are suppressed because the Faraday cage cancels parallel fields at its surface. The tests are performed by applying a voltage to the piezo and measuring the resulting change in electrostatic field.

The Noliac NAC2123-C06 piezos are constructed with a stack structure, made with many layers of interleaved electrodes with piezoelectric material in between. There are two contacts on the side of the piezo. Every other layer is connected to one contact while the others are connected to the other contact. One consequence of this geometry is that the stray field produced depends not only on the voltage applied to the piezo, but also on which electrode is on top. In one test, grounding the negative lead and applying +66 V to the positive lead produced a field of +35 V/m. Instead grounding the positive lead [11] and applying −66 V to the negative lead produced a

for ghosts. This makes finding reputable devices slightly more difficult.

[11] Piezos can handle larger voltages in the forward direction than in the reverse direction, so reverse-

Figure 8-11: Photo of the piezo mounted inside the Faraday cage. Aluminum foil is placed over the leads to shield them. The external soldering contacts on the sides of the piezo are not shielded.

field of -115 V/m. The difference in electric field amplitude is presumably due to the top electrode layer being connected to the negative lead.

The same test was also repeated after placing some UHV foil over the contacts on the sides of the piezo. In that case and the resulting fields were $+15$ V/m and -35 V/m for the positive and negative 66 V bias respectively. This clearly is an improvement, but could be better. It is worth noting that these piezos are rated to 200 V, and so fully biasing it would lead to even larger electric fields.

Following a suggestion from Boris Braverman, we decided to gold plate the piezos to further reduce the stray electric fields [90]. This was done at MIT's Microsystems Technology Laboratories (MTL) using an e-beam deposition machine. Kurt Broderick, a staff member at the MTL, suggested that a base layer of chrome be deposited on the piezos before applying the gold layer to improve the adhesion of the coatings. Following his advice, 50 nm of chrome then 100 nm of gold was deposited for each side of the piezo [12]. The piezo was mounted on a rotating platform angled at 45° to coat one flat face and the sides simultaneously. The piezo was then flipped and the procedure was repeated to coat both sides. To help reduce the number of possible locations for patch charges, some of the non-conductive parts of the cavity structure were coated with the piezos. Kapton tape was placed over any regions of those parts that needed to remain uncoated. Unfortunately the Kapton tape barrier approach did not work well for the piezos because the tape was too stiff to carefully outline the edges of its small contacts. Instead they were left uncovered and the leads were shorted by the coating. In an attempt to prevent that, a test run was perform in which a 1 μm thick layer of SiO_2 was applied to each side of the piezo before applying the chrome and gold. However this did not provide a sufficient insulating layer and the leads were still shorted by the coating. In the final coating run on the piezos used for the cavities, the SiO_2 layer was omitted. Instead a razor was used to scrape off

biasing them can lead to damage. Here the piezo remains forward-biased because the negative lead is brought to -66 V rather than the positive lead. The differential voltage between the electrodes is the same for the two measurements, it is just the common mode voltage that has changed.

[12] These values do not account for the projection angle of the piezos. Furthermore, the sides of the piezo were coated in both passes, i.e. before and after flipping the piezo, so the coating is thicker on the sides.

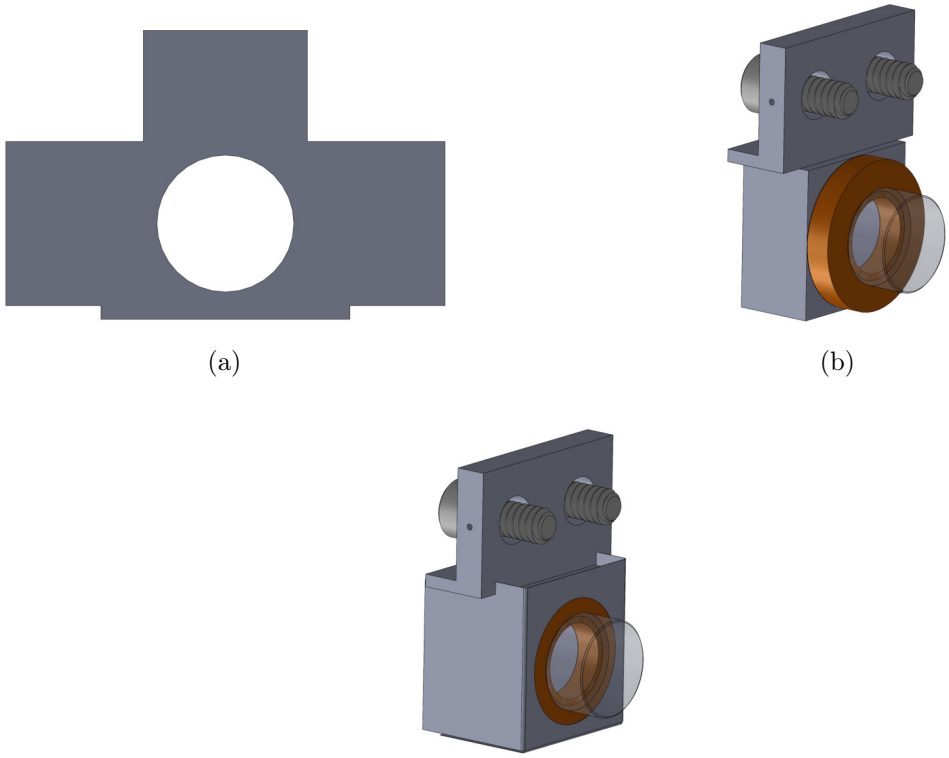
Piezo Coated	Contacts Shielded	E-Field	
		+66 V/GND	GND/−66 V
No	No	+35 V/m	−115 V/m
No	Yes	+15 V/m	−35 V/m
Yes	No	+20 V/m	−25 V/m
Yes	Yes	+3 V/m	−2 V/m

Table 8.6: Electrostatic field amplitude measured for various configurations of piezo shielding, as discussed in Section 8.7.1. Results are shown for two biases: one with +66 V on the positive lead and the negative lead grounded, then one with the positive lead grounded and −66 V on the negative lead. Results from repeated tests and tests with different piezos varied somewhat quantitatively, but they all agreed to within a factor of two and showed the same qualitative behavior. Placing shields over the piezo contacts reduces the electrostatic field amplitude, and so does coating the piezos in gold. Taking both measures significantly reduces the field amplitude.

the coating around the electrodes, which broke the short. After epoxying (EPO-TEK H27D) the piezos to their mount, the resistance between the different faces of the piezo and the mount were all ∼10 Ω.

The electrostatic field tests were repeated with the coating, and the results are summarized in Table 8.6. The coating clearly reduced the stray electrostatic field, providing a performance increase comparable to shielding the contacts of an uncoated electrode. When the coated piezo's contacts are also shielded, the electric field is particularly small and very near the noise floor of the electrostatic fieldmeter, leaving a field of only a few V/m. Those fields are likely sufficiently small as they are below the ∼10 V/m where the Stark shift of the rubidium $100S_{1/2}$ state enters the ∼10 MHz range, as mentioned in the introduction to this section. Therefore it was decided to include shields for the piezo in the final apparatus.

The design of the piezo shields is shown in Fig. 8-12 and Fig. 8-13. The shields are laser cut out of stainless steel foil in a procedure developed for the cavity length shims, which will be discussed in Section 8.9.1. In particular, stainless steel 316 is UHV-compatible and 1 mil thick foil made from that alloy is commercially available, so that was used to produce the shields. That material was found to produce shields that could survive the ultrasonic cleaner, wouldn't crinkle, could fold and maintain their shape, and were generally easy to work with. The shields are cut into a pattern



(a) (b)

(c)

Figure 8-12: The stainless steel foil piezo shields. (a) The flat pattern for the piezo foil shields. The shields are designed to be cut from a flat piece of stainless steel foil and can be made with a laser cutter. (b) A mirror mount with piezo before addition of the piezo shield. The piezo's leads are not shown. (c) A mirror mount with piezo and foil shield showing how the shield folds to wrap around the piezo.

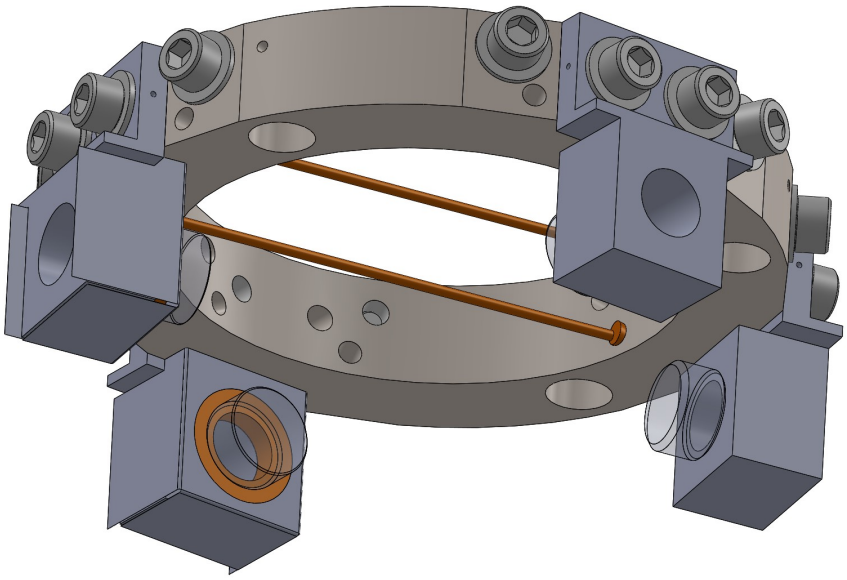


Figure 8-13: Cavity ring structure with piezo foil shields in place. The shields wrap around the top and sides of the mirror mount, and the extend to the surface of the ring structure. Kapton tape can be used as an insulation layer between the piezo contacts and the shield to avoid a short. The hole around the mirror has tolerance to allow for adjusting the position of the mirrors. The mirror holders without piezos have no need for these shields.

that can be wrapped around the piezo and mirror holder. This encloses the majority of the piezo and its contacts to provide significant shielding. The remaining exposed part of the piezo, as can be seen in Fig. 8-12(c) and Fig. 8-13, is shielded by the gold coating. Kapton tape can be applied to the piezo's contacts as a layer of insulation between them and the shield to avoid shorting the piezo. Care must be taken to route the leads to the piezos to avoid stray electrostatic fields from them. Twisting them together and running them around the outside of the ring structure should reduce their stray fields at the position of the atoms. Enclosing them in a grounded braid, such as Accu-Glass 110037 or 110200, would also significantly reduce their stray fields.

Further reduction of the stray electrostatic field due to the piezos is also possible using some other approaches. Firstly, as shown in the tests above, the direction of the electric field changes when the common mode voltage is changed. In other words, the field generated by the piezo can be tuned while keeping the piezo length fixed by increasing or decreasing the voltages applied to both electrodes while keeping the voltage difference across the piezo constant. By choosing the right common mode voltage, the electrostatic field produced by the piezos at the location of the atoms can be zeroed. The ratio of voltages applied to the piezo needed to null the field at the atoms should be independent of the voltage. Therefore there should exist a value a with $0 \leq a \leq 1$ such that the field is nulled when the potentials applied to the positive and negative electrode are aV and $-(1-a)V$ respectively. Here V is the desired voltage across the piezo, which sets its length. Thus, a piezo driver which applies those voltages, rather than grounding one of the piezo electrodes, would implement a feedforward cancellation of the stray electrostatic field produced by the piezo. Alternatively a feedforward compensation can be applied to the electrodes of the apparatus, which will be discussed in Section 8.7.2.

The stray field produced by the piezo is proportional to the voltage applied to it. Therefore another approach to reduce the stray field due to the piezo is to reduce the piezo voltage required to achieve the desired cavity length. This can be tuned by adjusting the temperature of the cavity ring. As the temperature of the ring changes, the cavity will expand or contract, thus tuning its length. By setting the

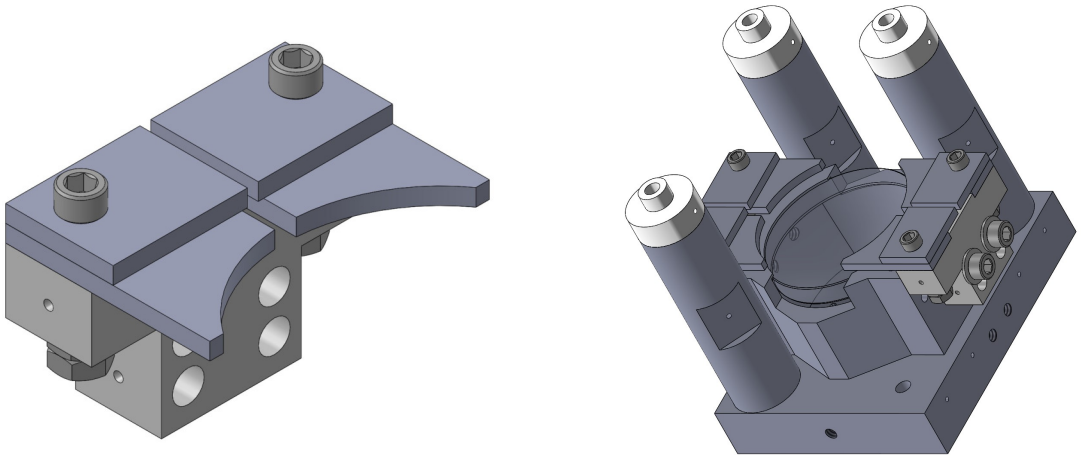
temperature to the appropriate value, the cavity length can be set to the desired value while the piezo voltage is zero. The light becomes resonant with the cavity each time the cavity's length changes by $\lambda/2$, so the cavity length only needs to be tuned by at most $\pm\lambda/4$. For 1560 nm light (the longest wavelength used with this cavity), that corresponds to a temperature change [13] of 1 K for a cavity $\approx$5 cm long with a thermal expansion coefficient of $\sim$10 ppm. That degree of temperature tuning is readily achievable.

The response time for tuning the cavity length by adjusting its temperature is very slow though, so the piezo will still be needed to suppress vibrations in the cavity length. However, the amplitude of vibrations is likely small compared to the low-frequency drifts in cavity length that could be cancelled via temperature stabilization. For that reason, stabilizing the cavity ring temperature to the appropriate value should significantly reduce the required piezo voltage and thus reduce the amplitude of its stray electrostatic field. Notably this technique is generally best for reducing the piezo voltage for one cavity at a time. If both cavities are needed simultaneously then they would likely need different ring temperatures to zero their piezo voltages. This may be possible to do to some extent using temperature gradients, though it is likely that sufficient gradients cannot be produced and further investigation would be required to determine the feasibility of that approach.

8.7.2 In-Vacuum Electrodes

Even if the stray electrostatic fields from the piezos are well-contained, the electric field at the atoms can still be too large due to other sources. In particular patch charges have been shown to be problematic, as has been seen for ion traps [146, 147, 148, 149] and optical lattice clocks [150]. Patch charges are regions of insulating materials where charge has accumulated, typically on their surface. Because the material is insulating, the charge does not dissipate or dissipates with an extremely

[13] Note that the goal here is to tune the cavity length to an (approximate) multiple of $\lambda/2$, not to any specific absolute length. That is in contrast to the discussion presented in Section 8.9.1 in which the goal is to tune the cavity length to a specific absolute length via temperature tuning.



(a) (b)

Figure 8-14: CAD images of the electrodes used to cancel electrostatic fields in the Rydberg cQED system. (a) A pair of electrodes on their Macor mount. Each consists of two stainless steel pieces, one which extends towards the atoms and one which covers most of the top of the Macor piece. The are secured to the Macor mount via screws and leads can be connected by sandwiching them between two nuts on the bottom of the screws. (b) The lower lens platform which mounts beneath the cavity ring. Two pairs of electrodes are mounted to the lens tube on opposite sides. Some components, including one of the support rods, are hidden for clarity.

long timescale. Patch charges can be created by various processes and can even be induced by light, particularly at short wavelengths [146, 147, 148, 149]. They can occur on any insulating component, including the cavity mirrors themselves [150]. To compensate for patch charges and other stray electrostatic fields, electrodes were added to the apparatus.

The electrodes are mounted to the lower lens tube (see Section 8.8) via insulating Macor ceramic pieces, as depicted in Fig. 8-14. This creates a circle of electrodes around the atoms, offset from the plane of the cavities by ≈15 mm. Applying voltages to the electrodes produces an electric field at the atoms which can be used to null any other electric fields.

Generally three electrodes are required to null the three components of the electric field vector, so four electrodes is sufficient for that purpose. Due to constraints from Maxwell's equations in free space, the electric field gradient has five degrees of

freedom [14] and so eight total electrodes would be needed to compensate both the electric field and its gradients. Mounting holes on the upper lens tube are included in the design, which could be used to include an additional set of four electrodes. The addition of those four electrodes would bring the total number to eight, which would be sufficient to cancel the electric field and its first order derivatives at the atoms. However the current plan is to omit the upper set of electrodes for simplicity, and they have not been manufactured. As a fallback, other in-vacuum electronic components could potentially be used as additional electrodes if necessary. For example, a voltage offset could be applied to the RF wires or microwave antenna. The common mode voltage of the piezos could also be adjusted to take advantage of their stray electrostatic fields, though those should be weak due to their shielding. One other option would be to add additional electrodes outside of the chamber, as was done in Ref. [150].

Although the electrodes make it possible to compensate for background electrostatic fields, it is generally better not to have those fields in the first place. This is particularly true if those fields drift over time, which would require periodic recalibration of the electrode voltages. Although light can create patch charges, it can also be used to remove them. Research has shown that shining UV light into an experiment can reduce patch charges [151, 150], presumably by ejecting electrons from negative patch charges or by ejecting electrons from metallic surfaces which can then land on positive patch charges to neutralize them. Periodically shining UV light into the vacuum chamber may provide a convenient method for minimizing patch charges. It may also be helpful to use a "static eliminator" during construction of an apparatus. These devices ionize air to create positive and negative charges. These charges then tend to deposit on patch charges with the opposite sign of charge, thus neutralizing them. Placing an apparatus near one of these devices before installing it into a vacuum chamber may reduce the initial amount and severity of patch charges.

[14] Each of the three electric field vector components can have a derivative in any of the three spatial directions, giving a total of nine derivatives. In free space, $\nabla \cdot \mathbf{E} = 0$, which is a scalar equation that provides one constraint between the derivatives. Similarly $\nabla \times \mathbf{E} = 0$ is a vector equation which provides three constraints (one for each vector component). The nine derivatives with four constraints leaves five degrees of freedom.

8.8 Imaging Systems

One of the features of the Rydberg cQED apparatus is a pair of imaging systems along the vertical axis, one on each side of the atoms. The design specs for the system have changed over time and will likely be different in the final apparatus. Because the design is not finalized, few quantitative results will be presented. Instead the qualitative features and overall reasoning for the design will be described. It is expected that design's structure will be the same in the final version, though some very specific details such as the lens material or asphere coefficients will likely be different.

The original goal was to include in-vacuum lenses with moderately large NA. There was no particular target NA in mind, but a larger NA was preferred. Furthermore the price target was set below the price point of professionally engineered in-vacuum commercial objectives. The atoms are also several centimeters from the surface of any viewport, which eliminates the possibility of using designs similar to many quantum gas microscopes. In those designs a large NA objective is placed just outside the vacuum chamber and the atoms are placed near the surface of the viewport. Therefore a cost-effective bespoke lens system design was developed based on using pairs of aspheres to perform 1-to-1 imaging between the atoms and points outside the vacuum chamber. Additional commercial optics and microscope objectives can then be used to further tailor the light, e.g. to add magnification. These components could be off-the-shelf because they do not need to be custom-designed for the vacuum chamber, thus significantly reducing their cost.

After some initial investigation it became clear that the best locations to add lenses were above and below the atoms along the vertical axis as those paths were unobstructed by the existing structure. The locations for the lenses are shown in Fig. 8-15. Due to some asymmetry with regards to how the diagonal MOT beams pass through the cavity ring and over some of the cavity mirror mounts, the lower lens could be made slightly larger and/or closer to the atoms than the upper lens. Those constraints make it possible to achieve an NA of $\approx$0.4 on the upper path and

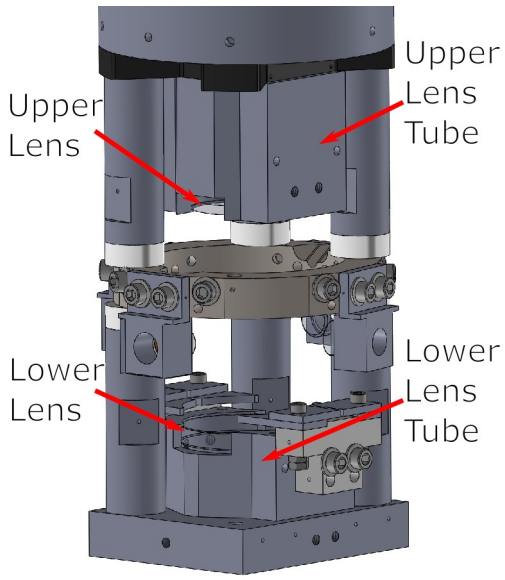


Figure 8-15: CAD rendering of a section of the Rydberg cQED apparatus showing the in-vacuum lenses and their lens tubes above and below the cavity ring structure. The lenses are rendered transparent, so structure behind them is visible. Some components, such as some of the support posts, are not displayed for clarity.

≈0.5 on the lower path. Interestingly for an NA of 0.5, the diffraction resolution limit $0.61\lambda/\text{NA}$ is ≈950 nm. That is not so far from the 780 nm lattice spacing of the trap light standing wave in the cavity. With that resolution atoms at adjacent lattice sites would not be well-resolved, but single holes in a lattice of atoms should be detectable. Therefore it should be possible to determine the occupation of individual lattice sites. Thus near-single-site resolution should be achievable, but it requires that the imaging be diffraction-limited.

Multi-element optics systems require more complicated mechanical designs and more careful tolerancing as there are more components which will not be perfectly positioned. To reduce complexity, a single-element design was selected with only one in-vacuum lens in each imaging path. Spherical singlets have significant spherical aberration for NA ≈ 0.5 and are not diffraction-limited, so aspheres were used instead. Aspheres generally work well on-axis, but their performance quickly falls off-axis. This was particularly evident when performing simulations to test using an in-vacuum asphere with an air-side asphere to simulate imaging light from the atoms onto a camera. When the air-side asphere was given a larger focal length than the in-vacuum

asphere (to provide magnification), the imaging worked well for a point source on the imaging axis. However, the simulated images included significant amounts of coma for atoms slightly displaced from the imaging axis. Coma is an odd-order aberration and thus can be significantly reduced by using a symmetric optical system, so a new approach was taken. Instead the same asphere design was used for both the in-vacuum lens and the air-side lens. One lens collimates the light, then the other focuses it back down, and by symmetry the odd-order aberrations are canceled out. This removed nearly all of the coma and significantly improved performance.

This setup also makes it possible to accommodate for some misalignment between the in-vacuum lens and the atoms in the cavity. The imaging system still produces more aberrations for atoms that are further off-axis and generally deteriorates for atoms that are more then a few tens of microns away from the center. That may seem problematic because positioning the in-vacuum lens relative to the atoms with a few tens of microns of precision would be a tall task. Fortunately that is not actually necessary. Although the in-vacuum lens should be aligned as well as possible, if it is off by a few hundred microns then the imaging quality can be mostly restored by adjusting the position and angle of the air-side asphere. Doing so adjusts where the low-aberration region of imaging is and can shift it by a few hundred microns. Put another way, for any given positioning of the two aspheres only a $\sim$30 μm $\times$ $\sim$30 μm square region will be in focus with low aberrations, but the center of that $\sim$30 μm $\times$ $\sim$30 μm square region can be shifted around by a few hundred microns by adjusting the position of the air-side lens. Some careful alignment of the in-vacuum lens may still be necessary, but the ability to adjust the position the air-side lens later on significantly improves the tolerance to in-vacuum lens positioning errors.

The two identical asphere setup provides a low-aberration 1-to-1 imaging path for light into and out of the chamber. Unfortunately this is generally not sufficient for imaging as magnification is usually required. Typically cameras have pixels that are a few or several microns across, so without magnification the finite pixel size would limit the system's resolution. However, additional optics can be added after the 1-to-1 imaging in order to provide magnification. At first glance it may seem then

that the 1-to-1 imaging with aspheres has not achieved anything, but it in fact has helped significantly. Because the atoms have been imaged to outside of the vacuum chamber, standard microscope objectives and other off-the-shelf components can be used to perform the magnification. That allows us to offload the difficult part of the optical system design to professionals, but without requiring a bespoke vacuum-compatible objective. That improves performance while keeping costs low. Basically the 1-to-1 imaging with aspheres takes the problem of imaging or generating arbitrary light patterns inside the vacuum chamber and reduces it to the problem of imaging or generating arbitrary light patterns outside of the vacuum chamber. In some sense the 1-to-1 imaging system creates a virtual copy of the atoms outside of the chamber where it is easier to create and modify optical setups. This makes the system very flexible as well. It is also worth pointing out that the imaging systems work both directions; they can collect light emitted from the atoms and guide it to outside of the chamber, or they can take light from outside of the chamber and guide it to the atoms.

One downside of using a single-element optics system inside the vacuum chamber is that it prohibits correcting for chromatic aberration. Typically chromatic aberration is corrected for by including two different elements made out of different glasses, but this is not an option when there is only one lens. For this reason, each of the two imaging systems is optimized for a particular wavelength. Throughout the design process the lower lens has been optimized for 780 nm light, as that wavelength will be used for imaging and for manipulating the internal state of atoms. Initially the upper lens with smaller NA was designed for 1064 nm light with the intent that it could be used to generate optical tweezers. That plan has since changed and those lenses will likely be optimized for 480 nm light instead, which is used for the Rydberg transition.

When the wavelength is adjusted from the design wavelength, the first issue that crops up for the imaging systems is that the lens focal lengths change slightly due to the change in index of refraction of their glass. This defocuses the light slightly and leads to larger spot sizes. That can be remedied by adjusting the position of the air-

side asphere to refocus, but then that messes up the alignment for beams at any other wavelength that follow the same path. A solution is to split the beam paths with a dichroic mirror. Each wavelength then uses the same in-vacuum asphere at the same fixed position, but each wavelength gets its own air-side asphere. Each air-side asphere can be positioned as needed for its beam's wavelength, independently of the air-side asphere for the other wavelengths. That makes it possible to have each beam focused properly simultaneously despite the chromatic focal shift. This approach does have some limits. When the wavelength is changed too far, the index of refraction changes enough such that choices for the lens radius of curvature and asphere coefficients are no longer appropriate, and aberrations begin to appear. However, this approach does greatly increase the spread of wavelengths which can be used simultaneously in one imaging path compared to a setup without a dichroic.

The asphere designs for the Rydberg cQED apparatus are 25 mm to 30 mm in diameter and have working distances that are in that same range. Lenses with custom dimensions are typically much more expensive than stock options due to the reduced cost of production at scale. However, aspheric lenses of this size are often made with CNC grinding machines rather than using molds or tools made for specific lens designs. Roughly speaking, to make a custom CNC lens design, the manufacturer simply has the type different numbers into the machine. There's no need to make new tools or molds. Therefore custom CNC aspheres can be bought at prices fairly comparable to stock options from some manufacturers. In particular Thorlabs was found to offer reasonable prices for our custom asphere designs. It is worth mentioning that the surface quality and accuracy of CNC grinders and polishers are excellent, but may not be sufficient for diffraction-limited performance. To achieve an even higher surface quality, magnetorheological finishing (MRF) is necessary, which is also offered by Thorlabs as well as other manufacturers.

Thus diffraction-limited performance at multiple nearby wavelengths is achievable with these custom aspheres. Performance at two vastly different wavelengths is achievable by designing one imaging path for each, with one positioned above the atoms and one positioned below the atoms. The system can be made at a fraction of the

cost of a bespoke professionally engineered system. The system is also flexible thanks to its ability to be significantly reconfigured by adjusting only air-side components after the 1-to-1 imaging. Future work on the system primarily involves finalizing the choice of design wavelengths and lens materials, then using design software to tune the asphere coefficients to the optimal values for those choices.

8.9 Other Design Aspects

In addition to the more significant design features described in the previous sections, the Rydberg cQED apparatus includes some other simpler features which are described in this section. Section 8.9.1 describes the design and construction of shims used to finely tune the cavity length, and Section 8.9.2 describes changes made to the vacuum chamber large viewports to improve optical access.

8.9.1 Cavity Length Shims

One of the original design goals for the Rydberg cQED system was to adjust the cavity length so that the FSR is approximately half the ≈6.834 GHz rubidium ground state hyperfine splitting. Doing so would make it possible to drive Raman transitions between the hyperfine manifolds with both legs of the Raman transition driven by cavity modes.

The actual target FSR was offset by a few megahertz from 3.417 GHz so that a magnetic field could be used to tune the Raman condition into and out of resonance. That makes it possible to turn off that resonance when it is undesirable. Setting the FSR to half of the hyperfine splitting rather than the full hyperfine splitting is important in order to drive the Raman transition. Neighboring cavity modes have opposite symmetry around the cavity center. Therefore the antinodes of one align with the nodes of the other near the center of the cavity where the atoms will be. That prohibits the atoms from maximally coupling to both modes simultaneously, and it is particularly a problem if the trapping lattice pins that atoms at the nodes of one of those cavity modes. When the next nearest cavity mode is used, then both

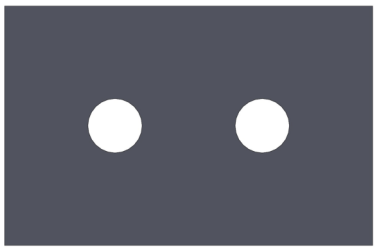


Figure 8-16: CAD rendering of the shim design for adjusting the cavity length. The shim can be placed between the cavity mirror holder and the cavity ring. Two #4 clearance holes are provided to allow through the screws used to attach the mirror holder to the cavity ring. The diameter of the clearance holes on the shim is smaller than the diameter of the clearance holes on the mirror holders. If the shims are well-positioned, then the edges of the clearance hole will not be pinched between the mirror holder and cavity ring. This is important because, as discussed in the text, the edges cut by the laser cutter are typically thicker than the rest of the shim.

modes have the same symmetry about the cavity center and their antinodes align near the center of the cavity. Thus, the cavity length should be set to make the FSR equal to approximately half the hyperfine splitting.

A cavity length of 43.83 mm was chosen as a target, which provides an FSR of (6.840/2) GHz. That sets the cavity modes Raman transition ≈6 MHz from the hyperfine splitting. A cavity length of 43.87 mm would put the Raman transition approximately on resonance, so the cavity length should be set with an accuracy better than 40 μm. That is comparable to the tolerances in the component dimensions, and the tolerances on several components add up to set the total error in cavity length. Therefore some additional effort is required to get the cavity length right.

A simple option would be to temperature-tune the cavity length, however reasonable temperature changes do not produce sufficient changes in cavity length. The coefficient of thermal expansion for stainless 316 steel is only ∼10 ppm/K, so a one degree Celsius change in temperature only changes the cavity length by ∼0.5 μm. Tuning the cavity length by 40 μm would require adjusting the temperature by ∼80 K, which is too large of a change to be practical. Similarly, the specified piezo travel range is only 3.3 μm which is not sufficient to tune the FSR significantly. A mechanical approach instead is necessary.

A simple mechanical approach to tune the cavity length is to insert shims between a cavity mirror mount and the cavity ring. Shims can only increase the cavity length, which is why the target cavity length was chosen to be slightly shorter than the length that would make the FSR equal to the ground state hyperfine splitting. At the time of the design, stainless 316 foil 1 mil (25.4 μm) thick and stainless 304 foil 0.5 mil (12.7 μm) thick were available from McMaster, and those thicknesses provide reasonable cavity length increments. Both of those alloys are UHV-compatible and thus are suitable material for shims. One difficulty is that cutting the shim stock to shape is nontrivial, particularly for the small geometry with internal holes shown in Fig. 8-16. Many different tools and techniques were tested for cutting the shims, including shears, scissors, razor blades, drills, punches and dies, paper hole punches, a laser cutter, and electroetching. Many tools tended to crumple the shims as they cut and the 0.5 mil stainless 304 foil was particularly prone to crumpling. Some of the hand tools also are not good for cutting internal holes, particularly small ones, and so they could not cut out the screw clearance holes. Many of the techniques also caused the edges of the shim to end up somewhat thicker than the bulk of the shim.

The only technique which produced thin edges and could produce the clearance holes was electroetching. To achieve the desired shape, a poor man's lithography procedure was implemented. Nail polish was applied to the shim stock, then it was ablated off using the laser cutter where the edges of the shim would be. The cutter's power and speed settings were set to ablate the nail polish without cutting the shim stock for this approach. The shim was then placed in a vinegar and salt solution and connected to the positive lead from a power supply. The negative lead was placed in the solution and 9 V was applied to the leads, which drove a few hundred milliamps of current. The exposed stainless steel then etched away, leaving behind the shim. The nail polish was then removed with acetone. The edges of the shim were somewhat tattered rather than perfectly straight, but they were very flat and had the same thickness as the rest of the shim. Unfortunately the electroetched shims had a distinct smell and therefore are unlikely suitable for UHV use. Other electroetching procedures with different chemicals may be able to produce UHV-compatible shims,

but this line of investigation was not continued further.

The only other tool that produced thin and flat edges without crumpling was plain office scissors, which worked better than shears. The edges were the same thickness as the bulk for shim stock cut with scissors, which was checked with a micrometer. The quality of the edges was also visible under a microscope. Unfortunately the scissors were not capable of cutting the small clearance holes necessary for the screws.

The only tool which gave reasonable results when cutting the screw clearance holes [15] was the laser cutter (Epilog Mini Model 8000 with 40 W CO_2 laser). However, the laser cutter melted the edges of the shim, which would cause material to ball up at the edges. The edges were then ≈3 mil thick for the 1 mil shim stock. The edge thickness could be reduced somewhat by sanding, but that tended to tear the shim eventually and may have affected the thickness of the center of the shims as well.

In the end a hybrid approach was taken. The shims were first cut out using the laser cutter. That conveniently produced the overall shape of the shim, including the clearance holes, but left thickened edges. The outer perimeter was then trimmed off with scissors to provide clean and thin outer edges. The clearance holes for the screws were cut with a somewhat small clearance; smaller than the clearance holes for the same screws on the mirror holder. As long as the shims are reasonably well-aligned to the mirror holder, the edges of the shim's clearance holes will not get pinched between the mirror holder and cavity ring with this design. In that case the thickness at the clearance hole edges does not matter; the mirror holder is pressed against only the thin and flat regions of the shim. Therefore this approach yields shims which are convenient to produce and have the screw holes courtesy of the laser cutter. However, they still have the thickness of shims cut only with scissors where it counts.

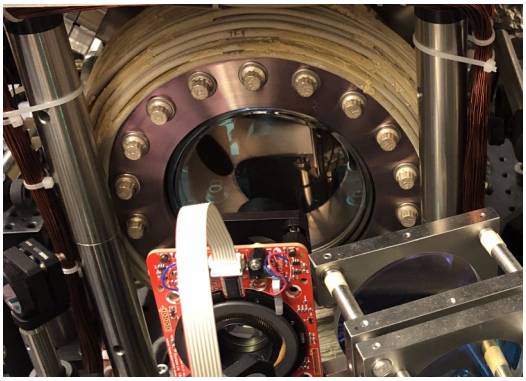


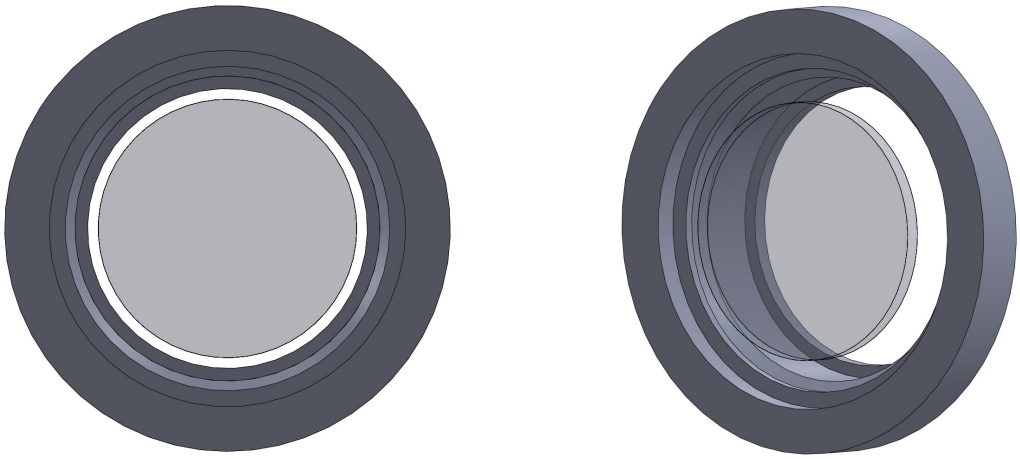
Figure 8-17: Photo of one of the large viewports on the vacuum chamber. The transparent fused silica piece does not extend as far outward radially as it could, which reduces optical access.

8.9.2 Improved Optical Access

Another goal of the Rydberg cQED apparatus upgrade was to improve optical access. To reduce costs and design effort it was decided to continue using the existing vacuum chamber in the lab, which placed some limits on how much the optical access could be improved. One possible place for improvement was that the large viewports on the side of the chamber had a somewhat thick outer metal ring, as shown in Fig. 8-17. That material restricted the optical access to $\approx \pm 24.5°$. In other words, the transparent part of the viewport did not extend radially outward as far as it could, which meant that any beam passing through the viewport and hitting the atoms would have to be incident on the viewport within 24.5° of perpendicular. It would be possible to squeeze in one more beam in the plane of the cavities if the optical access through the large viewports were not limited in this way. Therefore efforts were undertaken to improve optical access through the large viewports.

A straightforward way to improve the optical access would be to replace the viewport with another 6.75" CF viewport which has a larger fused silica window that extends further outward radially. However no commercially available options with a significantly larger window on a 6.75" CF flange were found. One company informed us that the fused silica windows were only available to them in fixed sizes, so custom

[15]Some preliminary tests using a drill also showed promise. The foil was sandwiched between two thicker pieces of steel, then the entire stack was drilled through. However that method requires further investigation to quantify how well it works.



(a) (b)

Figure 8-18: Simplified CAD renderings of the bucket windows with the old viewport's window overlaid for comparison seen (a) head-on and (b) at a skew angle. The transparent gray disk shows the where the window of the old viewport is, indicating its optical access. The white area within the circle of the bucket window shows where it will be possible to see with the bucket window but currently isn't possible to see with the old viewport. With the bucket window it will be possible to see through regions which were obscured by the old windows. This is particularly obvious in (b) where it is clear that the bucket window provides significantly more optical access at skew angles.

sizes were not feasible. Instead an alternative approach was taken. A window only slightly larger than the original was used, but the window was brought closer to the atoms. Moving a window closer to the atoms increases the maximum access angle for a given window diameter. The window could be brought in closer to the atoms without purchasing a new vacuum chamber through the use of a re-entrant viewport (aka "bucket window") as shown in Fig. 8-18, so this approach was taken. The 3.88" diameter window is replaced with a 4.20" window that is mounted 2.44" closer to the atoms by the re-entrant viewport. That brings the window in close enough so that its finite diameter does not limit the optical access. The window was not brought in further so that there would be space to mount a platform for rubidium dispensers, though that platform is now likely to be omitted from the final design. The bucket window improves the maximum optical access angle to $\approx \pm 32°$, a increase of $\approx 7.5°$ in each direction. The limit is again imposed by the outer metal ring of the flange with the knife edge and bolt clearance holes, though that ring's inner diameter is larger for the bucket window than for the old viewport.

The bucket windows have been ordered, but not yet installed. The plan is to install them at the same time as the Rydberg cQED experiment, as the vacuum chamber will need to be opened anyway.

Appendix A

Lattice Temperature Measurement

Measuring the temperature of atoms in a lattice is complicated by the harmonic oscillator ground state energy. The true thermodynamic temperature of the cloud is the value that appears in the Boltzmann factor which determines the relative populations of states with different energy. If a thermal atomic cloud has a temperature large compared to its trap frequency and its trap is suddenly turned off, then the cloud expands with a Gaussian velocity distribution with an RMS velocity set by $(1/2)k_{\mathrm{B}}T = (1/2)mv_{\mathrm{rms}}^2$ where m is the mass of the atom. The temperature can then be calculated by fitting a Gaussian to the atomic distribution after some time of flight and the temperature is simply proportional to v_{rms}^2. However, that is no longer the case when the temperature is small compared to the trap frequency. Even at zero temperature the cloud has nonzero v_{rms}. In particular, when an atom is cooled to the ground state of its lattice site (i.e. it has zero temperature) and the lattice site is then suddenly turned off, the atomic wave packet expands at a rate set by the kinetic energy of the harmonic trap ground state. When the temperature is small but nonzero, the dominant contribution of the kinetic energy during expansion is still the ground state kinetic energy. That makes it difficult to measure the temperature, as the signal can be swamped by the ground state energy.

More precisely, the mean kinetic energy of a cloud during expansion after suddenly turning off its harmonic trap can be calculated as follows. The energy of each state is $E_n = [n+(1/2)]\hbar\omega$ where ω is the angular trap frequency and n indexes the harmonic

trap level starting at $n = 0$. Of that energy, half is potential energy and half is kinetic energy. The population in each state can be calculated by summing the Boltzmann factors $\exp\left[-E_n/(k_{\mathrm{B}}T)\right]$ to determine the partition function Z

$$\begin{aligned} Z &= \frac{e^{\frac{-\hbar\omega}{2k_{\mathrm{B}}T}}}{1 - e^{\frac{-\hbar\omega}{k_{\mathrm{B}}T}}} \\ &= \frac{\sqrt{f}}{1-f} \end{aligned} \tag{A.1}$$

where $f = \exp\left(\frac{-\hbar\omega}{k_{\mathrm{B}}T}\right)$. The population in each state is then $\exp\left[-E_n/(k_{\mathrm{B}}T)\right]/Z$. In particular the population in the ground state p_0 is

$$\begin{aligned} p_0 &= \frac{e^{\frac{-\hbar\omega}{2k_{\mathrm{B}}T}}}{Z} \\ &= 1 - f \end{aligned} \tag{A.2}$$

which is a frequently-used result. The expected value for the kinetic energy can be calculated by averaging the kinetic energy over all of the states weighted by their population.

$$\begin{aligned} \langle \mathrm{KE} \rangle &= \frac{1}{Z} \sum_{n=0}^{\infty} \left(n + \frac{1}{2}\right) \frac{\hbar\omega}{2} e^{\frac{-\left(n+\frac{1}{2}\right)\hbar\omega}{k_{\mathrm{B}}T}} \\ &= \frac{\hbar\omega}{4} \left(\frac{1 + e^{\frac{-\hbar\omega}{k_{\mathrm{B}}T}}}{1 - e^{\frac{-\hbar\omega}{k_{\mathrm{B}}T}}} \right) \\ &= \frac{\hbar\omega}{4} \left(\frac{1+f}{1-f} \right) \end{aligned} \tag{A.3}$$

In the lab we often assign a sudden release "temperature" defined by $(1/2)k_{\mathrm{B}}T_{\mathrm{sud}} = \langle \mathrm{KE} \rangle$. This is essentially the temperature calculated when the effects of the ground state energy are neglected. Therefore the sudden release "temperature" T_{sud} is not actually the true thermodynamic temperature T but it approaches the true temperature in the limit that the temperature is large compared to the trap frequency. When the true thermodynamic temperature is zero, the sudden release "temperature" is $T_{\mathrm{sud}} = \hbar\omega/(2k_{\mathrm{B}})$. The relation between the two in general is plotted in Fig. A-1.

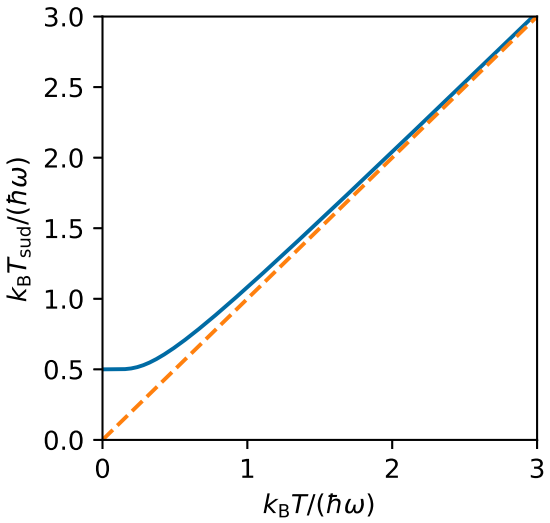


Figure A-1: The sudden release "temperature" T_{sud} as a function of the true thermodynamic temperature T. Both are normalized to the trap frequency. The minimum value of T_{sud} is $\hbar\omega/2$. A dashed line $T_{\text{sud}} = T$ is plotted as a guide to the eye, showing that $T_{\text{sud}} \approx T$ when $T \gg \hbar\omega$.

The above relations can be inverted to calculate f and the true thermodynamic temperature T as a function of the measurable quantities $\langle \text{KE} \rangle$ or T_{sud}. The value of f is

$$\begin{aligned} f &= \frac{4\langle \text{KE} \rangle - \hbar\omega}{4\langle \text{KE} \rangle + \hbar\omega} \\ &= \frac{2k_B T_{\text{sud}} - \hbar\omega}{2k_B T_{\text{sud}} + \hbar\omega} \end{aligned} \tag{A.4}$$

and the thermodynamic temperature is

$$\begin{aligned} k_B T &= \frac{\hbar\omega}{\ln\left[\frac{4\langle \text{KE} \rangle + \hbar\omega}{4\langle \text{KE} \rangle - \hbar\omega}\right]} \\ &= \frac{\hbar\omega}{\ln\left[\frac{2k_B T_{\text{sud}} + \hbar\omega}{2k_B T_{\text{sud}} - \hbar\omega}\right]} \end{aligned} \tag{A.5}$$

At small $T \lesssim \hbar\omega/2$, a small change in T_{sud} corresponds to a large change in T. This can be seen in Fig. A-1 and demonstrates the difficulty of measuring the true thermodynamic temperature when it is small compared to the trap frequency. A

method for improving the sensitivity of the measurement is to adiabatically ramp down the lattice instead of suddenly extinguishing it. The adiabatic release "temperature" can then be defined by $(1/2)k_{\mathrm{B}}T_{\mathrm{ad}} = \langle \mathrm{KE} \rangle$. This is the same as the sudden release "temperature", except that the kinetic energy is measured after turning off the lattice adiabatically. This approach is suggested and analyzed in Ref [152]. There the value of T_{ad} is calculated for a 1D system and for 3D system. The result for the 3D system is fairly algebraically complex, but the result for the 1D system is much simpler and gives approximately the same values. We often use the 1D system results, even the lattice is higher-dimensional, as this simplifies the algebra significantly and gives approximately correct results. The 1D adiabatic release temperature is

$$k_{\mathrm{B}}T_{\mathrm{ad}} = E_{\mathrm{r,trap}} \left(\frac{Q_0}{k_{\mathrm{trap}}} \right)^2 \frac{1 + 4f + f^2}{6(1-f)^2} \tag{A.6}$$

where $E_{\mathrm{r,trap}}$ is the recoil energy for a single trap photon, k_{trap} is the wavevector of a trap photon, and $Q_0 = 2\pi/a$ is the spatial frequency of the lattice, and a is the spacing between lattice sites. For two counterpropagating beams, the lattice spacing is $a = \lambda/2$ and so $Q_0/k_{\mathrm{trap}} = 2$.

In practice we've run into a few issues when attempting to measure the temperature via adiabatic release. The first few were technical. Small offsets in the intensity stabilization can lead to the trap beams staying on when their power is set to zero, or turning off earlier than expected when ramped towards zero. This is particularly difficult because the trap frequency scales with the square root of the trap beam powers, so to decrease the trap frequency by a factor of 10^3 requires reducing the trap beam intensity by a factor of 10^6. Doing that well would require zeroing the intensity stabilization system offset to one part in a million compared to its full scale amplitude. It is also important that any drifts in the offset are comparably small as well. In practice we achieved an offset stability and accuracy of roughly one part in 10^4 or possibly slightly better. Given that the initial trap frequencies are $\sim$100 kHz, a factor of 10^4 attenuation in power reduces the trap frequency to $\sim$1 kHz which is comparable to the recoil energy. Therefore 10^{-4} precision is marginally sufficient.

Another difficulty we ran into was associated with how the adiabatic ramps were generated. A uniformly adiabatic ramp would have an intensity ramp of the form [1] $I = I_0/(1+\Gamma t)^2$. In practice, due to the limited DAC resolution of our control system, we instead use a low-pass RC filter to generate a smooth exponential ramp. The exponential ramp is not uniformly adiabatic (the adiabaticity parameter $\epsilon = \dot{\omega}/\omega^2$ changes throughout the ramp) but can be adiabatic if its time scale is set slow enough. In particular, its time scale should be slow enough so that the ramp is still sufficiently adiabatic when the trap frequency drops down the the recoil energy. At that point tunneling starts to occur and the adiabaticity criteria changes. The control signals for the trap beam powers are connected to the setpoint inputs of the intensity stabilization system for each beam through an RC filter. Then by suddenly jumping the control signal to zero, the RC filter turns that step into a smooth exponential ramp. The intensity stabilization system then uses feedback to adjust the beam powers to follow that exponential ramp.

Although the ramp produced by an ideal RC filter is purely exponential, we found that the ramp produced by real RC filters becomes non-exponential after several decay times. This produced very measurable effects in the released cloud's properties as the trap powers were larger than expected after several decay times, leading to larger kinetic energies during the time of flight expansion. The unusual RC filter behavior was difficult to measure directly in the control signal. In particular the overdrive recovery time of many oscilloscopes makes it difficult to accurately measure the tail of the waveform where the signal amplitude has significantly reduced. The issue was that dielectric relaxation [153] in the capacitors caused small amounts of current to effectively flow through the capacitor with a decay time set by its dielectric's properties. That extra current developed a voltage drop across the resistor, changing the output voltage of the RC filter. A Tektronix TDS460A oscilloscope was found to have sufficiently fast overdrive recovery to measure the effects of dielectric relaxation

[1] This can be derived from the adiabaticity condition that $\dot{\omega}/\omega = -\epsilon\omega$ where $\epsilon \ll 1$ is required for adiabaticity. If ϵ is constant over time (i.e. the ramp is uniformly adiabatic), that can be symbolically integrated to give $\omega = \omega_0/(1+\Gamma t)$ where $\Gamma = \epsilon\omega_0$ and ω_0 is the trap frequency at the beginning of the ramp. The trap frequency scales in proportion to the square root of the trap beam powers, thus $I = I_0/(1+\Gamma t)^2$.

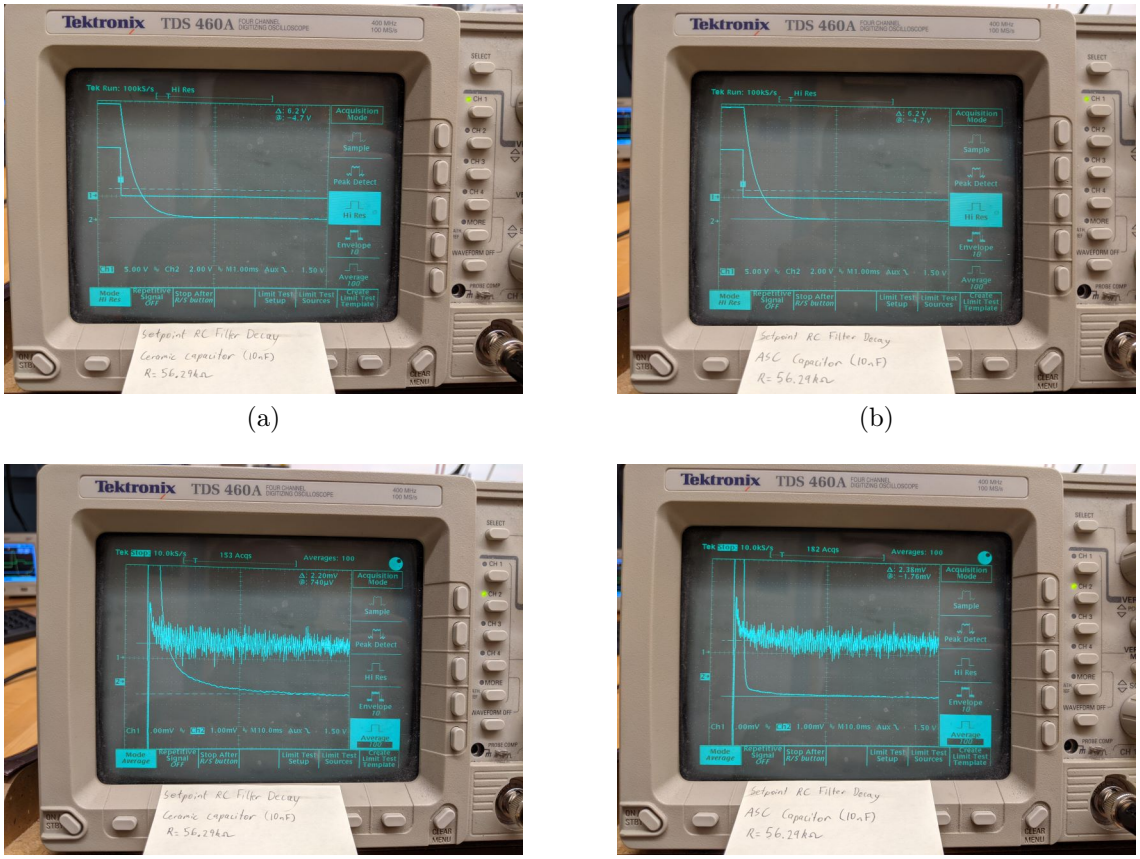


Figure A-2: Demonstration of effects due to dielectric relaxation in low-pass RC filter capacitors. (a) The full scale exponential decay when a step function is applied to the input of the RC filter. Both the input and output waveforms are displayed, and a marker is used to mark the asymptotic value. The horizontal timescale is 1 ms per division and the RC decay time constant is ≈0.5 ms. (b) Same as in (a), except a different capacitor with lower dielectric relaxation is used. At this scale the effects of dielectric relaxation cannot be seen so the waveform looks identical to that in (a). (c) Same as in (a) but zoomed in on the tail of the decay. The vertical scale has been decreased from 2 V per division to 1 mV per division and the horizontal timescale has been increased from 1 ms per division to 10 ms per division. There is clearly an exponential decay with a ∼10 ms timescale, which is much slower than the ≈0.5 ms RC time constant. (d) Same as in (c) but using the same capacitor as used for (b). This capacitor also shows some slow timescale decay, but with a much smaller amplitude than seen in (c).

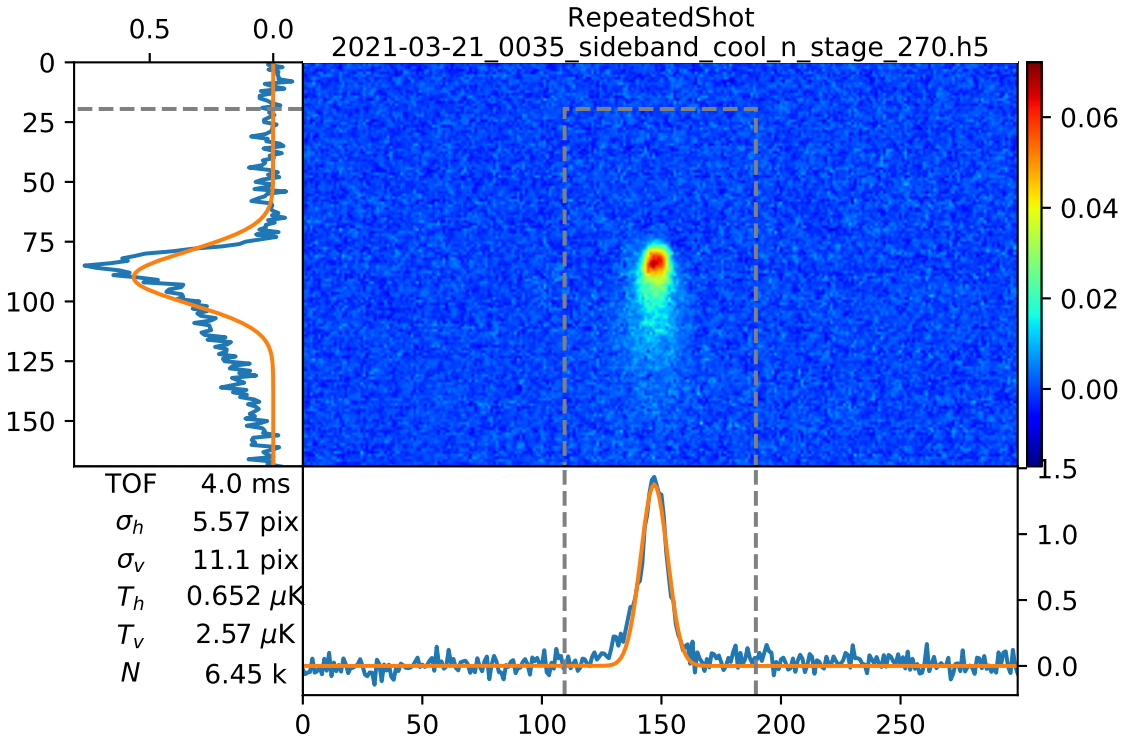


Figure A-3: Absorption image and integrated cross sections of a cloud after adiabatic release from a 3D lattice. A long tail of atoms along the vertical direction is visible. The tail is composed of atoms which escaped from the lattice before it was fully extinguished. This effect makes it more difficult to quantitatively assign a true thermodynamic temperature to the cloud. The adiabatic release was performed using an exponential ramp with a time constant of ≈0.5 ms for the beam powers.

in the tails of the RC decays. Alternatively the circuit suggested in Ref. [153] could be used to measure the behavior with an oscilloscope that has slower overdrive recovery. It was found that typical X7R ceramic capacitors exhibited this behavior more than other types. Fig. A-2 shows the differing performance for an X7R capacitor and a film capacitor. In the end NP0/C0G capacitors were used as they also showed little effects from dielectric relaxation and were readily available in our lab with various capacitances. This was useful as capacitors with different nominal values could be placed in parallel to make precisely-matched capacitances, which made it possible to match the RC decay times between the beams very precisely.

The last difficulty associated with performing adiabatic release temperature mea-

surements was the effect of gravity. The atoms are cooled by the adiabatic reduction in trap frequency only while they are in the lattice. If the lattice has a finite size then the atoms can escape it, at which point they will not be cooled by any further reduction in the trap frequency. Different atoms can also escape at different times, which then affects the final velocity distribution and spatial distribution. These effects are worsened by gravity, which adds a tilt to the potential that can drag atoms out of the trap earlier than they would otherwise escape. The effect of gravity is shown in Fig. A-3. These effects limit the extent to which the cloud can be adiabatically cooled during the release and complicate the data analysis. In particular it becomes difficult to calculate the actual thermodynamic temperature. The measured T_{ad} can still be used as a qualitative diagnostic though, as a colder cloud will likely be smaller after time of flight expansion and show less spilling as warmer atoms are more likely to escape earlier. The measured T_{ad} can also be used to calculate an upper bound on T, as the measure T_{ad} should be higher than the T_{ad} that would have been measured with a perfectly adiabatic release. It should be possible to cancel the effects of gravity by using a magnetic field gradient to apply a compensating force, though we have not done that as it would be difficult to do with our apparatus.

For these reasons we often used adiabatic release as a way to sensitively compare the performance between different sequences, particularly sequences with different trap frequencies. We then used careful sudden release measurements for more quantitative results.

Appendix B

Lab Wiki

I am extremely fortunate for having walked into a functioning lab rather than having to build one from scratch. For that, I am deeply indebted to my lab's previous members who put in years of work to make the machine what it is today. However, one of my biggest disappointments and frustrations in grad school was that my knowledge often built next to, rather than on top of, the knowledge of previous generations. The root cause of that difficulty was a lack of proper documentation practices and resources, and that lack had accumulated over a few decades. This appendix discusses some of the consequences of poor documentation practices and benefits of good documentation practices for an academic lab. A lab wiki is presented as a tool to solve many of those problems. A typical wiki page for our lab, the one for our homemade 780 nm reference laser mode stabilizer, is also included as an example.

B.1 Benefits of Documentation

In an academic lab, it is extraordinarily rare that a system is set up then never looked at again. Components often fail or need to be upgraded or duplicated. Even if none of those situations apply, sometimes a system needs to be examined when debugging an issue, just to eliminate it as a possible cause. Someone who is an expert on a system may be able to perform those tasks efficiently without documentation, but academic labs have a steady turnover of members. That means that all too often

the only expert on a system has long since gone. For the people who come later, often new and unexperienced students, proper documentation is critical for efficiently accomplishing whatever task they have for a system.

If the documentation about some system in the lab is missing, then any task involving it becomes much more laborious. This is a particularly frequent issue for our lab because we have historically built many devices that other labs buy commercially. That means that there is no commercial documentation to fall back on. At best, someone may have written some notes in their lab notebook. However that information is nearly inaccessible because over the years thousands of pages of handwritten lab notes have accumulated. That makes it impractical to look through all of it to find a section about the system in question. Even if a section about it is found, it may well be outdated and inaccurate. There's also a strong possibility that the desired information is not recorded anywhere, in which case any time spent looking for it would be sure to go to waste.

The difficulties of working without documentation are perhaps best demonstrated with an example. At some point during my work on the Rydberg cQED apparatus, which is the subject of Ch. 8, I needed to use an old homemade grating laser in the lab which I was not familiar with. It was set up with a homemade lockbox and homemade current controller as well. There is a folder full of schematics for various circuits in the lab, but often there is no way to figure out which schematic corresponds to which circuit. That meant that it was possible to find schematics that were probably similar to the circuits in the lockbox and current controller, but it was unlikely that the actual component values were the same. There was also no documentation on the laser itself. Furthermore, the labels on the some of the electronics and cabling didn't match, so it wasn't clear if the electronics were designed to work safely with that laser. For that reason I had to very carefully figure out various parameters like the gains of different inputs myself, rather than simply reading them off from a document. After a lot of time trying to figure out how things were set up, I managed to get the laser scanning, which was necessary for some tests on the Rydberg cQED apparatus cavities. I quickly found out that the scan range was too small for the cavity tests and had to

be increased. I then had to experiment with various scan rates and attempt to figure out its scanning bandwidth to see if the scan range increased at lower frequencies. After some time I found that it did increase for slower scans, but not by enough.

Looking at the current controller, I saw a pot labeled "gain" and decided to turn it up to increase the scan range. The rule that turning a knob right increases the parameter it controls doesn't always apply to our homemade circuits, so I tried turning it back and forth to increase the scan range. Unfortunately that pot didn't do what I expected; it only adjusted the feedforward current modulation applied as the piezo voltage is tuned to avoid mode hops. After toying with it, the laser began mode hopping more frequently as it scanned rather than scanning over a larger range. I was an inexperienced lab member at the time working on my own, so I did not understand what was happening. It wasn't until Vladan happened to stop by the lab and explain the role of that pot to me that I understood the issue. The pot did not have a readout dial so I could not simply return it to its previous setting. I then had to figure out a way to reoptimize the mode-hop-free scan range, then perform that procedure. I had no point of reference for what a typical mode-hop-free scan range was, so I had to play around with the settings until I convinced myself that the laser was not capable of doing any better. In the end the laser did not have scanning range or bandwidth adequate for my tests and I had to abandon it, but not until after having devoted a full day or more of work into figuring that out.

I'm sure that I was not the first person to do any of those tests on that laser, possibly not even the second. All of that work was duplicated effort. Each step in that process could have been skipped or would have been much faster if I knew what the person who made that laser knew. I also could have skipped the process entirely if the laser's scanning abilities, which could be summarized with just a few numbers, were recorded somewhere. Again, I am very lucky to have entered an existing lab and have greatly benefitted from the work of those before me. I do not aim to complain with this anecdote, but only to point out how the efficiency of labwork can be greatly enhanced by documentation. Often times just recording a few numbers where someone can find them can be done in a few minutes but save several hours

of work in the future. This is particularly true for academic labs where members are often new and inexperienced.

Experiences like the above are not uncommon in the lab and certainly are not unique to me. Many people in our group's labs, and presumably people from labs around the world, have faced similar difficulties. Without proper documentation, work often becomes tedious, redundant, and is more likely to fail to produce the desired results. On the other hand, when documentation exists and can be easily found, systems can be improved much faster. The documentation gives new lab members the opportunity to pick up on previously hard-learned lessons the easy way so that they can focus on new problems. Documenting new work is also much easier when previous documentation exists, as it can just be updated rather than written from scratch. That encourages people to continue to document their work. Even those who invest time writing the documentation can benefit from it. It can serve as a valuable resource for reminding oneself what they did and why, which can be helpful when taking a look at a system years after having worked on it. It also has the benefit that the writer can refer people to the documentation when they have questions, rather than having to explain a system to each person who asks about it individually.

B.2 Features of Good Documentation

Although any documentation is better than no documentation, there are many important features of a good lab documentation system. Some of those features are:

- Easy to share
- Easy to locate information about a given topic
- Well-organized information for a given topic
- Easily Editable
- Ability to embed images

- Ability to embed equations
- Ability to include attachments
- Ability to reference further resources

Often labs use a shared online notebook to achieve the above features, however we have found that a wiki has some benefits. Notebooks are usually organized by author and date. However, users typically want to look up information on a certain topic rather than look up what a particular person wrote on a particular day. A wiki is much better organized for that purpose as they are typically organized by subject rather than by author and date. Organizing by subject rather than by author also makes it easier for multiple people to contribute to documentation on the same subject. Of course online notebooks can also be organized by topic, but the design of a wiki intrinsically encourages users to organize information in that way.

In 2016, Boris Braverman and Jesse Amato-Grill set up a Confluence wiki for use by the MIT portion of the Center for Ultracold Atoms (CUA). In addition to the communal pages for various AMO topics, a section was added to the page tree for lab-specific information for individual labs. This turned out to be an excellent platform for our lab's documentation as it satisfied all of the bullet points listed above.

The wiki is accessible to all lab members, which makes it shareable. The pages are organized in a page tree which makes it easy to browse to the information on a particular topic or to just look around. Additionally a search feature is provided for convenience. Each page can host a table of contents for the information that it contains, which can be organized into sections with various heading levels and can contain collapsible sections. It is also possible to include a portion of the page tree in a given wiki page to conveniently link to its child pages, which is frequently helpful. For example, we often make a single wiki page for a given circuit design, then make an individual child wiki page for each device constructed from that design. In that case, a page tree excerpt included on the circuit design page provides a list of links to all devices that use that design.

Editing a wiki page is simple and done through the browser; no programming skills are required. That also makes it very easy to update information on a topic [1]. Furthermore, wiki pages can support embedding LaTeX equations as well as individual images or even galleries of images. Other files can be included as attachments, which is very convenient for keeping track of commercial test reports, datasheets, manuals, and so on. Published articles relevant to the topic can also be attached for convenience [2]. Wiki attachments are also great places to keep raw data from characterizations done in the lab. The wiki page is also a great place to include information which can be used to find even more information elsewhere. This can include links to other wiki pages or to external sites. It is also helpful to include part numbers so that more information about a device can be easily found online.

Even if a wiki page doesn't contain the information that a user needs, it can still be very helpful. For example, new students may not be familiar enough with a topic to know what keywords to search online to begin to learn about it. Having even a short blurb of text on the wiki page which provides some small amount of context provides new students with a jumping-off point. They can read it and then search online for more information about any terms or concepts that they are not familiar with. For example, a new student may have trouble with a laser mode hoping, but they may not know what is going on because they have not heard of mode hoping before. They could then go to the laser's wiki page to see a list of common issues with it and there they may find a short blurb about how the particular laser is prone to mode hops. They could then search online to figure out what a mode hop is and begin to understand their trouble with the laser. Without that prod in the right direction, the inexperienced student would likely struggle to understand the behavior of the laser. The wiki page, just by using the words "mode hop," can greatly assist the student in their efforts to understand their laser woes.

The wiki has other features as well. One example is that wiki pages have a basic

[1] This is a large advantage over recording information in theses and normal project writeups.

[2] This is one of the reasons that the wiki is not shared publicly. Ideally much of the information would be shared publicly for the benefit of all, but keeping wiki access restricted to CUA members avoids copyright issues. It also makes it possible to store some information that should not be public, although storing passwords on the wiki is still certainly *not* a good idea.

version tracking system. This can be used, for example, to restore a previous version after adding some erroneous changes. Another feature is that wiki pages can include excerpts from other wiki pages. This is sometimes more convenient than linking between pages and is far superior to copy/pasting content between pages. Copy/pasting information means that it needs to be updated in two locations whenever it is changed, which is easy to forget to do and is prone to error. The person making the update may not even know that there are multiple copies to change. When using excerpts it is only necessary to change the information in one location, then it is updated everywhere automatically.

Given all of the above features, the wiki constitutes a powerful documentation tool. The time saved by passing on information far outweighs the small effort of adding info to a wiki page. Its benefits are particularly helpful to inexperienced lab members, but more senior members can greatly benefit as well.

B.3 Sample Wiki Page: Mode Stabilizer

To provide a sense of what lab wiki pages are like, a sample wiki page is included below. It is the wiki page for a custom circuit designed to provide active feedback to our old homemade 780 nm reference grating laser in order to avoid mode hops. The formatting of the page is prettier when it is rendered with a web browser, but the reproduction below still includes the same material. Some small edits were made for the purpose of improving the formatting for this reproduction, which is why the "last edited" date is much later than the date that this circuit was last used. The links and attachments of course do not work in the reproduction below, but they do work on the actual wiki page. Furthermore, clicking on the images in the wiki page brings up a larger and higher-resolution version of the image for closer inspection, though that also does not work in the reproduction.

The example page includes a description of the problem that the circuit solves and explains how the circuit solves that problem. The final design, and the reasoning for the design, are included. All of this provides helpful context for future lab members

who otherwise may not know anything about the circuit. Photos of the device are also included to help lab members ensure that they are looking at the correct wiki page. The schematic and PCB layout are included as images for easy reference, and higher-resolution versions are attached which can be used to zoom in and look at details of particular sections. Justification for the choice of component values is included, which makes it easier to tailor future instances of the circuit for other lasers. Links for finding further information about the components are provided, as are links to wiki pages for related hardware in the lab. Characterization measurements are included as well. The procedures both for initially setting up the circuit and for using it on a daily basis are described. Common issues are mentioned, as are workarounds for those issues. A changelog is provided to keep track of resolved issues and other changes. Suggestions for improvements for future versions of the circuit are also provided.

All of the information in the wiki page already had to be determined in order to design and implement the circuit. Simply taking some time to record that information in a wiki page is a small overhead compared to the total time investment in the project. However, the time spent recording that information can pay off significantly. All of this information can be extremely helpful to someone encountering the circuit for the first time, and it is frequently even a helpful reference for people already familiar with the circuit as well. Simply recording this information, which had to be determined anyway, vastly reduces the amount of duplicated effort when a lab member must later work on the circuit. It also provides valuable information for anyone who would like to duplicate or improve upon the circuit. Over the last several years we have started to write many wiki pages covering various systems in the lab. It has become a powerful resource that significantly speeds learning, reduces duplicated effort, and improves productivity. Without further ado, the mode stabilizer wiki page is reproduced below:

Reference Laser Mode Stabilizer

Added by Zachary Michael Vendeiro, last edited by Zachary Michael Vendeiro on Jul 22, 2021 07:21

Introduction

Update: This circuit is no longer in use. It was built for our Rb Lab 780nm Homemade Grating Reference Laser, which has been replaced with Rb Lab 780 nm MOGLabs Reference Laser, which does not need this circuit.

Our homemade reference laser will tend to mode hop as the lab changes temperature and humidity over the course of the day, throwing it out of lock and messing up our sequence. This circuit provides feedback to keep the reference laser happy in its mode to prevent it from mode hopping and keep it in lock.

Working Principle

Alban pointed out that the reference laser's error signal gets noisy (noise in the ~50kHz range) when the laser gets near a mode hop. For a long time we used this as an indication that we should relock the reference before doing more runs of the experiment. We realized that we could use that noise as an error signal to determine how happy the laser is in its mode, and possibly feedback to adjust the piezo voltage or current to make the laser happier. However it wasn't clear how to implement this feedback with the laser's current controller. To do that we'd need to adjust the piezo voltage or diode current independently, but the current controller does not have a way to do that electronically. The modulation input adjusts both the current and piezo voltage, and it is already used to stabilize the reference laser's frequency. The piezo offset and current knobs are manual controls and cannot be adjusted electronically. Although the dither input on the current controller changes the current but not the piezo, it is AC-coupled so it cannot be used to tune the mode position. Given these limitations we could not figure out a way to perform the feedback for a long time.

At some point we thought of a way to non-intrusively adjust the laser's mode position electronically. The idea is that we can take the piezo output of the current controller and run it to another box instead of directly to the laser's piezo. That box could then add an offset to the piezo voltage to adjust the reference laser's mode position, and send the offset voltage to the laser's piezo. The piezo is much more robust than the laser diode, so it is relatively safe and non-intrusive to hijack its signal like this. By making sure that the new box has the same output impedance as the current controller's piezo output, we shouldn't change the frequency stabilization's feedback loop's loop gain, so the frequency lock should be mostly unaffected. Naively it would seem like adjusting the piezo voltage would mess up the laser's frequency. However if we adjust the piezo voltage slowly (compared to the bandwidth of the frequency stabilization feedback loop) while the laser is locked, the frequency stabilization will adjust both the piezo and current to keep the laser at the correct frequency. In the end, increasing the piezo offset with the mode stabilizer while the laser is locked might even decrease the actual piezo voltage, due to the way the frequency stabilization adjusts the piezo and current to keep the laser frequency fixed. In the end all that matters is that slowly adding an offset to the piezo voltage while the laser is locked adjusts the mode happiness without adjusting the laser's frequency. Since the laser goes out of lock on the timescale of hours, we don't need fast feedback for the mode stability and adjusting things slowly is no problem.

This circuit then works in the following way. It takes the signal from the balanced photodiode used for the frequency stabilization, then blocks its DC component and measures the RMS AC voltage using an AD637 integrated circuit. That RMS AC voltage is used as the discriminator and a setpoint RMS AC voltage is chosen with a potentiometer. The difference between the RMS AC voltage and the pot's setpoint voltage is used as the error signal and fed to a simple PI stage. The output of the PI stage is added to the voltage from the piezo output of the current controller, then sent to the reference laser's piezo.

Mode Feedback Setpoint

Choosing the Setpoint and Enabling the Stabilization

Choosing the correct setpoint for the feedback is important. If the setpoint is too low the laser won't be able to get its mode noise that low even when perfectly optimized, which will cause the mode stability feedback integrator to rail and throw the laser out of lock. Putting the setpoint too high will make the laser less happy than it could be, and seems to make the laser more prone to mode hopping when there is a physical perturbation, such as from working on the table.

To choose a good setpoint:

1. Lock the laser
2. Adjust the current knob on the current controller back and forth slightly and watch the error signal from the mode stabilizer box to see what it's maximum value is
3. Turn the current knob left from this optimum slightly until you see the error signal start to drop a bit
4. Now with the current slightly off from its maximum, adjust the setpoint knob on the mode stabilizer to make the error signal approximately zero. This makes sure that the desired setpoint is actually achievable
    a. I tend to lock the pot in place with its ten-turn dial. To unlock the dial, rotate the plastic nub on the side of it left slightly
5. Optionally lock the setpoint in place so that it doesn't get bumped or moved around by moving the plastic nub on the side of ten turn dial right slightly
6. Now you should have a good setpoint and you can turn on the mode stabilization feedback.
    a. Note: The minimum achievable mode noise depends on the which mode the laser is lasing in (i.e. how many nodes are in the standing wave between the grating and diode back facet). For simplicity it may be easy to just choose a setpoint that is reachable with the laser in its noisiest mode. Then you can just use that setpoint regardless of which mode the laser is running in, as it should be possible to get the mode noise that low regardless of which mode the laser is in.
    b. Note: Actually it seems that setting the desired noise level too high can make the laser more sensitive to perturbations such as bumping the table. It's best to set it to something somewhat near its minimum value given how happy the reference laser is in its mode that day.

Setpoint History

It may be helpful to keep track of where we put the setpoint, so we'll keep a log of it here

Date	Pot Setpoint	Corresponding Process Value Output	Notes
3/9/19	0.27	57 mV	Value used during initial setup and test of this circuit
3/9/19	0.40	104 mV	After turning the laser on/off, I had to move it to a new mode because the piezo knob on the current controller was all the way left. The laser was not as happy in this new mode, even at its optimal mode position. Therefore I had to increase the setpoint.
5/24/19	0.60	160 mV	After re-locking the laser, the mode noise at the optimal mode position was still too large for the system to reach its setpoint, so I had to increase the setpoint.
6/5/19	0.70	180 mV	The lab temperature is very high today (27.5° C near reference laser) and the reference laser isn't happy about it. The minimum mode noise is larger because of this so I had to increase the setpoint yet again.

Date	Pot Setpoint	Corresponding Process Value Output	Notes
6/11/19	0.40	112 mV	The air conditioner has been repaired after being broken for several weeks and the lab temperature is now stabilized at about 20°C or 21°C. Now the setpoint is much larger than necessary so I've put it back to 0.40 since that seemed to work well back when the lab was cold judging by the entries in this table from March.
7/10/19	0.70	173 mV	Tried turning off the reference laser over night, then turning it on in the morning with the rest of the machine to see if it would stay in lock with the mode stabilizer. After doing this, the minimum noise level was a bit higher in the afternoon, so I increased the setpoint.
September 2019	N/A	N/A	At this point we have a reasonable log of typical setpoints for this control. We have to change it semi-frequently as the lab environment changes, which in turn changes how happy the reference laser is. Now we'll start just adjusting daily as needed.
5 March 2020	1.70	57 mV	We still adjust the setpoint to whatever makes the error about +10 mV before turning on the mode stabilizer. However, We changed R_Setpoint_Limit_1 from 12 kΩ to 100 kΩ, so now the set point knob reads a much larger value for the same actual set point.

Design

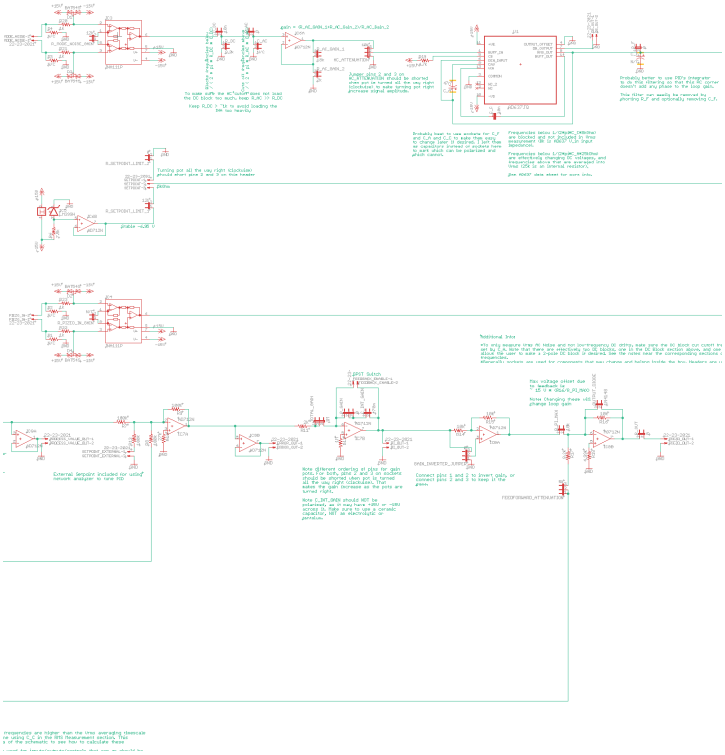
Schematic and PCB Layout

Eagle Files

- Project file: eagle.epf
- Schematic: RbLab_Reference_Mode_Stabilizer.sch
- Board: RbLab_Reference_Mode_Stabilizer.brd
- Gerbers for order from December 2018: RbLab_Reference_Mode_Stabilizer_Gerbers.zip

Schematic

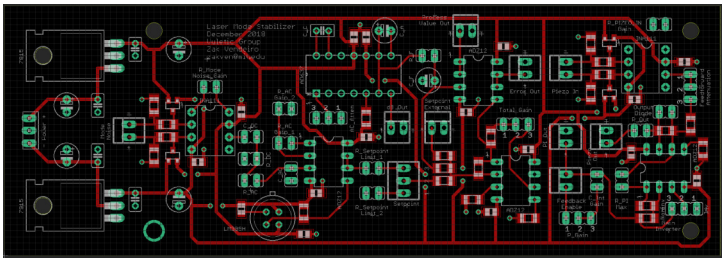
The schematic is kind of long, so I split the screenshot into two pieces. It may be more helpful to open the pdf: RbLab_Reference_Mode_Stabilizer_sch.pdf. Also, the purpose of each section is labeled on the schematic and those labels are visible in the pdf, but they are cut off in the images below.



Note that as drawn the Piezo Input to Piezo Output has gain -1, which is not good since this should have gain +1 there. To fix that I reversed the connections between the Piezo Input BNC and the header on the PCB. That works because the INA111 buffer stage lets both inputs float, so it just flips the sign on its output voltage.

Board

pdf: RbLab_Reference_Mode_Stabilizer_brd.pdf



Values of Adjustable Internal Components

Here we list the values of various components that are internal to the box and are adjustable, such as pots and components in sockets.

Pot Name	Nominal Max Value	Tuned Value
AC Attenuation	10 kΩ	Maximized, 0Ω (pin 1 to 2), 10.44 kΩ (pin 2 to 3)
Total Gain	1 MΩ	Maximized, 1.060 MΩ (pin 2 to 3)
P Gain	0 Ω	Shorted, 0 Ω (pin 1 to 2)
Feedforward Attenuation	5 kΩ	1.878 kΩ (pin 2 to 3)

And here's a table for resistors and capacitors in sockets:

Component Name	Value
R_Mode_Noise_Gain	11.94 kΩ
R_Plezo_In_Gain	N/C (∞ Ω)

Component Name	Value
R_Setpoint_Limit_1	100 kΩ (was 12 kΩ until 5 March 2020)
R_Setpoint_Limit_2	0 Ω
C_DC	10 nF
R_DC	8.2 kΩ
C_AC	N/C
R_AC	0 Ω
R_AC_Gain_1	18 kΩ
R_AC_Gain_2	2 kΩ
C_A	47 nF
C_C	10 nF
R_F	0 Ω
C_F	N/C
C_Int_Gain	47 nF (was 470 nF until 5 March 2020)
R_PI_Max	10 kΩ
R_Out	1 kΩ

Other Adjustable Components:

Component Name	Value
Gain_Inverter_Jumper	Inverting
Output_Diode	N/C

Box Drawing

Here is a drawing of the box that houses the circuit. Holes are labeled by their drill size. This drawing was done for a Hammond 1590D die cast aluminum box.

Reference_Laser_Mode_Stabilizer_Box_Drawing.pdf

AD637 References

When designing this circuit I often referred to a few documents about RMS to DC converters. Here they are for reference:

- The AD637 Datasheet: AD637_Data_Sheet.pdf
- Analog Devices Application Note "RMS-to-DC Converters Ease Measurement Tasks": RMS_to_DC_Converters_Ease_Measurement_Tasks.pdf
- Analog Devices "RMS to DC Conversion Application Guide" Appendix D: RMS_to_DC_Applicatiction_Guide_Appendix_D.pdf
    - This is just part of the appendix of the "RMS to DC Conversion Application Guide". There is a whole lot more to it, and it is available online for free. It's probably way more material than you'd want or need though.

Tuning

Feedforward Attenuation

The Feedforward Attenuation was tuned to give gain 1. This was done with the resistor R_Out shorted to avoid any effects from that, although its effects probably would have been minimal. I (Zak) did this by putting a sine wave into Piezo In and comparing the input with Piezo Out on a scope, then adjusted the Feedforward Attenuation pot until the signal overlapped well. You could do a better job by sending in a DC voltage and comparing the input to the output with our really nice Keithley 2000 multimeter, but this would probably be overkill.

DC Block and AC Cutoff

When I chose the values for this section I thought the mode noise was at around 20kHz, so I chose the DC block cutoff to be at about one tenth of that, i.e. 2kHz or so. I chose R_DC to be 8.2k so that is was about the same as the input impedance of the AD637's V_in input. That means that When C_DC and C_C are the same value, they have about the same cutoff frequency. That's not at all necessary, just convenient for comparing their corner frequencies. With C_DC set to 10nF, the DC cutoff is then 1.94 kHz. Similarly with C_C set to 10nF, that produces another DC block, this one with the cutoff at 1.99 kHz. The AC cutoff for high frequencies is not in use as of yet, so I just shorted its resistor socket and didn't add a capacitor.

AC Amplification and/or Attenuation

As discussed in the notes section below, the AD637 buffer starts to behave nonlinearly when its input exceeds 2.72 V peak-to-peak for some reason. The mode noise is roughly 20 mV peak-to-peak out of the photodiode when the laser mode happiness is optimized, so we want to choose the AC gain to make that signal appreciable without overdriving the AD637 buffer. With that in mind, I chose an overall gain of 50 or so to get about 1V peak-to-peak at the buffer. That should give enough overhead so that we can make the setpoint a bit higher than the minimum mode noise. To get this i put a 12kΩ as the gain resistor of the Mode Noise INA111, which makes its gain about five. Then I made the AC gain stage have a gain of 10 or so to get the remainder of the gain, and left the AC Attenuation pot maximized. This makes 50kHz well within the gain bandwidths of the INA111 and AD712 stages. Note that the INA111 gain is before the DC block, so if the DC level is too large it will rail. This shouldn't matter in practice though because the reference laser frequency lockbox will rail before that point because it has a gain 10 stage at its input, so if the INA111 in this circuit rails we've already had a problem with the frequency stabilization.

RMS Averaging

The value of C_A determines the cutoff between what frequency components are averaged into Vrms, and which are seen as changing DC levels. Since we only want to measure the change in amplitude of the AC noise, I decided to set the averaging corner frequency to something below the 2kHz cutoff of the DC block. Given that the noise at 50kHz dominates the signals at lower frequencies, this probably isn't strictly necessary but it doesn't hurt. With that in mind I chose C_A to be 47nF, which gives a 500 Hz cutoff frequency for the RMS averaging due to an internal 25 kΩ resistor of the AD637 (see its data sheet).

Normally when using the AD637 one would add R_F and C_F after the AD637's Vrms output to filter out its ripples. However, since we're using the AD637 in a feedback loop, this filter would add phase that could make the loop unstable if we aren't careful. Furthermore, the integrator of the PI stage should effectively do the job of filtering, since an integrator is a low pass filter, so these ripples shouldn't cause any issues in the feedback loop. To omit the R_F and C_F filter, I just shorted R_F and left the C_F socket empty.

PI Feedback Parameters

This circuit does not need large bandwidth at all, it just needs to be faster than the timescale over which the reference laser goes out of lock, which is ~ 1/(1 hr) ~ 10^-5 Hz. In fact it's good for this circuit to respond kind of slowly so that the frequency stabilization doesn't have to compensate its changes very quickly. With that in mind I (Zak) initially just set the Total Gain to ~1 MΩ and C_Int_Gain to 470 nF, which gives a 0.47 sec timescale to the PI response.

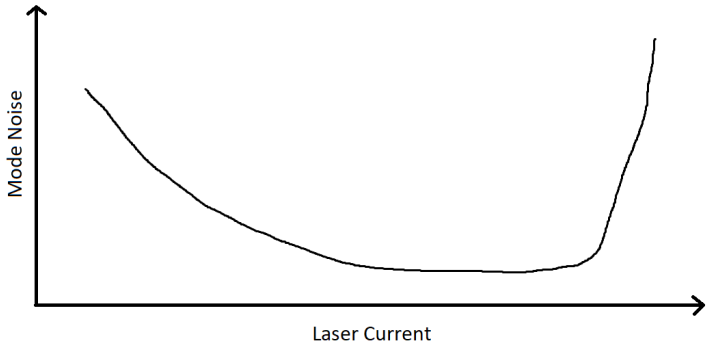
However, because the gain of the rest of the feedback loop is so small, that actually made the response time something more like ~ 30 seconds. That was a little annoying because if the setpoint was e.g. too low, then nothing would go wrong until ~1 minute after turning on the mode stabilizer at which point the reference laser would mode hop. Unfortunately, usually at that point we would have moved on to other things and wouldn't necessarily notice right away that the reference laser had gone out of lock. To make things happen a little bit faster, I (Zak) changed the C_Int_Gain capacitor from 470 nF to 47 nF.

Since we don't require fine tuning of the feedback loop, I did not adjust things beyond this. I didn't even add proportional gain, it's just a very slow integrator.

Gain Inversion

TLDR: The gain inversion section should be set to "Inv" to invert the signal to get the correct sign for the feedback.

The Mode Noise signal is asymmetric around its optimum. This can be seen by tuning the current knob on the current controller by hand and looking at the noise on the error signal (either the frequency feedback error signal's noise from the photodiode, or the ~DC process value output of the mode stabilizer). When the current knob is turned left from the optimum, the noise increases slowly and smoothly. However, when the knob is turned right the noise suddenly increases in magnitude quickly. This is depicted in the cartoon plot below. Since we're just trying to have some loose and slow feedback on the mode position, it's probably better to go to the side with the more gentle behavior and lock there. Note that the slope of the mode noise as a function of current has opposite signs on either side of the optimum, so we can pick which side we lock to by selecting the sign of the mode stabilization feedback gain.



To make sure we're locking to the more gentle side, I started by locking the laser (i.e. its frequency to the atomic transition) but leaving the mode stabilization feedback off. Then I adjusted the current by turning the current knob on the current controller left until there was a little bit more noise than at the optimum, then I adjusted the setpoint pot of the mode stabilizer to make the error signal zero without turning on the mode stabilization feedback. Doing this makes sure that the lock point is at a reasonable point above the minimum but without too much noise so that the laser is still mostly mode happy. I then turned the current knob even further left to increase the noise, then turned on the mode stabilization feedback. When the feedback has the correct sign of gain, the mode noise will decrease when the stabilization is turned on, otherwise it will increase. The response is pretty slow due to the way we tuned the PI to be very slow. Doing this showed that I had to set the gain inversion to "Inv" to invert it, i.e. I had to short the two pins on the gain inversion header nearest where it says "Inv" on the silk screen.

R_Out

The value of R_Out was chosen to match the output impedance of the piezo output of the reference laser's current controller Rb Lab 780nm Reference Laser Current Controller, which is about 1 kΩ. This was done so as to not affect the frequency stabilization feedback loop when inserting this circuit into the system.

Issues

This section lists some properties of this circuit that are not ideal.

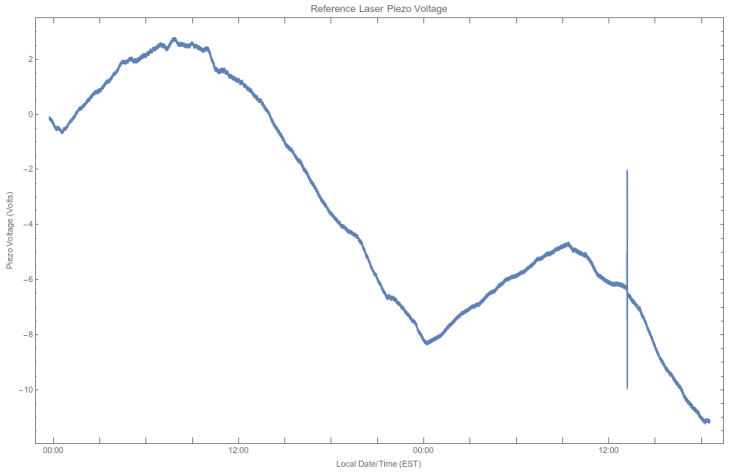
Piezo Output Railing

One issue with the circuit is that the Piezo Out voltage can never exceed ~±15V or so (probably closer to ±12V for an AD712 op amp). That means that if the piezo voltage required to keep the laser happy exceeds that value, the output will rail and the mode stabilization will no longer work. In practice it turns out that the required piezo voltage can drift a lot over time.

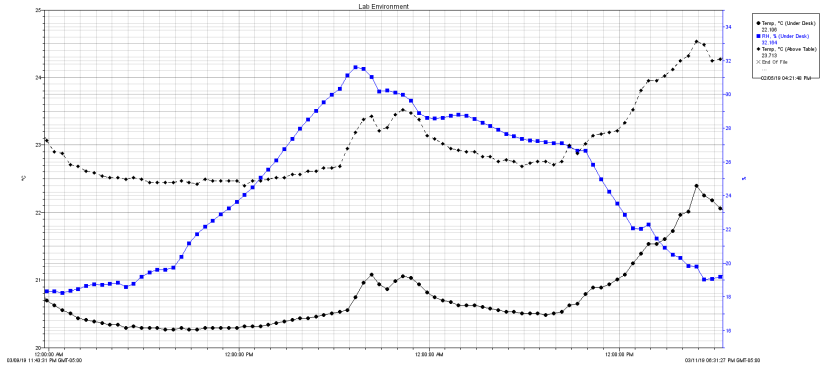
Below we have some logs of the reference laser piezo voltage and the lab environment taken with the mode stabilizer on. The piezo voltage was logged using a python script to log the readings from a Keithley 2000 multimeter connected in parallel with the Reference laser's piezo.

Log from March 10th 2019

Below is a log of the piezo voltage taken over a weekend with the laser locked and the mode stabilization on. Below that is a log of the lab's temperature and humidity taken using the Hoboware logger. A clear large drift towards negative voltages is apparent, which likely would have caused the Piezo Out voltage to rail not long after this log was stopped. The spike is due to someone re-locking the laser even though it had NOT gone out of lock. Interestingly the peaks in temperature seem to be reasonably well aligned with the minima in piezo voltage, and the humidity seems to be pretty uncorrelated. Keep in mind that in this data set, the temperature log records the air temperature below the control desk and above the optics table near the control desk, so the temperature around the reference laser in that box may be a bit different.



(Data from this file: Keithley2000Log_1552189410.csv, the first column is the time in Unix time and the second column is the piezo voltage in volts)



(Subset of data from this log file: Experiment_20190403.hobo, the two bumps are from turning the machine on)

Log from 1 May 2019

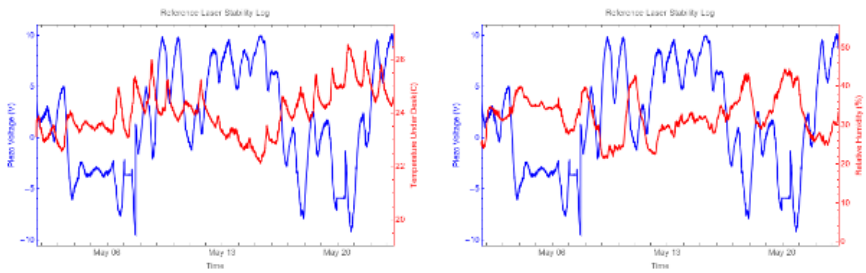
We bought another temperature sensor for the Hoboware data logger and mounted it to measure the air temperature right near the reference laser, so we took another log. Below are three plots, all plotting the piezo voltage as a function of time, but one is overlaid with the temperature near the reference laser, one is overlaid with the humidity (measured under the control computer desk), and one is overlaid with the temperature measured below the control computer desk. The reference laser piezo voltage is definitely well correlated with the temperature of the air near it. It's not a unique 1-to-1 mapping from temperature to piezo voltage though, as there are multiple times with the same temperature that have differing piezo voltages. The humidity seems correlated as well, but to a lesser degree. It may just be that the humidity is correlated with temperature, which in turn is correlated with piezo voltage.

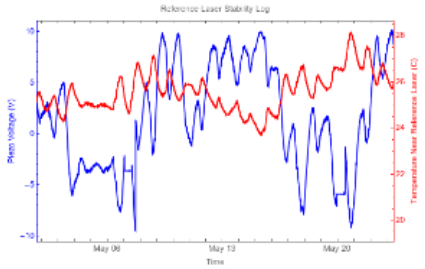
Here is the piezo voltage log: Keithley2000Log_1556744758.csv

Here is the exported temperature/humidity log data: Experiment_Environment_20190501-20190607.txt

Here is a Mathematica notebook to read in and plot the data: Reference_Laser_Log_Plotting.nb

Log from 1 May 2019





Log from 12 June 2019

For the previous logs, the lab temperature stability was very bad. We found out this was because one of the two AC units in the ceiling above the lab stopped working because its fan tripped a breaker and the lab temperature control couldn't reach its setpoint with the one remaining weaker AC unit (see Rb Lab Lab Temperature and Humidity for more details). After this repair the temperature does not vary nearly as much over the course of the day, so we've taken another log to see how things behave when the temperature is more stable. The jumps and flat regions in the piezo voltage log are due to it going out of lock, us re-locking the laser, and the multimeter going out of range for a bit (it was at a fixed range, so I changed it to autorange).

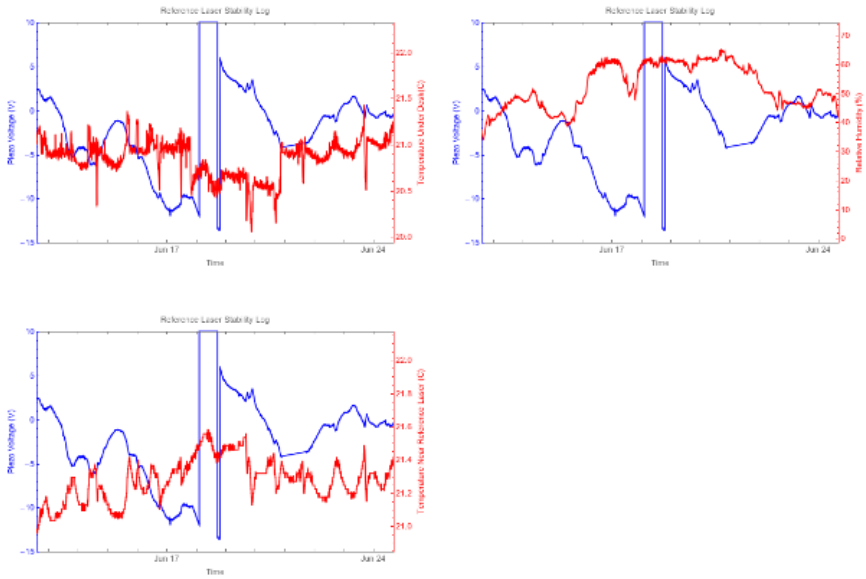
As can be seen below, the temperature near the reference laser is much more stable than before, staying between 21.0°C and 21.6°C for the duration of the log. In fact, the temperature variations are small enough for us to resolve the discrete steps in the Hoboware logger's ADC. Despite the temperature stability, the piezo voltage still moves around quite a lot. Furthermore, the piezo voltage is well anti-correlated with the relative humidity, which does vary by ~30% over the course of the log. Given that piezos are known to interact badly with humidity (it shortens their lifetimes, especially when operated at higher voltages), it's not crazy to think that the drifts are due to the humidity interacting with the piezo material. In fact, it's likely that this is the major source of drifts. We should perform another log like this in the winter when the humidity is more stable (it's always very low so it likely won't vary much) and see if the piezo voltage drifts less.

Here are the piezo voltage logs (windows update restarted logging computer, so we had to restart the logging): Keithley2000Log_1560362875.csv and Keithley2000Log_1561152752.csv

Here is the exported temperature/humidity log data: Experiment_20190612-20190708.txt

Here is a Mathematica notebook to read in and plot the data: Reference_Laser_Log_Plotting.nb

Log from 12 June 2019



Conclusion

We'll probably just have to live with this. If it's really an issue, I can only think of two ideas that may solve it and neither seems very easy or particularly promising.

- One idea would simply be to put a temperature-controlled and/or humidity-controlled box around the reference laser, or temperature control the breadboard beneath the laser. That should help if these drifts are due to temperature or humidity drifts, but won't help if they're due to something else (e.g. a slow sagging of the grating due to glue creep or a gradual tilting of its mirror mount). Given that the required piezo voltage moves up and down, I don't think the issue is creep or sagging. Furthermore, there is a clear oscillation period of 1-day in the log, which makes it likely that it is temperature.
- Another option would be to drive the piezo with a high-voltage amplifier to increase the range over which it can be scanned. However, there are some things to consider before doing that
    - That would change the piezo gain for the frequency stabilization and it would change the ratio of piezo voltage per diode current. To accommodate this, we'd either have to (a) decrease the gain on the feedforward path of the mode stabilizer to cancel the gain of the HV amplifier, or (b) re-tune the frequency stabilization lockbox and current controller. Option (a) is much more straightforward and may be worth doing.
    - The frequency stabilization control signal from the lockbox will still be limited, and if it rails, then the HV amplifier on the piezo won't help. That said, I suspect that what happens over the day is that the current doesn't change a whole lot because it is relatively stable, while the piezo voltage drifts a lot to cancel out large drifts in the length of the lasing cavity. If this is the case, then the output of the frequency stabilization lockbox shouldn't change much even when the mode stabilizer drifts very far. This should be checked by logging the frequency stabilization lockbox output voltage though before committing to building an HV amplifier.
    - Furthermore, I'm not sure exactly what the max voltage of the piezo is. It looks like it may be a AE0505D08F from Thor labs, but I'm not sure. Also, it seems to be hooked up backwards to the BNC, in which case the high-voltage amplifier would need to go to large **negative** voltages, not positive.
    - Also note that our usual HV amplifier design is inverting, so we'd need to flip some gains. We can use the gain inverter in the mode stabilizer to flip the gain on the mode stabilization feedback path. To flip the gain on the frequency stabilization path, we could reverse the connections to the INA111 input buffer on the piezo feedforward path again.

Monitor Signal Output Oscillations

At one point while looking at the "Process Value Out" and "Error Out" signals on an oscilloscope, we noticed that they were oscillating wildly with ~MHz frequency. The laser was still in lock and the mode stabilization still seemed to be working while this was happening, so it seemed like the AD712 buffers for those output were oscillating but everything in the feedback loop was working fine. Furthermore, disconnecting the BNCs from those monitor outputs then reconnecting them stopped the oscillations, again without causing any trouble for the frequency or mode stabilization, further supporting the idea that it is only an issue with the buffers. Interestingly, disconnecting the BNC cables at the other end (i.e. the ends connected to the scope) did not stop the oscillations. It's possible that the capacitance of the BNC cables causes the AD712 buffers to oscillate. If this becomes a frequent issue, it may be worth it to just try a different op amp, at least for the buffer outputs (IC9 in the circuit). If that doesn't work, putting a resistor in series with the BNC outputs might kill these oscillations, at the cost of increasing the output impedance.

Update 9 April 2019: I swapped out IC9 to replace the AD712 there with an OPA2227. So far it seems stable, but we'll keep an eye on it. Another good replacement candidate would be an AD827 which is supposed to be stable for all capacitive loads. It's a very nice and expensive (~$15) op amp though so its kind of overkill for a buffer.

Notes

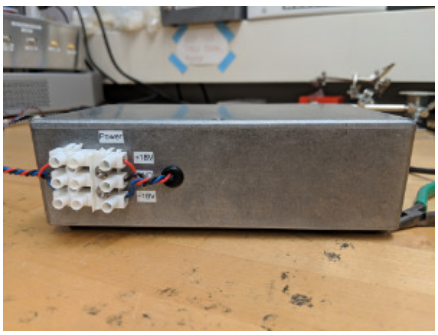
- If the input to the AD637's buffer (pin 1) exceeds 2.72 V peak-to-peak (i.e. 1.36 V amplitude, 0.962 V rms), then the buffer's output begins to distort. The distortion is assymetric and is worse on the negative side. That messes up the RMS voltage measurement slightly, and makes it increase less than it should when the signal is increased beyond that point. The response of the rms measurement section still increases monotonically with input amplitude though, it's just not linear beyond that point. The AD637 RMS output is 930mV at this point, i.e. when the input to it is 2.72 V peak-to-peak. Note that the 2.72 V is measured right at pin 1 on the AD637, so the Mode Noise input to the box will be different due to the gains of the stages in between.
    - This distortion happens at a lower voltage than I expected. I suspect that the issue is that it can't drive the V_In input of the AD637 well because it has an 8 kΩ input impedance (for signal frequencies which see C_C approximately as a short). I confirmed this by temporarily removing C_C from its socket so that the buffer wouldn't see the 8 kΩ load, and then the buffer's output was no longer distorted.
    - From the spec sheet, the buffer output current min and max currents are -0.13 mA and +5 mA respectively. For negative signals, (-0.13 mA) * (8 kΩ) = -1.04 V, which is somewhat close to the 1.36 V where we start to see distortion. However, we see distortion on the positive side too where 5mA should be plenty. Also, -0.13 mA seems very small so I'm suspicious that I'm interpreting this wrong. However, the signal going into the buffer looks fine even when the buffer's output is distorted, so the issue seems to be with the AD637's buffer. It's possible that this is just a defect of this specific AD637 and other AD637's may work fine, I haven't tried any other ones.
- When the Piezo In and Piezo Out are hooked up as drawn in eagle, the pizeo voltage is inverted. To fix this, I reversed the connections between the Piezo In BNC jack and the INA111 buffer. I actually swapped the crimp pins' positions in the header socket connector thing, so the connector goes on the correct way (in other words put the connector on normally, do NOT try to force it on backwards)
- Maybe a temperature controlled box around the reference laser would limit how far the laser drifts around. That would help alleviate the issue that the piezo output can rail

Photos

Box Photos

Here are some photos of the box:

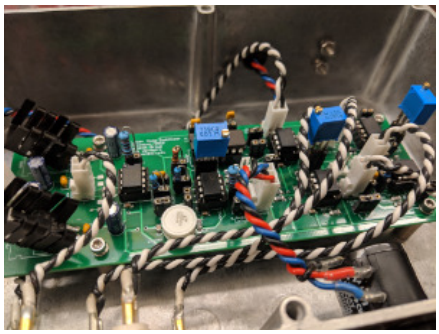
Box Photos

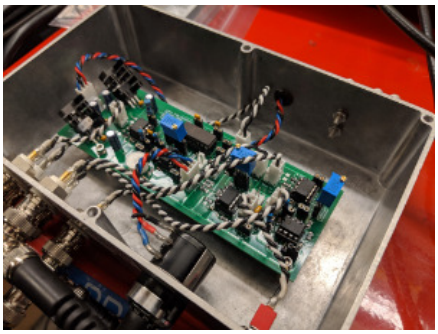


PCB Photos

Here are some photos taken with the box open so that the PCB is visible:

PCB Photos





Other Possible Upgrades

This section is intended to list ideas that could be implemented in future versions of this circuit, possibly with new versions of the PCB.

- Any upgrades you can think of that would help with issues from the "Issues" section above. Some suggestions are included in those sections

- Ditch the AD637's buffer and use an op amp buffer instead
    - This is to fix the fact that the AD637's buffer seems too weak to drive its own input to voltages over ≈1.36V (see the "AC Amplification and/or Attenuation" subsection of the "Tuning" section above)
    - To leave the buffer unused, ground it's input (Pin 1) and leave its output floating (Pin 14).
- The Mode Noise amplitude is proportional to the intensity of the light from the reference laser. This is annoying because if the laser's output power increases, then the minimum mode noise (i.e. when the laser is happiest) can end up being higher than the setpoint causing the laser to go out of lock.
    - One way to fix this would be to add a photodiode to the reference laser's path to measure its power, then run that signal to the Mode Stabilizer circuit and use a divider IC to divide the Mode Noise by the laser's DC power. That would require redesigning the circuit of course, but would make the mode happiness error signal power independent.
        - Actually this probably won't help, the mode noise seems to be more a function of which mode you're in rather than the power of the laser. And once you're in one mode, I don't think the power changes very much so I don't think this is a large issue with regards to that. It's more an issue that the ideal setpoint is very different for different lasing modes.
    - Another option would be to do something more clever with how we get the error signal so that we actually lock to where the mode noise is minimized. This would be kind of like doing a PDH lock to a cavity rather than locking to the side of its transmission. Doing this for the mode happiness would probably be difficult. One way would be to sweep the piezo offset around slowly, then mix that Mode Noise with the LO of that slow sweep (this would be too slow to use a normal mixer, we'd probably have to do it with a DAC and an arduino or something), which I think would give an error signal with a zero at the optimum if done correctly.
- It may be helpful to add a TTL control to turn the feedback on and off, like we have on our lockbox circuits. This would be helpful if we ever decided to bring our lab into the 21st century and make it possible for our computers to lock the lasers for us. If we end up doing this, it wouldn't be too hard to hack in a relay or FET switch in parallel with the mechanical switch, but it would be nicer to actually have a devoted input for this, especially if we end up using this circuit for other lasers.
- The mode noise amplitude from the photodiode is lower than I had thought. I must have been wrong when I (Zak) said that the noise was roughly 20mV pk-to-pk from the photodiode? At any rate it might be helpful to do one of the following:
    - Increase the gain of the input stages
    - Increase the value of R_Setpoint_Limit_1 to make it easier to set the Setpoint pot to a good value
    - Instead of connecting the "Mode Noise" input directly to the balanced photodiode signal, connect it to the error monitor of the frequency stabilization lockbox. That lockbox has some gain at its input and makes the noise ~10 times larger
- Maybe we should flip around the PIezo In input connections so that it has gain +1 instead of -1 when hooked up without having to swap the header pins + and - orientation relative to all the others.
- A feedforward piezo offset voltage BNC input might be helpful.
    - This would be helpful for locking the laser remotely.
        - Connecting a DAC controlled by the computer (e.g. an NI card output) to this input would allow you to put the laser into the correct mode remotely.
            - This would replace the role of us adjusting the current knob on the current controller since that can't easily be done remotely.
        - Of course additional connections, namely to the reference laser lockbox, would be necessary for locking this laser remotely.

Change Log

- On 5 March 2020 the box seemed to not receive the Mode Noise input, so we made the following changes:
    - When testing on the bench, the circuit worked fine. The socketed components were all a little loose because their lead diameters were too small for the sockets, so I (Zak) suspected that maybe one had been jostled and temporarily lost connection to the circuit. In case that was the issue, I soldered thicker header leads onto all of the socketed components to make them more secure.
    - I changed R_Setpoint_Limit_1 from 12 kΩ to 100 kΩ so that the set point knob position was larger and easier to set. Previously it always needed to be ~0.1
    - I changed C_Int_Gain from 470 nF to 47 nF. This was done to make the mode stabilizer settle faster after being turned on. Previously it would take ~1 minute to settle, which was usually fine. However, if the set point was set too low or too high then the laser mode hops and goes out of lock. It was annoying to have to wait ~1 minute to find out if that would happen or not, and decreasing C_Int_Gain by a factor of 10 should make that happen 10 times faster.

Wiki To-Do

- Nothing at the moment

Appendix C

Lock Monitor Blacs Plugin

Lock monitor is a plugin designed to detect when a laser goes out of lock, then automatically pause the experiment, relock the laser, and resume the experiment. In particular it is a plugin written for blacs, a software component of the Labscript suite [85, 86, 87]. Blacs is the program in charge of running the hardware to perform shots of the experiment sequence. It also has a plugin architecture which makes it possible to implement additional features. As blacs is in charge of interacting with hardware when shots are run, it was the natural place to implement lock monitor.

Lock monitor has made a significant impact to our lab productivity. It makes it possible to take data for extended periods without constant human supervision. This is particularly helpful when performing long-running optimizations with M-LOOP [32, 33], as a several-hour-long optimization attempt can be essentially ruined by a laser going out of lock. That was no longer an issue after implementing lock monitor. Users are freed to do other tasks rather than having to babysit their apparatus as it runs, which increases lab productivity. Furthermore, it makes it possible to take data when no one is around to monitor the apparatus, even overnight or when running the apparatus remotely. That aspect was critical while taking data during the COVID-19 pandemic. Lock monitor is also a morale booster. It can be frustrating to reach the end of a data collection run and then find out that it needs to be redone because a laser went out of lock. With lock monitor, users can start a data run confident that they will have results at the end of it.

Although lock monitor has not yet been adopted by other labs as of this writing, it is written in a general enough way so as to be used by any lab. Adoption does require writing some code, but documentation, working examples, and base classes are provided. The need for users to write some code stems from the fact that lock monitor is written to be as flexible as possible. It is designed such that a lab can use it with any hardware they have to lock a laser to any signal they have. The user then must write the code to communicate with their specific hardware (acquire a waveform from an oscilloscope, enable the feedback from a lockbox, etc.) and the code to interpret their error signal. Lock monitor handles calling that code, provides GUI controls, pauses and resumes the experiment as needed, provides a settings interface, and provides a text output terminal for monitoring lock checks and locking attempts. It also provides a template class for the common scan/zoom approach to locking lasers, which makes it easier for users to write code to lock lasers in that manner. The locking procedure for our 795 nm laser described in Section 3.2.2 follows that template. On the other hand, the locking procedure for our 780 nm power amplifier lasers uses an entirely different approach (described in Section 3.2.1), which demonstrates lock monitor's flexibility. Furthermore, although it was developed with monitoring laser locks in mind, lock monitor can be used for other monitoring applications as well. For example, it could be used to detect when a fiber coupling efficiency drops below some threshold value then trigger a fiber coupling alignment procedure to run.

As of this writing, it is still an open question as to whether or not lock monitor will be packaged with blacs in the future [1]. If it is included, then users will be able to enable lock monitor simply by editing their labconfig file, which is the configuration file for Labscript. If lock monitor is not packaged with blacs, then I will make it available as a 3rd party blacs plugin and post it to a repository associated with my Github account [2] with installation instructions.

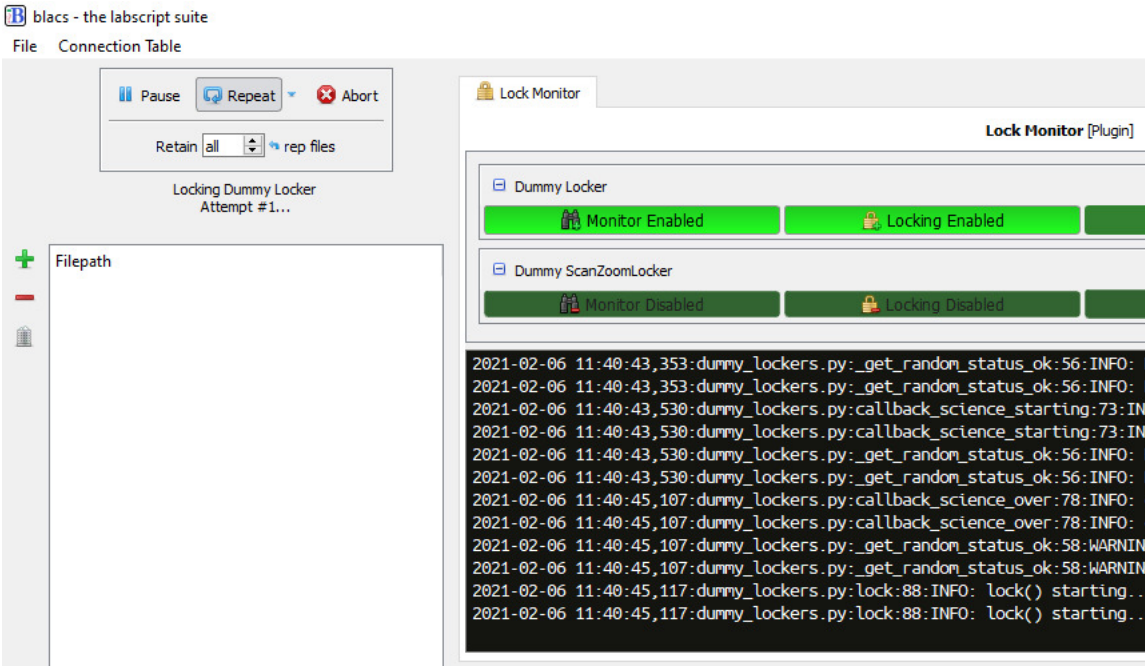
An introduction to lock monitor and its capabilities is presented in its README. The README is written in reStructuredText but has been rendered into a human-

[1] The Github pull request discussing this is available at https://github.com/labscript-suite/blacs/pull/88.

[2] https://github.com/zakv

readable form for inclusion in this thesis. The rendered form of the README is presented below.

Lock Monitor



Introduction

Lock monitor is a blacs plugin designed as a framework to help users automatically pause the blacs queue when a laser goes out of lock and optionally attempt to lock the laser automatically. The role of lock monitor is possibly best explained by listing what is does and doesn't do.

What Lock Monitor Does:

- It calls user-written code to check if lasers are in lock at various points during the execution of a shot.
- It temporarily pauses the blacs queue and optionally calls user-written code to lock a laser when it is determined to be out of lock.
    - The queue is unpaused if the laser is successfully locked, or remains paused if the laser fails to lock after a few attempts.
- It provides a blacs tab as a GUI for controlling various lockers.
- It provides a settings interface for adding/removing lockers.
- It provides base classes for monitoring and locking lasers, which include methods that the user must flesh out.
- It sandboxes user-written code in separate worker processes to reduce the chance that bugs there can crash blacs.

What Lock Monitor Does NOT Do:

- It does NOT include code to communicate with hardware, such as function generators and oscilloscopes, as the possible variety of hardware is too large. The user must use 3rd party libraries and/or write the code to communicate with the required hardware themselves.
- It does NOT perform the required signal analysis.
    - In particular it does NOT identify features oscilloscope traces of spectroscopic features to determine where to lock the laser.

Although lock monitor was written with locking lasers in mind, it likely has other uses. For example, a user could monitor the power transmitted through a fiber and pause the blacs queue if it drops below a certain level. They could even write code to automatically realign light into the fiber if the mount has electronic actuators, which lock monitor could automatically call.

Lock monitor took inspiration from, but is distinct from, the labwatch blacs plugin.

Using Lock Monitor

This section briefly describes how lock monitor is used. More detailed instructions and information on how to write the code required to interface lock monitor with your system are provided in other sections below.

Once the user has written their locker classes, using lock monitor is very simple. First, lock monitor is enabled by adding the line `lock_monitor = True` to the `[BLACS/plugins]` section of the labconfig. Instances of user-written locker classes are then added by entering their import path in the "Lock Monitor" tab of the blacs preferences menu, accessed via File -> Preferences -> Lock Monitor. A restart of blacs is required for those changes to take effect.

When lock monitor is enabled, a blacs tab is created with the title "Lock Monitor". The tab contains one collapsible section per locker, each one containing the controls for the given locker. The controls include the following:

- `Monitor Enable`: This controls whether or not the methods to check the laser's lock status are enabled.
- `Locking Enable`: This controls whether or not lock monitor should attempt to lock the laser if it is found to be out of lock. If monitoring is enabled but locking is disabled and the laser is found to be out of lock, then the current shot will be aborted and requeued, and the queue will be paused.
- `Force Lock`: When clicked, this will force lock monitor to lock the laser when the next shot is executed, regardless of whether or not the laser is deemed to be out of lock. Note that, as of this writing, the laser will not be locked immediately; it will be locked when the next shot is run. Once the locker has been forced to lock, this control will automatically be turned off again.
- `Restart Worker`: When clicked, this will instruct lock monitor to restart the worker process for the locker. Restarting the worker will re-import the locker and re-run its initialization methods, which can be helpful if something goes wrong or if you've made changes to the locker's code. Note that, as of this writing, this will not be done immediately; the worker will be restarted when the next shot is run. Once the worker has begun restarting, this control will automatically be turned off again.

Approach

Lock monitor works by calling methods of user-written python classes at certain points in the execution of a shot as the blacs state machine moves between states. These methods are referred to as callbacks, as they are run when blacs executes its plugin callbacks. Each user-written callback should return either `True` or `False`. When `True` is returned, blacs is allowed to continue on running the shot as usual. However, when `False` is returned, the following steps are taken:

1. Lock monitor attempts to abort the shot, though the shot may continue to run if it is too late to abort.
2. The blacs queue is paused.
3. Whether or not the shot was successfully aborted, it is NOT sent to lyse and is instead prepended to the top of the blacs queue.
4. Lock monitor then calls user-written code to lock the laser if such code is provided.
5. If the lock is successful, as determined by more user-written code, the queue is then unpaused. If the lock is unsuccessful, a few more attempt to lock are made. If none of those attempts are successful, or if no automatic locking code is provided for the laser, then the queue is left in the paused state.

Put another way, the user writes callbacks which are methods of a locker class. The callbacks should communicate with whatever hardware is necessary then return `False` if they determine that a laser is out of lock and `True` otherwise. For automatic locking to work, the user must also write a `lock()` method for their class, which lock monitor will call when necessary.

Making `Locker` Classes

The `Locker` Class

The code used to check and lock lasers should be organized into a locker class. User-created locker classes should inherit from the provided `Locker` class which can be imported from `blacs.plugins.lock_monitor.lockers`. That class provides the minimum set of properties and basic methods that locker classes should have. The class properties and methods are described in detail in the docstrings, which should be referred to when writing a custom locker class. Briefly, the most important attributes are:

- `self.logger` for logging progress, results, etc. with commands such as `self.logger.info("Doing some step...")`. Logging isn't strictly necessary but it is extremely helpful when debugging and so it is strongly encouraged that users add logging statements to their locker class methods.
- `self.display_name` which is used to label the controls, as well as a few other things.
- `self.plot_dir` is the suggested directory in which to save plots, which is a subdirectory of the directory specified by `blacs.plugins.lock_monitor.lockers.LOG_PATH`, which in turn is a subdirectory of the default labscript logging directory specified by `labscript_utils.setup_logging.LOG_PATH`. The subdirectory path also includes `self.plot_root_dir_name`, as well as subdirectories for the year, month, day, and full date/time. As with logging, generating plots isn't strictly necessary but can be extremely helpful in debugging when things go wrong, so doing so is strongly encouraged.
- `self.plot_root_dir_name` is the name of a folder in the path to `self.plot_dir`, which is created in `blacs.plugins.lock_monitor.lockers.LOG_PATH`. Note that the suggested full path for where to save plots is `self.plot_dir`, not `self.plot_root_dir_name`.
- `self.auto_close_figures` is used to indicate whether or not methods should close figures after generating them. This isn't enforced, but should be done, so the user's code should check the value of this property and close any figures generated if it is set to `True`. Generally this should be set to `True` when the locker class is used with lock monitor so that plots aren't left open, potentially consuming a lot of memory. Setting this to `False` can be useful though when testing/developing locker classes e.g. in a Jupyter notebook

The most important methods of the `Locker` class are the following:

- `self.__init__()`: If a user writes an `__init__()` method for their locker class, they should make sure that it calls the `__init__()` method of `Locker` by calling `super().__init__()`.
- `self.init()`: Not to be confused with `__init__()` (note the presence/lack of double underscores), the `init()` method is called by lock monitor when it starts up. For reasons that will be apparent later on, the code included in the class's `__init__()` method will run when the module with the user's locker class is imported. It is probably not desirable to connect to hardware in the `__init__()` method then because then any code that imports that module will immediately open a connection to the hardware. To work around this, put the code to open connections to hardware in the `init()` method of the locker class so that it won't be called when the module is imported but it will still be run by lock monitor.
- `self.close()`: Lock monitor will call each locker's `close()` method when it shuts down. That makes it the appropriate place to put code to disconnect from hardware and perform any other tasks that should be done when closing.
- `callback_*`: Callback methods are the methods called as shots are run to check if the laser is in lock. See the docstring of each of those methods to determine at what point in running a shot a given callback method is executed.
    - Not all of these methods need to do anything useful; it's perfectly fine to e.g. check the lock status in one callback and have the others do nothing.
    - Each callback method should generally return `True` except when it detects that the laser is out of lock, in which case it should return `False`. Returning `True` indicates to lock monitor that it should continue on as normal, while returning `False` will cause lock monitor to begin handling the fact that the laser is out of lock. Unused callbacks, or callbacks which perform some tasks but do not themselves determine whether or not the laser is locked, should just always return `True`.
    - Keep in mind that although the order in which callbacks are run is always the same for every shot, some callbacks may be skipped. For example, this may happen if the shot is aborted. Therefore do NOT assume that every callback will run for every shot. It is good practice to have `callback_pre_transition_to_buffered()`, which is the first callback run for any shot, do any clean up required if some callbacks from a previous shot were skipped, if that is necessary.

- `lock()`: When a callback method returns `False` to indicate to lock monitor that the laser is out of lock, lock monitor will call the locker's `lock()` method.
    - If the user's locker class does not support automatic locking, then this method should simply return `False` which indicates to lock monitor that the laser is still not in lock.
    - If the user's locker class does support automatic locking, then the `lock()` method should attempt to lock the laser and return `True` if it succeeded or `False` otherwise. Because the code to check the lock status presumably already exists in the `calback_*()` methods, it is usually convenient for `lock()` to call one or more of those methods after attempting to lock the laser. As mentioned above, it is very useful for the `lock()` method to use `self.logger` to log its progress and ideally it should save plots as well, typically in the directory specified by `self.plot_dir`.
    - Before saving any figures, `lock()` should call `self.update_save_time()` which will set a new timestamp to use when generating the year/month/day/time subdirectories in `self.plot_dir`.

To write a custom `MyLocker` class, the user should create a class which inherits from `Locker` then override some or all of the methods mentioned above. The methods should be written to implement the behavior described in the docstrings for the corresponding methods of the `Locker` class. The code below shows a typical starting point:

```
from blacs.plugins.lock_monitor.lockers import Locker

class MyLocker(Locker):
    def init(self):
        # Do some logging for debugging purposes.
        self.logger.info(f"init() called for {self.display_name}.")

        # Do any other desired setup here.

        # init() doesn't need to return anything.
        return

# Create an instance of the MyLocker class which lock monitor will use. Of
# course make sure to provide the actual initialization arugment values.
mylocker_instance = MyLocker(<initialization args>)
```

Generally it is best to start with `init()` and/or `__init__()` (make sure to call `super().__init__()`) and use it to add the ability for the locker class to communicate with hardware. Next, override one or more of the `callback_*` methods to have the locker check if the laser is in lock. Don't forget to add some logging statements using `self.logger` to help with debugging! It is not necessary to override all of the `callback_*` methods; just override as many as you need. Once that is done, optionally implement a `lock()` method so that lock monitor can automatically lock the laser when it goes out of lock. In addition to logging, you'll likely also want to generate some plots and save them to `self.plot_dir`, which can be very helpful for debugging as well. Make sure to close those plots if `self.auto_close_figures` is set to `True`.

Developing the code for a `Locker` class can take a lot of trial and error. For that reason it is recommended to develop that code in an interactive environment, such as a Jupyter notebook. When doing so, it can be helpful to set `auto_close_figures` to `False` so that they appear in the notebook.

The `ScanZoomLocker` Class

For convenience, lock monitor also includes a `ScanZoomLocker` class which inherits from `Locker` but includes some additional template methods for locking to a spectral feature. Again, the user is in charge of implementing most of the methods to acquire and analyze the required data.

`ScanZoomLocker` is designed to lock a laser in analogy with how they are often locked by hand. Often a user sweeps the laser's frequency and observes some spectrum on an oscilloscope. The user then looks at the oscilloscope trace and identifies the target feature that they'd like to lock to. Once the target feature is identified, the user repeatedly reduces the amplitude of the scan and adjusts the scan's offset to center the scan around the target feature. At some point the user often adjusts the setpoint (aka error offset) as well. After zooming in sufficiently far, the feedback is enabled to engage the lock. Lastly the user verifies that the laser successfully locked.

`ScanZoomLocker` includes methods which the user fills out to perform each of those steps that are normally performed when locking a laser to s spectral feature manually. It also includes an implementation of the `lock()` method to perform those steps in order. Of course the user is free to override the `lock()` method as well if desired.

In addition to the attributes inherited from the `Locker` class, `ScanZoomLocker` has the following notable attributes, many of which are initialization arguments of the class:

- `zoom_factor` sets by what factor the scan range is shrunk by during each iteration.
- `n_zooms` sets the number of zooming iterations to perform before enabling the feedback loop to lock the laser.
- `n_zooms_before_setpoint` sets the number of zooming iterations to perform before adjusting the setpoint (aka error offset).
    - Often the target spectral feature is very narrow in the initial scan, and sometimes it is wider than the final zoom's scan range. In such cases, performing some of the zooming iterations before adjusting the setpoint makes it possible to examine the target feature and determine the desired setpoint.
- `initial_scan_amplitude` sets the amplitude of the scan used at the beginning of the first zooming iteration.
- `initial_scan_feedforward` sets the feedforward control value used at the beginning of the first zooming iteration.

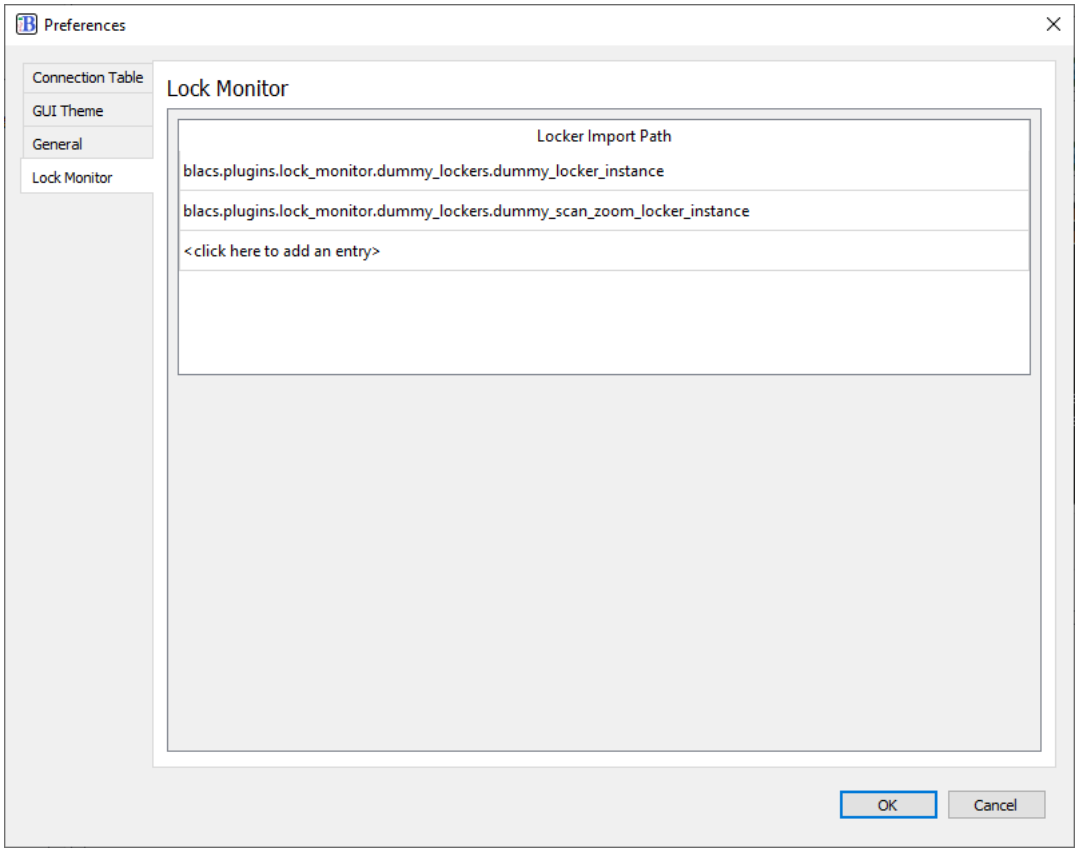
In addition to the methods inherited from the `Locker` class, `ScanZoomLocker` locker also includes the following template methods which should be implemented by the user:

- `check_lock()`
- `disable_feedback()`
- `enable_feedback()`
- `get_feedforward()`
- `get_lockpoint_feedforward_value()`
- `get_lockpoint_setpoint_value()`
- `get_scan_amplitude()`
- `get_setpoint()`
- `lock()`
- `set_feedforward()`
- `set_scan_amplitude()`
- `set_setpoint()`
- `zoom_in_on_lockpoint()`

For most of those methods their purpose is clear from their name. For any that aren't clear, consult the method's docstring for a description of the method's intended behavior. The code for the `DummyScanZoomLocker` class (discussed below) is also a good reference when implementing the methods required by `ScanZoomLocker`. The `zoom_in_on_lockpoint()` and `lock()` methods are already implemented (though the user is welcome to override them as desired), but the user must implement the other methods listed above in order for `ScanZoomLocker.lock()` to work.

Adding Lockers to Lock Monitor

Once your locker is ready for use, lock monitor must be configured to use it. Lockers are added to lock monitor using its settings interface



To add a locker to lock monitor, follow the steps below:

1. Create a python module in which an instance of your locker class is created.
    - It's typically convenient, but not required, to create the instance of the locker class in the same module in which it is defined.

2. Next, make sure that the python module can be imported from the python interpreter.
    - When labscript is installed, the `pythonlib` folder in the `userlib` directory is added to `PYTHONPATH`, so that can be a good place to put the module.

3. Add the instance of your locker class to lock monitor's list of lockers.
    - To do so, open blacs and select File -> Preferences then navigate to the "Lock Monitor" tab. Once there add the import path of your locker instance (the instance of your class, NOT the class itself!) to the table and click "OK".

4. Restart blacs for the changes to take effect.

Following from the `MyLocker` example above, suppose that code is included in a file `my_locker_module.py` which is placed in the directory `pythonlib/lockers`. To add that locker to lock monitor, you would add the import path `pythonlib.lockers.my_locker_module.mylocker_instance` in the lock monitor settings tab. Again note that the import path ends with `mylocker_instance` (the instance of the class) and NOT `MyLocker` (the class itself).

After blacs restarts, a new collapsible tool palette should be added to the lock monitor blacs tab which contains the controls for the locker and is labeled by the display name passed during `mylocker_instance`'s initialization. If no such tool palette appears, see the troubleshooting section below.

For convenience, lock monitor also includes some "dummy" locker instances in `blacs.plugins.lock_monitor.dummy_lockers` which do not actually control any hardware but can still be added as lockers to lock monitor. This can be useful for testing/debugging, or even just for seeing how lock monitor behaves when everything works correctly. The `DummyLocker` class in that module is a simple locker which does nothing but randomly pretend to be out of lock on occasion. It then pretends to lock simply by pausing for a few seconds. An instance of that class can be added to lock monitor by including the import path `blacs.plugins.lock_monitor.dummy_lockers.dummy_locker_instance` in the settings. The dummy locker module also includes the `DummyScanZoomLocker` class, which simulates using the `ScanZoomLocker` class. It also randomly pretends to be out of lock on occasion. Additionally it pretends to lock by generating simulated data of a spectroscopic signal and zooming in on a dispersive feature. In the process it produces log messages and plots which can be seen in the log folder. An instance of that class can be added with the import path `blacs.plugins.lock_monitor.dummy_lockers.dummy_scan_zoom_locker_instance`.

Creating a Logger

Due to the importance of good logging, the `Locker` class requires an instance of the `logging.Logger` class for one of its initialization arguments. This will be straightforward for those familiar with Python's `logging` module, and an example of how to set up a logger is included below to help other users. The example sets up a logger that will both print messages to console and record them to a file in labscript's default log directory. Of course feel free to change any of the options as desired, particularly the name for the log file.

```
# Make the required imports.
import os
import logging
import sys
from blacs.plugins.lock_monitor.lockers import LOG_PATH

# Get a logger. This is what should be passed to ``Locker`` during
# initialization.
logger = logging.getLogger(__name__)

# Configure some options for the logger. In particular set it to process
# messages of priority DEBUG or higher, and set the format of the log
# messages so that they include a lot of helpful data such as the time, the
# name of the file, the name of the calling function, and so on.
logger.setLevel(logging.DEBUG)
formatter = logging.Formatter(
    '%(asctime)s:%(filename)s:%(funcName)s:%(lineno)d:%(levelname)s: %(message)s'
)

# Set up a console handler, which will take the messages sent to the logger
# and print them to the console. It is set to only do this for messages of
# priority level INFO or higher.
console_handler = logging.StreamHandler(sys.stdout)
console_handler.setLevel(logging.INFO)
console_handler.setFormatter(formatter)
logger.addHandler(console_handler)

# Set up file handler, which will take the messages sent to the logger and
# record them to a log file called 'my_locker.log'. It will record messages
# of priority DEBUG or higher. The file is opened with mode 'w' which makes
# it overwrite any existing copy of the log file on startup.
full_filename = os.path.join(LOG_PATH, 'my_locker.log')
file_handler = logging.FileHandler(full_filename, mode='w')
file_handler.setLevel(logging.DEBUG)
file_handler.setFormatter(formatter)
logger.addHandler(file_handler)

# Make the first log entry.
logger.debug('Logger configured.')
```

Once the logger has been configured and a locker has been instantiated with it, it is stored as the locker's `self.logger` attribute. The locker can log messages from within the locker class methods as demonstrated in the example log messages below:

```
from blacs.plugins.lock_monitor.lockers import Locker

class MyLocker(Locker):
    def init(self):
        # Logging at the debug level is good for recording tedious details
        # or small steps of progress. Python f-strings are very helpful
        # here.
        self.logger.debug("Starting some small step...")
        self.logger.debug("Finished some small step.")
        self.logger.debug("Record some very detailed information here...")

        # Logging at the info level is great for keeping track of higher-
        # level program flow and recording some useful information, such as
        # the feedback loop error signal amplitude.
        self.logger.info("Starting some series of steps...")
        self.logger.info("Finished some series of steps.")
        self.logger.info("Measured some important value to be... ")

        # Logging at the warning level is good for indicating something
        # seems wrong.
        self.logger.warning("Some signal is starting to get a bit high...")

        # Logging at the error level is good to do when things do go wrong.
        self.logger.error("Something has gone wrong; here's some info:...")

        # The logger.exception method can be used in except blocks to record
        # the traceback of the error that was caught.
        try:
            # Something that will throw an error.
            raise Exception("Throwing an error for demonstration purposes.")
        except Exception:
            # Record a message which will automatically also include the
```

```
            # exception's traceback.
            logging.exception("Caught the following exception:")
```

As mentioned previously, in addition to logging information it can be very helpful to save plots as well. The example below includes a method that shows how this can be done.

```
import matplotlib.pyplot as plt
from blacs.plugins.lock_monitor.lockers import Locker

class MyLocker(Locker):
    def save_a_plot(self):
        # This example method assumes that some other method has stored the
        # data from an oscilloscope trace as the attribute self.scope_trace.
        # Another method, likely lock(), should probably call
        # self.update_save_time() as well so that a new plot directory is
        # created with the appropriate timestamp.

        # Create a figure and axes to plot on.
        fig = plt.figure()
        axes = fig.add_subplot(111)

        # Plot the data.
        axes.plot(self.scope_trace)

        # Add text information so that people know what the plot shows!
        axes.set_title("Scope Trace of...")
        axes.set_xlabel("Time (arbitrary units)")
        axes.set_ylabel("Spectroscopy Signal (Volts)")

        # Save the figure as an image. Of course feel free to change the
        # filename to something more appropriate and informative!
        filename = os.path.join(self.plot_dir, 'scope_trace.png')
        fig.savefig(filename)

        # Lastly, remember to close the figure if configured to do so!
        if self.auto_close_figures:
            plt.close(fig)
```

Tips

Below are some tips for setting up and using lock monitor:

- Test out the dummy lockers included with lock monitor to get a sense of what it's like to use lock monitor before writing any code yourself.
    - The dummy lockers don't control any real hardware but can be added to lock monitor for testing/debugging purposes.
    - The dummy locker can be added by including the import path `blacs.plugins.lock_monitor.dummy_lockers.dummy_locker_instance` in the lock monitor settings menu.
    - Similarly the dummy scan/zoom locker can be added by including `blacs.plugins.lock_monitor.dummy_lockers.dummy_scan_zoom_locker_instance`.
    - Once you are done playing around with the dummy lockers, it's best to turn off their monitoring using the control in the blacs tab, or just remove them from the list of lockers and restart blacs. Otherwise they will continue to occasionally pretend to be out of lock and then pretend to lock themselves, which will slow down data taking. On rare occasions they may also pretend to fail to lock several times in a row, which will cause lock monitor to give up and pause the blacs queue.

- Read the docstrings of the `Locker` and/or `ScanZoomLocker` methods when you implement them to understand what they should do.
    - The code for the dummy lockers can also be a helpful reference.

- Develop your code for your locker class in an interactive environment, such as a Jupyter notebook, for quick testing and debugging. Once you have a basic working example, then organize the code into the appropriate locker methods.
- Search the internet for code examples on how to communicate with the hardware you plan to use. If you are lucky you may even find a python package already exists for communicating with your device.
- Have the `init()` method of your locker class reset any hardware to its default state before adjusting settings.
    - This ensures that the instrument always starts in the same state, no matter what options someone may have changed since the last time it was used, which makes the code much more robust.

- Many instruments with controls on the instrument itself support a command to "lock" (aka disable) the manual controls. Locking the manual controls during `init()` also makes the system more robust, as it makes it less likely that someone will manually adjust some settings that could mess up your locker class's code.
    - It can also be helpful to unlock the manual controls during the locker's `close()` method to restore manual operation when blacs is closed.

- Things go wrong sometimes, especially when first setting up lock monitor. And when things go wrong, having a log of what happened is invaluable. So make sure to add plenty of logging statements with `self.logger` to your locker classes.

    - Also don't forget to save plots, typically in the `self.plot_dir` directory.
    - This is particularly helpful to do during the `lock()` method, but can be done at other times as well.

- Sometimes hardware communication can be I/O bound and slow, but threading can often be used to allow your code to run in parallel with the rest of blacs.
    - For example, often transferring the data from an oscilloscope can take a while, but the computer spends most of that time sitting idle waiting for the oscilloscope. If that communication all happens during one callback, blacs will have to sit and wait until it finishes, which will reduce the repetition rate of the system. Instead it can be quicker to use one callback to start a separate thread which requests data from the scope, then returns so that blacs can carry on. Then another callback can join the thread and retrieve the data.

- Keep in mind that not all callbacks are run for every shot; some are skipped if shots are aborted for example.
    - Make sure your code is robust to callbacks being skipped.
    - It can often be a good idea to reset some things, such as attributes/variables that store data, to default empty values during `callback_pre_transition_to_buffered()` since that is always the first callback run for any shot. That ensures that any data from previous runs is cleared to avoid any confusion.
    - This is also important to keep in mind if any callbacks are supposed to join threads.

- Avoid sharing hardware between different locker classes.
    - The locker callbacks are run in parallel in different processes, so setting up two lockers to talk to the same piece of hardware can cause them to interfere with each other.
    - That said, it's not impossible for different lockers to use the same piece of hardware, but be prepared for some debugging!

- Just as with analog and digital outputs in blacs, locker controls can be "locked" to their current state by right clicking on them and selecting "lock".
    - When a control is "locked" it ignores left mouse clicks, which can be useful for making sure you don't accidentally disable monitoring or automatic locking.
    - A locked control can be unlocked by right clicking on it and selecting "unlock".

Troubleshooting

Below are some general troubleshooting tips when things go wrong.

- Add plenty of logging statements when writing locker classes, and read through the log when things go wrong.
    - Including actual values of data in log messages can also make them much more helpful.
    - Depending on how the logger is configured, some message levels (e.g. `DEBUG`) may be recorded in the log file but NOT printed to the console, or vice versa.

- Check blacs's log file as it can contain additional information, including information logged by the lock monitor plugin itself.
    - This can be a a good place to look when a locker included in the settings doesn't appear in the lock monitor tab.
    - This is a particularly good place to look if lock monitor itself won't start.

- Start blacs from the command line and keep an eye on the text that it prints to that console.
    - This information is included in the blacs log as well, but much less information is printed to the console which can make it easier to find the important parts.

- Generate and save relevant figures in your locker's `lock()` method.
    - If a locker attempts but fails to lock, refer to the saved figures to help determine why.

FAQs

There is no "Lock Monitor" tab in blacs

First, ensure lock monitor is enabled by adding the line `lock_monitor = True` to the `[BLACS/plugins]` section of the labconfig and restart blacs. If the "Lock Monitor" tab still doesn't appear after that, try starting blacs from the command line and see if any error information is printed. The blacs log file can also be a good place to check for error messages related to lock monitor.

My locker doesn't appear in the lock monitor blacs tab

Perform the following checks:

- Ensure that the import path for the locker is added in the settings for lock monitor by clicking "File" -> "Preferences" and navigating to the "Lock Monitor" settings tab. Restart blacs after adding it if it was not already present.
    - As a debugging check you may want to try adding one of the dummy lockers mentioned above to ensure that this step works ok.

- Check that the locker can be imported in python.

    - Open a python interpreter and type `import <import_path>`, replacing `<import_path>` with the string provided to the lock monitor settings menu. If a `ModuleNotFoundError` or an `ImportError` is raised then you've found your problem.
    - If importing your module opens a connection to hardware it may be best to close blacs before performing this test to avoid trying to connect to the hardware twice at the same time.

- Ensure that the import path provided to lock monitor imports an instance of your locker class, NOT the locker class itself.
    - For example, if you write a class called `MyLocker`, then your module should create an instance of the class with `my_locker_instance = MyLocker(...)`, then the import path should end with `.my_locker_instance` and NOT `MyLocker`.

- Run blacs from the command line to see its console output and/or examine its log files to look for errors.
    - If a locker fails to start, it won't be added to the lock monitor blacs tab, but the error raised during its startup should be logged.

How do I remove a locker from lock monitor?

Open the settings menu in blacs ("File" -> "Preferences" -> "Lock Monitor") then select the import path that you'd like to remove. Simply erase all of the text and press enter to delete the entry. Finally restart blacs for the changes to take effect.

Appendix D

Absorption Imaging PCA

Absorption imaging is a common technique [154] for imaging ultracold atomic clouds. However, one way that it is commonly implemented has a normalization issue which leads to systematic errors in the calculated atomic distributions. This appendix first provides some background information about absorption imaging and its difficulties with interference fringes. It then describes a common approach to reducing the effect of those interference fringes based on principal component analysis (PCA). A normalization issue associated with a common modification to the algorithm is then pointed out. This appendix then provides a corrected modification to that approach which resolves the normalization issue. The text here is intended to provide enough context for the reader to understand the modifications, but other resources may provide a more pedagogical introduction to absorption image processing with PCA.

D.1 Introduction

The principle behind absorption imaging is straightforward. Simply shine a beam of light through the cloud and onto a camera so that the cloud leaves a shadow in the beam. Then take two images, one with and one without the atoms, and compare them so see the shadow. I'll refer to those as the "atom" image and "beam" image respectively. A third image can also be taken with the beam off, then that image is subtracted from the other two to remove any background light sources. That image

will be denoted the "background" image. The cloud's spatial distribution can then be inferred from the shape of its shadow.

It is a well-known fact that absorption imaging is troubled by fringes in the acquired images. The difficulty is that the light must be narrow in frequency compared to the atomic transition in order for it to interact with the atoms effectively. Imaging transitions often have linewidths in the $\sim$ MHz range, which implies that the coherence length of the light can be hundreds of meters. The beam generally passes through various optics, and each interface generally produces multiple reflections of the beam. Some of these reflections also hit the camera, leading to interference patterns due to the light's coherence. Anti-reflection coatings reduce the significance of these multiple reflections, but do not fully eliminate them.

The interference patterns are not problematic if they are the same in the images with and without atoms, but that is often not quite the case. If the path lengths for different reflections change slightly, even by just a fraction of a wavelength, then the interference patterns shift. Vibrations, thermal drifts, and other sources can cause changes in path lengths, thus shifting the interference patterns in the time between the atom and beam images. Pointing noise can also lead to a change of the interference pattern in a similar manner. The most straightforward implementation of absorption image processing involves simply taking the log of the ratio of the intensity in the beam and atom images, which gives the optical depth of the could. That approach cannot distinguish between an atom's shadow and a change in image intensity due to a shift in the fringe pattern. Therefore the changing fringe patterns appear in the calculated atom distribution, reducing the SNR of the image.

D.2 Principal Component Analysis

Although other techniques exist [155], the most common way to reduce the presence of these fringes is through principal component analysis (PCA) [1]. Qualitatively, the

[1]Interestingly PCA is a common statistical technique and is also the basis of the popular Eigenfaces facial recognition algorithm [156, 157].

approach is as follows. Simple absorption image processing just uses the single beam image taken after an atom image as the reference when calculating the cloud's optical depth (OD). A better reference can be tailor-made for the atom image by carefully combining many different beam images from different shots of the machine. By comparing all of the beam images, one can see how the beam profile tends to vary from shot-to-shot. Then one can look at the region of the atom image where there are no atoms (i.e. far from the position of the atomic cloud) and reconstruct what the rest of the image would look like if it weren't for the cloud's shadow. That reconstruction is done based on the observations of how the beam profile tends to vary. Once the reconstruction is done, the OD is calculated by taking the log of the ratio of the reconstruction to the atom image. The reconstruction generally has a fringe pattern more similar to that in the atom image than the single absorption image does, which reduces the amplitude of the fringes in the calculated atom distribution. This approach effectively uses the information contained in many beam images to make a better guess at what the atom image would have looked like without atoms, which then leads to a better estimate of the atomic distribution.

The above procedure can be explained more rigorously as follows [2]. Each image is a 2D array of real numbers, one for each pixel. Let j and k be the number of rows and columns in the image respectively. Each number in the array represents the number of photons detected by the corresponding pixel and is proportional to the light intensity incident on the pixel. Those 2D arrays form a vector space in which each image is a vector. It can be somewhat confusing that a 2D array is a vector as it seems more analogous to a matrix. It can be helpful to instead imagine taking all of the columns in the image and stacking them on top of each other to make one large column vector. Doing that makes it more clear how to treat the image as a vector. In practice that actually is how it is done on the computer, as numerical libraries typically assume vectors are in that form [3].

[2] I'll use similar notation to that used in Ref. [157], which also explains this procedure.

[3] Depending on the implementation, the image vector may actually be the rows of the image concatenated into one long row to make a row vector. The column vector is then made by taking the transpose. This detail is not important to the algorithm, it just determines how to convert back and forth between 2D images and image vectors.

Label the beam image [4] column vectors as $\mathbf{\Gamma}_i$. Then let $\bar{\mathbf{\Gamma}}$ be the average of those vectors. Define $\mathbf{\Phi}_i = \mathbf{\Gamma}_i - \bar{\mathbf{\Gamma}}$, which is the difference between each beam image vector from the average beam image vector. Then define the matrix $\mathbf{A} = [\mathbf{\Phi}_1, \mathbf{\Phi}_2, \ldots, \mathbf{\Phi}_m]$ where m is the number of beam images used in the PCA. The principal components are then the eigenvectors $\mathbf{U}_i$ of the covariance matrix $\mathbf{A}\mathbf{A}^\intercal$ where the "T" superscript denotes the transpose. The principal components represent the ways in which the beam images tend to vary. In practice they tend to vary due to the interference patterns changing in amplitude and shifting around. This can be seen by converting the $\mathbf{U}_i$ to 2D image arrays and then plotting them. They typically show clear interference patterns. See Ref. [157] for some examples.

In practice a trick is usually used to calculate the $\mathbf{U}_i$. The dimension of $\mathbf{A}\mathbf{A}^\intercal$ is $jk \times jk$ which is typically large, so calculating its eigenvectors can be very time-consuming. However, those eigenvectors can be calculated from the eigenvectors $\mathbf{V}_i$ of the matrix $\mathbf{A}^\intercal\mathbf{A}$ via the relation [5] $\mathbf{U}_i = \mathbf{A}\mathbf{V}_i$. Notably $\mathbf{U}_i$ and $\mathbf{V}_i$ have the same eigenvalue. The matrix $\mathbf{A}^\intercal\mathbf{A}$ has dimension $m \times m$. Often (though not always) $m \ll jk$, i.e. the number of beam images is small compared to the number of pixels in each image. If $m < jk$, then $\mathbf{A}^\intercal\mathbf{A}$ is smaller than $\mathbf{A}\mathbf{A}^\intercal$ and its eigenvectors can be calculated faster. In that case, the $\mathbf{U}_i$ are typically determined [6] by first calculating $\mathbf{V}_i$, then left-multiplying those vectors by $\mathbf{A}$.

Notably $\mathbf{A}^\intercal\mathbf{A}$ has only m eigenvectors while $\mathbf{A}\mathbf{A}^\intercal$ has jk eigenvectors, so not all eigenvectors can be calculated this way. It can be shown that the $\mathbf{U}_i$ that can be calculated this way are the $\mathbf{U}_i$ with the largest eigenvalues. Therefore those $\mathbf{U}_i$ are the largest principal components, i.e. they represent the main ways that the beam images change from shot to shot. The principal components with smaller eigenvalues represent ways in which the beam images change very little and often only contain

[4] If the third "background" image with the beam off is taken, it should be subtracted from the beam and atom images first then the result should be used in this algorithm.

[5] This relation is easy to show. By the definition of an eigenvector, $\mathbf{A}^\intercal\mathbf{A}\mathbf{V}_i = \lambda\mathbf{V}_i$ where λ is the eigenvalue. Left-multiplying each side by $\mathbf{A}$ gives $\mathbf{A}\mathbf{A}^\intercal\mathbf{A}\mathbf{V}_i = \mathbf{A}\lambda\mathbf{V}_i$, then inserting $\mathbf{U}_i = \mathbf{A}\mathbf{V}_i$ gives $(\mathbf{A}\mathbf{A}^\intercal)\mathbf{U}_i = \lambda\mathbf{V}_i$. This shows that $\mathbf{U}_i$ is then an eigenvector of $\mathbf{A}\mathbf{A}^\intercal$ with the same eigenvalue λ.

[6] Note that $\mathbf{U}_i$ will not be normalized if the $\mathbf{\Phi}_i$ or $\mathbf{V}_i$ are not normalized. In that case the $\mathbf{U}_i$ should be scaled to normalize them.

random noise. Therefore the reconstruction can actually be more accurate when those principal components are omitted, so the fact that they aren't calculated isn't much of a problem. Often even some of the m eigenvectors (the ones with smallest eigenvalues) calculated using this trick above are dropped, which speeds up the analysis and can reduce noise. Let n be the number of principal components that are kept.

Once a set of principal components $\mathbf{U}_i$ has been calculated, the reference image must be constructed. The most straightforward way to do this is as follows. First calculate $\mathbf{\Phi}_{\text{atom}} = \mathbf{\Gamma}_{\text{atom}} - \bar{\mathbf{\Gamma}}$ where $\mathbf{\Gamma}_{\text{atom}}$ is the atom image vector. Then construct the matrix $\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_n]$. The reconstructed reference image can then be calculated [7] as $\mathbf{\Phi}_{\text{ref}} = \mathbf{U}\,(\mathbf{U}^\intercal \mathbf{\Phi}_{\text{atom}})$. This formula has a fairly physical interpretation. Each element i of the vector $(\mathbf{U}^\intercal \mathbf{\Phi}_{\text{atom}})$ is the projection of $\mathbf{\Phi}_{\text{atom}}$ onto the principal component $\mathbf{U}_i$. That projection is a measure of how much the image has varied from the mean beam image in the particular way described by $\mathbf{U}_i$, i.e. how much of the fringe pattern represented by $\mathbf{U}_i$ is in the image. Left-multiplying the vector of those projections by $\mathbf{U}$ creates a weighted sum of the $\mathbf{U}_i$ where the weights are the projections $\mathbf{U}_i^\intercal \mathbf{\Phi}_{\text{atom}}$. Put another way, $\mathbf{\Phi}_{\text{ref}}$ is the projection of $\mathbf{\Phi}_{\text{atom}}$ onto the subspace spanned by $\mathbf{U}_i$. After calculating $\mathbf{\Phi}_{\text{ref}}$, the reference image $\mathbf{\Gamma}_{\text{ref}}$ is calculated as $\mathbf{\Gamma}_{\text{ref}} = \mathbf{\Phi}_{\text{ref}} + \bar{\mathbf{\Gamma}}$. Finally the optical depth for each pixel is calculated as $\log\left(\mathbf{\Gamma}_{\text{ref}}/\mathbf{\Gamma}_{\text{atom}}\right)$ where both the division and logarithm are calculated element-wise. The atom distribution is then given by the optical depth for each pixel.

Unfortunately the straightforward reconstruction method of the last paragraph has a systematic issue. The problem is that the shadow of the atoms, which is present in $\mathbf{\Gamma}_{\text{atom}}$ and $\mathbf{\Phi}_{\text{atom}}$, affects the projection weights calculated from $(\mathbf{U}^\intercal \mathbf{\Phi}_{\text{atom}})$. This leads to systematic errors in the constructed $\mathbf{\Gamma}_{\text{ref}}$ and introduces noise into the final results. A common method to avoid that is to set all the pixels in $\mathbf{\Phi}_{\text{atom}}$ to zero wherever there are atoms [8]. The is equivalent to left-multiplying $\mathbf{\Phi}_{\text{atom}}$ by a mask

[7] Note that the number of operations necessary to calculate this result depends on the order in which the elements are multiplied. It is better to calculate $\mathbf{U}^\intercal \mathbf{\Phi}_{\text{atom}}$ first, then multiply that result by $\mathbf{U}$. Instead multiplying $\mathbf{U}$ by $\mathbf{U}^\intercal$ then multiplying that by $\mathbf{\Phi}_{\text{atom}}$ takes many more operations and uses more memory. Both orderings are mathematically equivalent of course, but they require different numbers of steps to evaluate.

[8] In principal it isn't known beforehand where the atoms are. However often one knows from past experience where they should be in the image. As a fallback one can always perform the

matrix [9] $\mathbf{M}$ which is a diagonal matrix with zeros and ones on the diagonal and zeros off of the diagonal. The elements with ones correspond to pixels without atoms which should be used to calculate the reconstruction weights, and the elements with zeros correspond to pixels with atoms which should be set to zero. However, masking the atom region this way also leads to systematic errors, as can be clearly demonstrated with a constructed example.

Suppose that the "atom" image actually has no atoms. In that case the reconstructed image ideally should be identical to the "atom" image. Consider the case that the atom image just happens to be [10] $\mathbf{\Gamma}_{\text{atom}} = \bar{\mathbf{\Gamma}} + \mathbf{U}_1$. In that case $\mathbf{\Phi}_{\text{atom}} = \mathbf{U}_1$. Ideally then $\mathbf{\Phi}_{\text{atom}}$ would project perfectly onto $\mathbf{U}_1$. Then the reconstructed $\mathbf{\Gamma}_{\text{ref}}$ would be equal to $\bar{\mathbf{\Gamma}} + \mathbf{U}_1$ which would be exactly equal to the "atom" image as hoped. However, that is not the case when the atom region is masked. In that case, the projection onto $\mathbf{U}_1$ is $\mathbf{U}_1^{\intercal}\mathbf{M}\mathbf{U}_1 \neq 1$ which throws off the reconstruction. It leads to the wrong weight for $\mathbf{U}_1$ because it gives a value less than one. It also leads to the wrong weights for the other $\mathbf{U}_i$ as they should all be zero because they are orthogonal to $\mathbf{U}_1$, but they get nonzero weights because they are not orthogonal to $\mathbf{M}\mathbf{U}_1$. Therefore, the reconstructed image will not match the "atom" image $\mathbf{\Gamma}_{\text{ref}} \neq \mathbf{\Gamma}_{\text{atom}}$. A slightly different approach to masking the atom region is necessary to avoid that effect.

D.3 Proper Atom Masking

The fundamental issue with the atom masking approach discussed in Section D.2 is that although the set of vector $\{\mathbf{U}_i\}$ are orthonormal, the set of masked vectors $\{\mathbf{M}\mathbf{U}_i\}$ are not orthonormal. Equivalently they are not orthonormal after projecting them into the column space of $\mathbf{M}$, which is the space of all images where the pixels

straightforward reconstruction to get some image (with some systematic issues), then use that image to determine where the atoms are. Then the reconstruction can be redone with the known atom region to reduce the systematic issues.

[9] A more memory-efficient way to achieve the same result is to store the diagonal of $\mathbf{M}$ as a column vector the perform element-wise multiplication between it and $\mathbf{\Phi}_{\text{atom}}$.

[10] The systematic errors are a problem for essentially all possible $\mathbf{\Gamma}_{\text{atom}}$; it is just easiest to see that there are issues when considering this simple value for $\mathbf{\Gamma}_{\text{atom}}$.

with atoms are set to zero. This is why the reconstruction weights don't quite come out right. The corrected approach uses a slightly different definition for $\mathbf{U}_i$ so that $\{\mathbf{U}_i\}$ are *not* orthonormal but $\{\mathbf{M}\mathbf{U}_i\}$ are.

The technique is essentially the following. The atom region of the beam images is set to zero, then those images are used to construct an orthonormal basis. Next it is determined how to write that orthonormal basis as a superposition of the input beam images, with the atom region still set to zero. Then the same superposition of images is taken, except without setting the atom region to zero. This gives a basis $\{\mathbf{U}_i\}$ that is not orthonormal in the usual sense, but is orthonormal over the background region. If the atom region is set to zero then the basis is orthonormal, i.e. $\{\mathbf{M}\mathbf{U}_i\}$ are orthonormal. The projection can then be done by calculating the weights $(\mathbf{M}\mathbf{U})^\intercal \mathbf{\Phi}_{\text{atom}}$, which goes nicely because $\{\mathbf{M}\mathbf{U}_i\}$ are orthonormal. Then those weights can be used to sum the $\mathbf{U}_i$ to give the reconstructed image.

Mathematically, the technique is implemented as follows. First the matrix

$$\mathbf{A}_{\text{mask}} = \mathbf{M}\mathbf{A} = \Big[(\mathbf{M}\mathbf{\Phi}_1), (\mathbf{M}\mathbf{\Phi}_2), \ldots, (\mathbf{M}\mathbf{\Phi}_m)\Big] \tag{D.1}$$

is constructed and the eigenvectors $\mathbf{V}_{i,\text{mask}}$ of $\mathbf{A}_{\text{mask}}^\intercal \mathbf{A}_{\text{mask}}$ (which has dimension $m \times m$) are calculated. At this point the $\mathbf{V}_{i,\text{mask}}$ could be converted to $\mathbf{U}_{i,\text{mask}}$ using $\mathbf{U}_{k,\text{mask}} = \mathbf{A}_{\text{mask}}\mathbf{V}_{i,\text{mask}}$. The $\mathbf{U}_{i,\text{mask}}$ will then be eigenvectors of $\mathbf{A}_{\text{mask}}\mathbf{A}_{\text{mask}}^\intercal$, but they will be zero in the masked region, which is not the goal. The aim is to construct a "basis" which is not orthonormal or masked (so that its entries can be summed to reconstruct the image in the atom region), but becomes an orthonormal basis after being masked. Basically one needs to figure out what weights are used to sum all the $(\mathbf{M}\mathbf{\Phi}_i)$ to get $\mathbf{U}_{i,\text{mask}}$, then use the same weights to instead sum the $\mathbf{\Phi}_i$ without first multiplying them by $\mathbf{M}$. Alban Urvoy pointed out to me that this is easy, the weights are the entries of $\mathbf{V}_{i,\text{mask}}$. That is clear when writing out the expression for

$\mathbf{U}_{i,\text{mask}}$:

$$\begin{aligned}\mathbf{U}_{i,\text{mask}} &= \mathbf{A}_{\text{mask}}\mathbf{V}_{i,\text{mask}} \\ &= \Big[(\mathbf{M}\mathbf{\Phi}_1), (\mathbf{M}\mathbf{\Phi}_2), \ldots, (\mathbf{M}\mathbf{\Phi}_m)\Big] \begin{bmatrix} V_{i,\text{mask},1} \\ V_{i,\text{mask},2} \\ \vdots \\ V_{i,\text{mask},m} \end{bmatrix} \\ &= V_{i,\text{mask},1}(\mathbf{M}\mathbf{\Phi}_1) + V_{i,\text{mask},2}(\mathbf{M}\mathbf{\Phi}_2) + \ldots + V_{i,\text{mask},m}(\mathbf{M}\mathbf{\Phi}_m)\end{aligned} \tag{D.2}$$

where $V_{i,\text{mask},j}$ is entry j of $\mathbf{V}_{i,\text{mask}}$. This shows that the weights used when summing the all the $(\mathbf{M}\mathbf{\Phi}_i)$ to get $\mathbf{U}_{i,\text{mask}}$ are indeed the entries in $\mathbf{V}_{i,\text{mask}}$. Therefore, the unmasked $\mathbf{U}_i$ can be calculated as $\mathbf{A}\mathbf{V}_{i,\text{mask}}$. Note that $\mathbf{A}$, not $\mathbf{A}_{\text{mask}}$, is used in that step. The $\mathbf{U}_i$ should then be "normalized", i.e. divided by the magnitude of $(\mathbf{M}\mathbf{U}_i)$. As before, $\mathbf{U}_i$ corresponding to smaller eigenvalues can be dropped if desired, leaving n different $\mathbf{U}_i$.

It is easy to show that the $\mathbf{U}_i$ constructed this way have the desired properties, namely they are nonzero [11] in the atom region and are orthonormal after being masked. The first of those properties is clear from their construction. Writing out the definition in the same manner used in Eqn. D.2 shows

$$\begin{aligned}\mathbf{U}_i &= \mathbf{A}\mathbf{V}_{i,\text{mask}} \\ &= \Big[\mathbf{\Phi}_1, \mathbf{\Phi}_2, \ldots, \mathbf{\Phi}_m\Big] \begin{bmatrix} V_{i,\text{mask},1} \\ V_{i,\text{mask},2} \\ \vdots \\ V_{i,\text{mask},m} \end{bmatrix} \\ &= V_{i,\text{mask},1}\mathbf{\Phi}_1 + V_{i,\text{mask},2}\mathbf{\Phi}_2 + \ldots + V_{i,\text{mask},m}\mathbf{\Phi}_m\end{aligned} \tag{D.3}$$

The $\mathbf{\Phi}_i$ are *not* multiplied by $\mathbf{M}$ here so there is nothing setting any of the entries to zero. That means that when they are summed the resulting entries for $\mathbf{U}_i$ are generally nonzero. The second property is also easy to show. Note that the set $\{\mathbf{U}_{i,\text{mask}}\}$ is

[11] The entries can be zero, here I just mean to say that they are not forced to be zero by the mask.

orthonormal because it is an eigenbasis. It is also easy to see from Eqn. D.2 and Eqn. D.3 that $\mathbf{M}\mathbf{U}_i = \mathbf{U}_{i,\text{mask}}$. Then it must be that $\{\mathbf{M}\mathbf{U}_i\}$ are orthonormal since that set is the same as $\{\mathbf{U}_{i,\text{mask}}\}$. Thus the constructed $\mathbf{U}_i$ are orthonormal after being masked.

Once the $\mathbf{U}_i$ are obtained, $\mathbf{\Phi}_{\text{ref}}$ can be calculated as

$$\begin{aligned} \mathbf{\Phi}_{\text{ref}} &= \mathbf{U}(\mathbf{M}\mathbf{U})^\intercal \mathbf{\Phi}_{\text{atom}} \\ &= \mathbf{U}\mathbf{U}^\intercal \mathbf{M} \mathbf{\Phi}_{\text{atom}} \end{aligned} \tag{D.4}$$

Note that $\mathbf{M}^\intercal = \mathbf{M}$ because it is diagonal. The rest of the procedure to determine the atom distribution is the same as in the standard approach. The reconstructed image is then $\mathbf{\Gamma}_{\text{ref}} = \mathbf{\Phi}_{\text{ref}} + \bar{\mathbf{\Gamma}}$ and the optical depth for each pixel is calculated as $\log\left(\mathbf{\Gamma}_{\text{ref}}/\mathbf{\Gamma}_{\text{atom}}\right)$ where the division and logarithm are calculated element-wise.

Note that the corrected approach described above requires using the PCA "trick" in which the eigenvectors of $\mathbf{A}_{\text{mask}}\mathbf{A}_{\text{mask}}^\intercal$ are calculated by first calculating the eigenvectors of $\mathbf{A}_{\text{mask}}^\intercal\mathbf{A}_{\text{mask}}$. This is typically helpful because usually $m < jk$. For the somewhat unusual case that $jk < m$, it is faster to calculate the eigenvectors of $\mathbf{A}_{\text{mask}}\mathbf{A}_{\text{mask}}^\intercal$ directly. In that case it may be reasonable to calculate eigenvectors $\tilde{\mathbf{U}}_{i,\text{mask}}$ of $\mathbf{A}_{\text{mask}}\mathbf{A}_{\text{mask}}^\intercal$ directly, then convert them to $\mathbf{V}_{i,\text{mask}}$ via [12] $\mathbf{V}_{i,\text{mask}} \propto \mathbf{A}_{\text{mask}}^\intercal \tilde{\mathbf{U}}_{i,\text{mask}}$. The same procedure as above can then be implemented since it does not matter how the $\mathbf{V}_{i,\text{mask}}$ were calculated; it just matters that they are the eigenvectors of $\mathbf{A}_{\text{mask}}^\intercal\mathbf{A}_{\text{mask}}$. More thought would be necessary to determine if this approach is numerically efficient or if a more efficient approach for the case $jk < m$ exists.

Again the ordering of operations can affect the speed of evaluation and memory required, so proper use of the associative property can speed up evaluation of Eqn. D.4. Evaluating the product two matrices with sizes $a \times b$ and $b \times c$ produces a matrix of size $a \times c$, and each term of the resulting matrix requires evaluating b multiplications and $(b-1)$ additions. Therefore the number of operations scales as $\propto ac(2b-1) \sim abc$.

[12] This is easy to show as well. It was already shown that $\mathbf{U}_i = \mathbf{A}\mathbf{V}_i$, and left-multiplying that by $\mathbf{A}^\intercal$ gives $\mathbf{A}^\intercal\mathbf{U}_i = \mathbf{A}^\intercal\mathbf{A}\mathbf{V}_i = \lambda\mathbf{V}_i$. The second equality there holds true because $\mathbf{V}_i$ is an eigenvector for $\mathbf{A}^\intercal\mathbf{A}$.

Evaluating Eqn. D.4 left-to-right is inefficient. Calculating $\mathbf{U}\mathbf{U}^\intercal$ requires $\sim(jk)^2 n$ operations and produces a $jk \times jk$ matrix. Taking the product of that with $\mathbf{M}$ (which is $jk \times jk$) produces another $jk \times jk$ matrix and requires $\sim(jk)^3$ operations. Finally taking the product of that with $\mathbf{\Phi}_{\text{atom}}$ produces a $jk \times 1$ vector and requires $\sim(jk)^2$ operations. Therefore, performing the product left-to-right in this manner requires $\sim(jk)^3$ operations (assuming $n \ll jk$, which is typically the case), limited by the time taken to multiply two $jk \times jk$ matrices.

On the other hand, evaluating terms in Eqn. D.4 right-to-left is much faster. Evaluating $M\mathbf{\Phi}_{\text{atom}}$ requires $\sim(jk)^2$ operations, then left-multiplying that by $\mathbf{U}^\intercal$ requires $\sim njk$ operations, then left-multiplying that by $\mathbf{U}$ requires $\sim njk$ operations as well. Again assuming $n \ll jk$, the scaling is then dominated by the steps to calculate $\mathbf{M}\mathbf{\Phi}_{\text{atom}}$ which is $\sim(jk)^2$. That is faster than the other ordering by a factor of jk, which can be large. Furthermore, a trick can be used to improve the scaling further by taking advantage of the fact that $\mathbf{M}$ is diagonal. Rather than storing $\mathbf{M}$ as a full 2D matrix, its diagonal can be stored as a 1D column vector. Then taking the element-wise product (not matrix product or dot product) of that 1D column vector and $\mathbf{\Phi}_{\text{atom}}$ gives the same result as multiplying $\mathbf{M}\mathbf{\Phi}_{\text{atom}}$, but only requires jk operations rather than $\sim(jk)^2$. Basically this is a way to skip the multiplications involving the off-diagonal elements of $\mathbf{M}$ which are all zero and so do not contribute to the result. With this trick, the right-to-left evaluation order then scales as $\sim njk$, which is significantly faster than the $\sim(jk)^3$ scaling of the left-to-right evaluation. The scaling for the full calculation is somewhat slower because calculating the product $\mathbf{A}\mathbf{V}_{i,\text{mask}}$ requires $\sim mjk$ operations and $m \geq n$. The time required to calculate the eigenvectors may also be significant, though that will likely depend on the algorithm used and the values of m and jk. In my experience in our lab using thousands of beam images with $\sim 10,000$ pixels each, the $\mathbf{A}\mathbf{V}_{i,\text{mask}}$ calculation is often the slowest step. Again for convenient reference, jk is the total number of pixels in the images (which have j rows and k columns when expressed as 2D arrays), m is the number of beam images, and n is the number of principal components used.

It is worth pointing out that this corrected method properly handles the con-

structed example from Section D.2. There it was pointed out that the reconstructed image should ideally exactly match the "atom" image if there are actually no atoms in the "atom" image. When applied to the hypothetical image $\mathbf{\Gamma}_{\text{atom}} = \bar{\mathbf{\Gamma}} + \mathbf{U}_1$, it was shown that he common approach to masking the region where there may be atoms did not produce the desired result. In particular the reconstructed $\mathbf{\Gamma}_{\text{ref}}$ was shown not to be equal to $\mathbf{\Gamma}_{\text{atom}}$. With the corrected approach, if $\mathbf{\Gamma}_{\text{atom}} = \bar{\mathbf{\Gamma}} + \mathbf{U}_1$, then $\mathbf{\Phi}_{\text{atom}} = \mathbf{U}_1$ and so $\mathbf{M}\mathbf{\Phi}_{\text{atom}} = \mathbf{M}\mathbf{U}_1$. Then by construction $\mathbf{U}^\intercal(\mathbf{M}\mathbf{U}_1) = [1, 0, 0, \ldots, 0]^\intercal$ because $(\mathbf{M}\mathbf{U}_1)$ is orthogonal to all the $\mathbf{U}_i$ except $\mathbf{U}_1$. Left-multiplying by $\mathbf{U}$ then returns $\mathbf{\Phi}_{\text{ref}} = \mathbf{U}_1$ and so $\mathbf{\Gamma}_{\text{ref}} = \bar{\mathbf{\Gamma}} + \mathbf{U}_1 = \mathbf{\Gamma}_{\text{atom}}$. As desired, the reconstructed image exactly matches the "atom" image when using the corrected approach for this constructed example.

An implementation of this corrected approach in Python is available online [158], as is a MATLAB version [159].

Appendix E

List of Symbols

Below is a list of symbols used throughout this thesis. Primarily symbols which are used in many places are included here for convenient reference. Symbols which are only used locally in small sections of this thesis generally are not included.

Δ_{R} The single-photon detuning of the Raman beams from the atomic resonance to the intermediate excited state, as an angular frequency. This quantity is not to be confused with δ_{R}.

$\Delta\mathbf{k}$ Difference between the wavevectors of the Raman beams. Equal to $\mathbf{k}_{\mathrm{R1}} - \mathbf{k}_{\mathrm{R2}}$ where $\mathbf{k}_{\mathrm{R1}}$ is the wavevector of the absorbed (up-leg) photon and $\mathbf{k}_{\mathrm{R2}}$ is the wavevector of the emitted (down-leg) photon.

Δ Detuning of the optical pumping beam from atomic resonance, as an angular frequency.

E_{r} Single-photon recoil energy.

$E_{\mathrm{r}\Delta\mathbf{k}}$ Recoil energy associated with momentum kick of $\hbar\Delta\mathbf{k}$. $E_{\mathrm{r}\Delta\mathbf{k}} = \left(\hbar\Delta\mathbf{k}\right)^2/(2m)$.

$E_{\mathrm{r,trap}}$ Recoil energy for a photon from a trapping beam.

FWHM The full-width-half-max of a Lorentzian.

Γ_{cycle} Cycle rate of Raman cooling, i.e. number of Raman cooling cycles per second. Not to be confused with cooling power which depends on the energy removed per cycle as well.

Γ_{sc} Scattering rate of the optical pumping beam, in number of scatterings per unit time (photons/sec, effectively rad/sec).

$\mathcal{H}$ A Hamiltonian.

KE Kinetic energy.

OD The optical depth measured via absorption imaging.

OD_{p} The peak optical depth in an absorption image.

Ω_{R} Two-photon Raman transition Rabi frequency. Note that this is *not* the generalized Rabi frequency $\sqrt{\Omega^2 + \Delta^2}$.

Ω_{Rabi} Generalized Rabi frequency $\sqrt{\Omega^2 + \Delta^2}$.

Ω_{sc} Rabi frequency (as an angular frequency) of the optical pumping beam.

PSD The peak phase space density of an atomic cloud.

PSD_{c} The peak classical phase space density of an atomic cloud. In other words, this is the PSD calculated for a classical gas with a given atom number, temperature, and trap potential. Effects of Bosonic and Fermionic statistics are ignored when calculating this quantity. Instead the particles are treated as distinguishable classical particles, which often makes this quantity easier to calculate than the true PSD. The value of PSD_{c} is approximately equal to the value of PSD when $\text{PSD} \ll 1$ as quantum statistics does not play a strong role in that regime, and $\text{PSD}_{\text{c}} \sim 1$ at the onset of quantum degeneracy.

T_{ad} "Temperature" measured after releasing atoms from a lattice by extinguishing it slowly and adiabatically. Note that this is *not* equal to the true thermodynamic temperature of the atom while it was in the lattice. See Appendix A for more information.

T_h Temperature in the horizontal direction, as defined by the imaging plane. Related: T_v.

T_loose The temperature along the tightly-confined direction of a lattice site. See T_tight.

T_r The standard recoil temperature.

$T_\mathrm{r}^\mathrm{eff}$ Effective recoil temperature for Raman cooling, accounting for the Raman momentum kick amplitude and branching ratio for spontaneous decay into the initial state.

T_sud "Temperature" measured after releasing atoms from a lattice by extinguishing it suddenly. Note that this is *not* equal to the true thermodynamic temperature of the atom while it was in the lattice. See Appendix A for more information.

T_tight The temperature along the tightly-confined direction of a lattice site. This is primarily used to describe temperatures in 1D and 2D optical lattices, as there are tightly-confined and loosely-confined directions in those lattice sites. If atoms in these lattices are cooled faster than the temperatures in different directions cross-thermalize, then the temperatures in the tightly-confined and loosely-confined directions may differ. Related: T_loose.

T_v Temperature in the vertical direction, as defined by the imaging plane. Related: T_h.

δ_D Doppler shift, as an angular frequency.

δ_R Detuning of the two-photon Raman transition from the final state in lab frame, as an angular frequency. It is equal to the difference in energy between the two photons driving the Raman transition minus the difference in energy between the two ground states. This quantity is not to be confused with Δ_R.

δ_a Similar to δ_R, but as seen in the atom's moving frame rather than in the lab frame. In other words, it is δ_R plus the Doppler shift of the two-photon Raman transition.

δ_z Zeeman shift, as an angular frequency.

k_B Boltzmann's constant, $\approx 1.38 \times 10^{-23}$ kgm^2/(s^{2}K).

$\mathbf{k}_{R1}$ Wavevector of the beam driving the up-leg of a Raman transition.

$\mathbf{k}_{R2}$ Wavevector of the beam driving the down-leg of a Raman transition.

k_{trap} Wavevector of a trapping beam.

$\mathbf{k}$ Wavevector of a beam of light, which has magnitude $2\pi/\lambda$.

$\Delta E_{\mathbf{p}}$ Light shift due to optical pumping beam.

ω_L Angular frequency of a laser beam.

ω_{loose} The angular trap frequency along the tightly-confining direction of a trap. See ω_{tight}.

ω_0 Resonant angular frequency of a transition.

ω_r Single-photon recoil frequency.

ω_{tight} The angular trap frequency along the tightly-confining direction of a trap. This is primarily used to describe a trap frequency in a 1D or 2D lattice, for which there are tightly-confining and loosely-confining directions. Related: ω_{loose}.

ω_t The angular trap frequency of a harmonic trap.

$\mathbf{p}$ Momentum vector of an atom.

$v_{\Delta K=0}$ Critical atomic velocity at which a Raman cooling cycle causes no change in kinetic energy on average. Atoms moving faster than this velocity are cooled on average while atoms moving slower are heated on average.

v_{rms} RMS velocity.

$\mathbf{v}$ Velocity vector of an atom.

Bibliography

[1] Samuel M Brewer, J-S Chen, Aaron M Hankin, Ethan R Clements, Chin-wen Chou, David J Wineland, David B Hume, and David R Leibrandt. Al+ 27 quantum-logic clock with a systematic uncertainty below 10- 18. *Physical review letters*, 123(3):033201, 2019.

[2] Boulder Atomic Clock Optical Network BACON Collaboration. Frequency ratio measurements at 18-digit accuracy using an optical clock network. *Nature*, 591(7851):564–569, 2021.

[3] Tobias Bothwell, Dhruv Kedar, Eric Oelker, John M Robinson, Sarah L Bromley, Weston L Tew, Jun Ye, and Colin J Kennedy. Jila sri optical lattice clock with uncertainty of. *Metrologia*, 56(6):065004, 2019.

[4] CJ Ballance, TP Harty, NM Linke, MA Sepiol, and DM Lucas. High-fidelity quantum logic gates using trapped-ion hyperfine qubits. *Physical review letters*, 117(6):060504, 2016.

[5] John P Gaebler, Ting Rei Tan, Y Lin, Y Wan, R Bowler, Adam C Keith, S Glancy, K Coakley, E Knill, D Leibfried, et al. High-fidelity universal gate set for be 9+ ion qubits. *Physical review letters*, 117(6):060505, 2016.

[6] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell. Observation of Bose-Einstein Condensation in a Dilute Atomic Vapor. *Science*, 269(5221):198–201, 1995.

[7] K. B. Davis, M. O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle. Bose-Einstein Condensation in a Gas of Sodium Atoms. *Phys. Rev. Lett.*, 75:3969–3973, Nov 1995.

[8] M Vengalattore, JM Higbie, SR Leslie, J Guzman, LE Sadler, and DM Stamper-Kurn. High-resolution magnetometry with a spinor bose-einstein condensate. *Physical review letters*, 98(20):200801, 2007.

[9] Immanuel Bloch, Jean Dalibard, and Sylvain Nascimbène. Quantum simulations with ultracold quantum gases. *Nature Physics*, 8(4):267–276, apr 2012.

[10] Iulia M Georgescu, Sahel Ashhab, and Franco Nori. Quantum simulation. *Reviews of Modern Physics*, 86(1):153, 2014.

[11] Kyle S Hardman, Patrick J Everitt, Gordon D McDonald, Perumbil Manju, Paul B Wigley, MA Sooriyabandara, Carlos CN Kuhn, John E Debs, John D Close, and Nicholas P Robins. Simultaneous precision gravimetry and magnetic gradiometry with a bose-einstein condensate: a high precision, quantum sensor. *Physical review letters*, 117(13):138501, 2016.

[12] Dennis Becker, Maike D. Lachmann, Stephan T. Seidel, Holger Ahlers, Aline N. Dinkelaker, Jens Grosse, Ortwin Hellmig, Hauke Müntinga, Vladimir Schkolnik, Thijs Wendrich, André Wenzlawski, Benjamin Weps, Robin Corgier, Tobias Franz, Naceur Gaaloul, Waldemar Herr, Daniel Lüdtke, Manuel Popp, Sirine Amri, Hannes Duncker, Maik Erbe, Anja Kohfeldt, André Kubelka-Lange, Claus Braxmaier, Eric Charron, Wolfgang Ertmer, Markus Krutzik, Claus Lämmerzahl, Achim Peters, Wolfgang P. Schleich, Klaus Sengstock, Reinhold Walser, Andreas Wicht, Patrick Windpassinger, and Ernst M. Rasel. Space-borne Bose–Einstein condensation for precision interferometry. *Nature*, 562(7727):391–395, oct 2018.

[13] Florian Schäfer, Takeshi Fukuhara, Seiji Sugawa, Yosuke Takasu, and Yoshiro Takahashi. Tools for quantum simulation with ultracold atoms in optical lattices. *Nature Reviews Physics*, 2(8):411–425, 2020.

[14] DJ Wineland and H Dehmelt. Proposed $1014\delta\nu/\nu$ laser fluorescence spectroscopy on tl+ mono-ion oscillator iii (side band cooling). *Bull. Am. Phys. Soc*, 20(4):637–637, 1975.

[15] T.W. Hänsch and A.L. Schawlow. Cooling of gases by laser radiation. *Optics Communications*, 13(1):68–69, jan 1975.

[16] David J Wineland, Robert E Drullinger, and Fred L Walls. Radiation-pressure cooling of bound resonant absorbers. *Physical Review Letters*, 40(25):1639, 1978.

[17] Paul D Lett, Richard N Watts, Christoph I Westbrook, William D Phillips, Phillip L Gould, and Harold J Metcalf. Observation of atoms laser cooled below the doppler limit. *Physical review letters*, 61(2):169, 1988.

[18] Jean Dalibard and Claude Cohen-Tannoudji. Laser cooling below the doppler limit by polarization gradients: simple theoretical models. *JOSA B*, 6(11):2023–2045, 1989.

[19] A. Aspect, E. Arimondo, R. Kaiser, N. Vansteenkiste, and C. Cohen-Tannoudji. Laser Cooling below the One-Photon Recoil Energy by Velocity-Selective Coherent Population Trapping. *Physical Review Letters*, 61(7):826–829, aug 1988.

[20] Alain Aspect, Ennio Arimondo, R Kaiser, N Vansteenkiste, and C Cohen-Tannoudji. Laser cooling below the one-photon recoil energy by velocity-selective coherent population trapping: theoretical analysis. *JOSA B*, 6(11):2112–2124, 1989.

[21] Markus Lindberg and Juha Javanainen. Temperature of a laser-cooled trapped three-level ion. *JOSA B*, 3(7):1008–1017, 1986.

[22] Mark Kasevich and Steven Chu. Laser cooling below a photon recoil with three-level atoms. *Physical Review Letters*, 69(12):1741–1744, sep 1992.

[23] Simon Stellmer, Benjamin Pasquiou, Rudolf Grimm, and Florian Schreck. Laser Cooling to Quantum Degeneracy. *Physical Review Letters*, 110(26):263003, jun 2013.

[24] Jiazhong Hu, Alban Urvoy, Zachary Vendeiro, Valentin Crépel, Wenlan Chen, and Vladan Vuletić. Creation of a Bose-condensed gas of ^{87}Rb by laser cooling. *Science (New York, N.Y.)*, 358(6366):1078–1080, nov 2017.

[25] Pablo Solano, Yiheng Duan, Yu-Ting Chen, Alyssa Rudelis, Cheng Chin, and Vladan Vuletić. Strongly correlated quantum gas prepared by direct laser cooling. *Physical review letters*, 123(17):173401, 2019.

[26] Marshall T DePue, Colin McCormick, S Lukman Winoto, Steven Oliver, and David S Weiss. Unity occupation of sites in a 3d optical lattice. *Physical review letters*, 82(11):2262, 1999.

[27] K. Burnett, P. S. Julienne, and K.-A. Suominen. Laser-Driven Collisions between Atoms in a Bose-Einstein Condensed Gas. *Physical Review Letters*, 77(8):1416–1419, aug 1996.

[28] Alban Urvoy, Zachary Vendeiro, Joshua Ramette, Albert Adiyatullin, and Vladan Vuletić. Direct laser cooling to bose-einstein condensation in a dipole trap. *Phys. Rev. Lett.*, 122:203202, May 2019.

[29] Jan Rudolph, Waldemar Herr, Christoph Grzeschik, Tammo Sternke, Alexander Grote, Manuel Popp, Dennis Becker, Hauke Müntinga, Holger Ahlers, Achim Peters, et al. A high-flux bec source for mobile atom interferometers. *New Journal of Physics*, 17(6):065001, 2015.

[30] Gregory A Phelps, Anne Hébert, Aaron Krahn, Susannah Dickerson, Furkan Öztürk, Sepehr Ebadi, Lin Su, and Markus Greiner. Sub-second production of a quantum degenerate gas. *arXiv preprint arXiv:2007.10807*, 2020.

[31] Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P Adams, and Nando De Freitas. Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175, 2015.

[32] Paul B Wigley, Patrick J Everitt, Anton van den Hengel, John W Bastian, Mahasen A Sooriyabandara, Gordon D McDonald, Kyle S Hardman, Ciaron D Quinlivan, P Manju, Carlos CN Kuhn, et al. Fast machine-learning online optimization of ultra-cold-atom experiments. *Scientific reports*, 6(1):1–6, 2016.

[33] Aaron D Tranter, Harry J Slatyer, Michael R Hush, Anthony C Leung, Jesse L Everett, Karun V Paul, Pierre Vernaz-Gris, Ping Koy Lam, Ben C Buchler, and Geoff T Campbell. Multiparameter optimisation of a magneto-optical trap using deep learning. *Nature communications*, 9(1):1–8, 2018.

[34] Vladan Vuletić, Cheng Chin, Andrew J. Kerman, and Steven Chu. Degenerate Raman Sideband Cooling of Trapped Cesium Atoms at Very High Atomic Densities. *Physical Review Letters*, 81(26):5768–5771, dec 1998.

[35] I. Bouchoule, M. Morinaga, C. Salomon, and D. S. Petrov. Cesium gas strongly confined in one dimension: Sideband cooling and collisional properties. *Physical Review A*, 65(3):033402, feb 2002.

[36] EA Curtis, Christopher W Oates, and Leo Hollberg. Quenched narrow-line laser cooling of 40 ca to near the photon recoil limit. *Physical Review A*, 64(3):031403, 2001.

[37] T Binnewies, G Wilpers, U Sterr, F Riehle, J Helmcke, TE Mehlstäubler, Ernst Maria Rasel, and Wolfgang Ertmer. Doppler cooling and trapping on forbidden transitions. *Physical review letters*, 87(12):123002, 2001.

[38] Uwe Sterr, Tomas Binnewies, Carsten Degenhardt, Guido Wilpers, Jürgen Helmcke, and Fritz Riehle. Prospects of doppler cooling on forbidden lines. *JOSA B*, 20(5):985–993, 2003.

[39] Andrew J. Kerman, Vladan Vuletić, Cheng Chin, and Steven Chu. Beyond Optical Molasses: 3D Raman Sideband Cooling of Atomic Cesium to High Phase-Space Density. *Physical Review Letters*, 84(3):439–442, jan 2000.

[40] Zachary Vendeiro, Joshua Ramette, Alyssa Rudelis, Michelle Chong, Josiah Sinclair, Luke Stewart, and Vladan Vuletić. Machine-learning-optimized laser and evaporative cooling to bose einstein condensation in 575 ms. *Work in progress*, 2021.

[41] J. Reichel, F. Bardou, M. Ben Dahan, E. Peik, S. Rand, C. Salomon, and C. Cohen-Tannoudji. Raman Cooling of Cesium below 3 nK: New Approach Inspired by Lévy Flight Statistics. *Physical Review Letters*, 75(25):4575–4578, dec 1995.

[42] H. J. Lee, C. S. Adams, M. Kasevich, and S. Chu. Raman Cooling of Atoms in an Optical Dipole Trap. *Physical Review Letters*, 76(15):2658–2661, apr 1996.

[43] H Perrin, A Kuhn, I Bouchoule, T Pfau, and C Salomon. Raman cooling of spin-polarized cesium atoms in a crossed dipole trap. *Europhysics Letters (EPL)*, 46(2):141–147, apr 1999.

[44] V. Boyer, L. J. Lising, S. L. Rolston, and W. D. Phillips. Deeply subrecoil two-dimensional Raman cooling. *Phys. Rev. A*, 70:043405, Oct 2004.

[45] Dian-Jiun Han, Steffen Wolf, Steven Oliver, Colin McCormick, Marshall T. DePue, and David S. Weiss. 3D Raman Sideband Cooling of Cesium Atoms at High Density. *Physical Review Letters*, 85(4):724–727, jul 2000.

[46] Yichao Yu, Nicholas R Hutzler, Jessie T Zhang, Lee R Liu, Jonathan D Hood, Till Rosenband, and K-K Ni. Motional-ground-state cooling outside the lamb-dicke regime. *Physical Review A*, 97(6):063423, 2018.

[47] Y Castin, H Wallis, and Jean Dalibard. Limit of doppler cooling. *JOSA B*, 6(11):2046–2057, 1989.

[48] HJ Lee and S Chu. Atomic spin polarization by raman cooling. *Physical Review A*, 57(4):2905, 1998.

[49] F Diedrich, JC Bergquist, Wayne M Itano, and DJ Wineland. Laser cooling to the zero-point energy of motion. *Physical Review Letters*, 62(4):403, 1989.

[50] Ch Monroe, DM Meekhof, BE King, Steven R Jefferts, Wayne M Itano, David J Wineland, and P Gould. Resolved-sideband raman cooling of a bound atom to the 3d zero-point energy. *Physical review letters*, 75(22):4011, 1995.

[51] S. E. Hamann, D. L. Haycock, G. Klose, P. H. Pax, I. H. Deutsch, and P. S. Jessen. Resolved-Sideband Raman Cooling to the Ground State of an Optical Lattice. *Physical Review Letters*, 80(19):4149–4152, may 1998.

[52] Kathryn Moler, David S Weiss, Mark Kasevich, and Steven Chu. Theoretical analysis of velocity-selective raman transitions. *Physical Review A*, 45(1):342, 1992.

[53] Claude Cohen-Tannoudji, Jacques Dupont-Roc, and Gilbert Grynberg. *Atom-photon interactions: basic processes and applications*. John Wiley & Sons, Ltd, 1998.

[54] CR Monroe, EA Cornell, CA Sackett, CJ Myatt, and CE Wieman. Measurement of cs-cs elastic scattering at t= 30 μk. *Physical review letters*, 70(4):414, 1993.

[55] NR Newbury, CJ Myatt, and CE Wieman. s-wave elastic collisions between cold ground-state rb 87 atoms. *Physical Review A*, 51(4):R2680, 1995.

[56] Huang Wu and Christopher J Foot. Direct simulation of evaporative cooling. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 29(8):L321, 1996.

[57] Wolfgang Ketterle and N.J. Van Druten. Evaporative Cooling of Trapped Atoms. *Advances In Atomic, Molecular, and Optical Physics*, 37:181–236, jan 1996.

[58] J. I Cirac, M Lewenstein, and P Zoller. Collective laser cooling of trapped atoms. *Europhysics Letters (EPL)*, 35(9):647–652, sep 1996.

[59] Y. Castin, J. I. Cirac, and M. Lewenstein. Reabsorption of Light by Trapped Atoms. *Physical Review Letters*, 80(24):5305–5308, jun 1998.

[60] L Santos, Z Idziaszek, J I Cirac, and M Lewenstein. Laser-induced condensation of trapped bosonic gases. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 33(19):4131–4148, oct 2000.

[61] Steffen Wolf, Steven J. Oliver, and David S. Weiss. Suppression of Recoil Heating by an Optical Lattice. *Physical Review Letters*, 85(20):4249–4252, nov 2000.

[62] J. Schuster, A. Marte, S. Amtage, B. Sang, G. Rempe, and H. C. W. Beijerinck. Avalanches in a Bose-Einstein Condensate. *Physical Review Letters*, 87(17):170404, oct 2001.

[63] TA Savard, KM O'hara, and JE Thomas. Laser-noise-induced heating in far-off resonance optical traps. *Physical Review A*, 56(2):R1095, 1997.

[64] Rudolf Grimm, Matthias Weidemüller, and Yurii B Ovchinnikov. Optical dipole traps for neutral atoms. *Advances in atomic, molecular, and optical physics*, 42:95–170, 2000.

[65] H. Schmaljohann, M. Erhard, J. Kronjäger, M. Kottke, S. van Staa, L. Cacciapuoti, J. J. Arlt, K. Bongs, and K. Sengstock. Dynamics of F = 2 Spinor Bose-Einstein Condensates. *Physical Review Letters*, 92(4):040402, jan 2004.

[66] M.-O. Mewes, M. R. Andrews, N. J. van Druten, D. M. Kurn, D. S. Durfee, and W. Ketterle. Bose-Einstein Condensation in a Tightly Confining dc Magnetic Trap. *Physical Review Letters*, 77(3):416–419, jul 1996.

[67] E. A. Burt, R. W. Ghrist, C. J. Myatt, M. J. Holland, E. A. Cornell, and C. E. Wieman. Coherence, Correlations, and Collisions: What One Learns about Bose-Einstein Condensates from Their Decay. *Physical Review Letters*, 79(3):337–340, jul 1997.

[68] Andrew James Kerman. *Raman sideband cooling and cold atomic collisions in optical lattices*. PhD thesis, Stanford University, 2002.

[69] H.J. Metcalf and P. van der Straten. *Laser Cooling and Trapping*. Graduate Texts in Contemporary Physics. Springer New York, 2012.

[70] Daniel A. Steck. Quantum and Atom Optics. *available online at http://steck.us/teaching (revision 0.13.4, 24 September 2020)*, sep 2020.

[71] Daniel A. Steck. Rubidium 87 D Line Data. *available online at http://steck.us/alkalidata (revision 2.2.1, 21 November 2019)*, nov 2019.

[72] David J Wineland and Wayne M Itano. Laser cooling of atoms. *Physical Review A*, 20(4):1521, 1979.

[73] PP Craig, JG Dash, AD McGuire, D Nagle, and RR Reiswig. Nuclear resonance absorption of gamma rays in ir 191. *Physical Review Letters*, 3(5):221, 1959.

[74] Igor Teper. *Ultracold atoms in microfabricated magnetic traps near surfaces and inside optical resonators*. PhD thesis, Stanford University, 2006.

[75] Yu-Ju Lin. *Ultracold atoms magnetically trapped inside an optical resonator on a microfabricated chip*. PhD thesis, Stanford University, 2006.

[76] Ian Daniel Leroux. *Squeezing collective atomic spins with an optical resonator*. PhD thesis, Massachusetts Institute of Technology, 2011.

[77] Monika Helene Schleier-Smith. *Cavity-enabled spin squeezing for a quantum-enhanced atomic clock*. PhD thesis, Massachusetts Institute of Technology, 2011.

[78] Hao Zhang. *Generation and Measurement of Entangled Atomic Ensembles with an Optical Cavity*. PhD thesis, Massachusetts Institute of Technology, 2015.

[79] Jiazhong Hu. *Light-induced many-body correlations in ultracold gases*. PhD thesis, Massachusetts Institute of Technology, 2017.

[80] OMJH Morsch, JH Müller, M Cristiani, Donatella Ciampini, and Ennio Arimondo. Bloch oscillations and mean-field effects of bose-einstein condensates in 1d optical lattices. *Physical Review Letters*, 87(14):140402, 2001.

[81] Antoine Browaeys, H Häffner, Callum McKenzie, SL Rolston, Kristian Helmerson, and WD Phillips. Transport of atoms in a quantum conveyor belt. *Physical Review A*, 72(5):053605, 2005.

[82] Alexander Franzen. Componentlibrary: a vector graphics library for illustrations of optics experiments. http://www.gwoptics.org/ComponentLibrary/. Accessed: 9 July 2021.

[83] C Wieman and Th W Hänsch. Doppler-free laser polarization spectroscopy. *Physical Review Letters*, 36(20):1170, 1976.

[84] Leonardo Ricci, Matthias Weidemüller, Tilman Esslinger, Andreas Hemmerich, Claus Zimmermann, Vladan Vuletic, Wolfgang König, and Theodor W Hänsch. A compact grating-stabilized diode laser system for atomic physics. *Optics communications*, 117(5-6):541–549, 1995.

[85] Philip Thomas Starkey, Christopher James Billington, Shaun Phillip Johnstone, Martijn Jasperse, K Helmerson, Lincoln David Turner, and Russell Paul Anderson. A scripted control system for autonomous hardware-timed experiments. *Review of Scientific Instruments*, 84(8):085111, 2013.

[86] Christopher James Billington. *State-dependent forces in cold quantum gases*. PhD thesis, Monash University, 2018.

[87] Philip Thomas Starkey. *A software framework for control and automation of precisely timed experiments*. PhD thesis, Monash University, 2019.

[88] JL Hall, L Hollberg, T Baer, and HG Robinson. Optical heterodyne saturation spectroscopy. *Applied Physics Letters*, 39(9):680–682, 1981.

[89] Akio Kawasaki. *Towards spin squeezed ^{171}Yb atomic clock beyond the standard quantum limit*. PhD thesis, Massachusetts Institute of Technology, 2017.

[90] Boris Braverman. *Cavity quantum electrodynamics with ensembles of ytterbium-171*. PhD thesis, Massachusetts Institute of Technology, 2018.

[91] Paul Horowitz and Winfield Hill. *The Art of Electronics*. Cambridge university press Cambridge, 1989.

[92] Russell Paul Anderson, Ethan Payne, Josh Morris, Chris Billington, Philip Thomas Starkey, and Christopher James Billington. analysislib-mloop. https://github.com/rpanderson/analysislib-mloop.

[93] Grace H Zhang, Boris Braverman, Akio Kawasaki, and Vladan Vuletić. Fast compact laser shutter using a direct current motor and three-dimensional printing. *Review of Scientific Instruments*, 86(12), 2015.

[94] B. DeMarco and D. S. Jin. Onset of Fermi Degeneracy in a Trapped Atomic Gas. *Science*, 285(5434):1703–1706, 1999.

[95] M. D. Barrett, J. A. Sauer, and M. S. Chapman. All-Optical Formation of an Atomic Bose-Einstein Condensate. *Physical Review Letters*, 87(1):010404, jun 2001.

[96] Tetsuya Ido, Yoshitomo Isoya, and Hidetoshi Katori. Optical-dipole trapping of Sr atoms at a high phase-space density. *Physical Review A*, 61(6):061403(R), may 2000.

[97] T. W. Hijmans, G. V. Shlyapnikov, and A. L. Burin. Influence of radiative interatomic collisions on dark-state cooling. *Physical Review A*, 54(5):4332–4338, nov 1996.

[98] Jahn Rührig, Tobias Bäuerle, Axel Griesmaier, and Tilman Pfau. High efficiency demagnetization cooling by suppression of light-assisted collisions. *Optics Express*, 23(5):5596, mar 2015.

[99] Franco Dalfovo, Stefano Giorgini, Lev P. Pitaevskii, and Sandro Stringari. Theory of Bose-Einstein condensation in trapped gases. *Reviews of Modern Physics*, 71(3):463–512, apr 1999.

[100] Wolfgang Ketterle and N. J. van Druten. Bose-Einstein condensation of a finite number of particles trapped in one or three dimensions. *Physical Review A*, 54(1):656–660, jul 1996.

[101] Robert P. Smith, Robert L. D. Campbell, Naaman Tammuz, and Zoran Hadzibabic. Effects of Interactions on the Critical Temperature of a Trapped Bose Gas. *Physical Review Letters*, 106(25):250403, jun 2011.

[102] Nicholas P. Robins, Cristina Figl, Matthew Jeppesen, Graham R. Dennis, and John D. Close. A pumped atom laser. *Nature Physics*, 4(9):731–736, sep 2008.

[103] J. F. Barry, D. J. McCarron, E. B. Norrgard, M. H. Steinecker, and D. DeMille. Magneto-optical trapping of a diatomic molecule. *Nature*, 512(7514):286–289, aug 2014.

[104] Loïc Anderegg, Benjamin L. Augenbraun, Yicheng Bao, Sean Burchesky, Lawrence W. Cheuk, Wolfgang Ketterle, and John M. Doyle. Laser cooling of optically trapped molecules. *Nature Physics*, 14(9):890–893, sep 2018.

[105] R. J. C Spreeuw, T Pfau, U Janicke, and M Wilkens. Laser-like Scheme for Atomic-Matter Waves. *Europhysics Letters (EPL)*, 32(6):469–474, nov 1995.

[106] Maxim Olshanii, Yvan Castin, and Jean Dalibard. A model for an atom laser. In Inguscio M., Allegrini M., and Sasso A., editors, *Proc. XII Conf. on Laser Spectroscopy*, pages 7–12. World Scientific, 1996.

[107] H. Jelassi, B. Viaris de Lesegno, and L. Pruvost. Photoassociation spectroscopy of $^{87}Rb_2$ $(5s_{1/2} + 5p_{1/2})$ $0g^-$ long-range molecular states: Analysis by Lu-Fano graph and improved LeRoy-Bernstein formula. *Physical Review A*, 73(3):032501, mar 2006.

[108] H. Jelassi, B. Viaris de Lesegno, and L. Pruvost. Photoassociation spectroscopy of $^{87}Rb_2$ $(5s_{1/2} + 5p_{1/2})$ $0u^+$ long-range molecular states: Coupling with the $(5s_{1/2}+5p_{3/2})$ $0u^+$ series analyzed using the Lu-Fano approach. *Physical Review A*, 74(1):012510, jul 2006.

[109] Haikel Jelassi and Laurence Pruvost. Weakly bound $^{87}Rb_2$ $(5s_{1/2} + 5p_{1/2})$ $1g$ molecules: Hyperfine interaction and LeRoy-Bernstein analysis including linear and nonlinear terms. *Physical Review A*, 89(3):032514, mar 2014.

[110] Joschka Wolf, Markus Deiß, Artjom Krükow, Eberhard Tiemann, Brandon P Ruzic, Yujun Wang, José P D'Incao, Paul S Julienne, and Johannes Hecker Denschlag. State-to-state chemistry for three-body recombination in an ultracold rubidium gas. *Science (New York, N.Y.)*, 358(6365):921–924, nov 2017.

[111] Chun-Chia Chen, Rodrigo González Escudero, Jiří Minář, Benjamin Pasquiou, Shayne Bennetts, and Florian Schreck. An ultracold bose-einstein condensate in steady state. *arXiv preprint arXiv:2012.07605*, 2020.

[112] Simon Stellmer, Rudolf Grimm, and Florian Schreck. Production of quantum-degenerate strontium gases. *Phys. Rev. A*, 87:013611, Jan 2013.

[113] Richard Roy, Alaina Green, Ryan Bowler, and Subhadeep Gupta. Rapid cooling to quantum degeneracy in dynamically shaped atom traps. *Physical Review A*, 93(4):043403, 2016.

[114] Nicholas R Hutzler, Lee R Liu, Yichao Yu, and Kang-Kuen Ni. Eliminating light shifts for single atom trapping. *New Journal of Physics*, 19(2):023007, 2017.

[115] Ippei Nakamura, Atsunori Kanemura, Takumi Nakaso, Ryuta Yamamoto, and Takeshi Fukuhara. Non-standard trajectories found by machine learning for evaporative cooling of 87 rb atoms. *Optics express*, 27(15):20435–20443, 2019.

[116] Adam J Barker, Harry Style, Kathrin Luksch, Shinichi Sunami, David Garrick, Felix Hill, Christopher J Foot, and Elliot Bentine. Applying machine learning optimization methods to the production of a quantum gas. *Machine Learning: Science and Technology*, 1(1):015007, 2020.

[117] ET Davletov, VV Tsyganok, VA Khlebnikov, DA Pershin, DV Shaykin, and AV Akimov. Machine learning for achieving bose-einstein condensation of thulium atoms. *Physical Review A*, 102(1):011302, 2020.

[118] Yadong Wu, Zengming Meng, Kai Wen, Chengdong Mi, Jing Zhang, and Hui Zhai. Active learning approach to optimization of experimental control. *Chinese Physics Letters*, 37(10):103201, 2020.

[119] Jasper Snoek, Oren Rippel, Kevin Swersky, Ryan Kiros, Nadathur Satish, Narayanan Sundaram, Md. Mostofa Ali Patwary, Prabhat, and Ryan P. Adams. Scalable bayesian optimization using deep neural networks, 2015.

[120] Matthias Weidemüller, Tilman Esslinger, Maxim A Ol'shanii, Andreas Hemmerich, and Theodor W Hänsch. A novel scheme for efficient cooling below the photon recoil limit. *EPL (Europhysics Letters)*, 27(2):109, 1994.

[121] D Boiron, C Triché, DR Meacher, P Verkerk, and G Grynberg. Three-dimensional cooling of cesium atoms in four-beam gray optical molasses. *Physical Review A*, 52(5):R3425, 1995.

[122] DS Petrov, GV Shlyapnikov, and JTM Walraven. Regimes of quantum degeneracy in trapped 1d gases. *Physical Review Letters*, 85(18):3745, 2000.

[123] Isabelle Bouchoule, Karen V Kheruntsyan, and GV Shlyapnikov. Interaction-induced crossover versus finite-size condensation in a weakly interacting trapped one-dimensional bose gas. *Physical Review A*, 75(3):031606, 2007.

[124] Peter Krüger, S Hofferberth, IE Mazets, Igor Lesanovsky, and Jörg Schmiedmayer. Weakly interacting bose gas in the one-dimensional limit. *Physical review letters*, 105(26):265302, 2010.

[125] Toshiya Kinoshita, Trevor Wenger, and David S. Weiss. Observation of a One-Dimensional Tonks-Girardeau Gas. *Science*, 305(5687), 2004.

[126] Belén Paredes, Artur Widera, Valentin Murg, Olaf Mandel, Simon Fölling, Ignacio Cirac, Gora V. Shlyapnikov, Theodor W. Hänsch, and Immanuel Bloch. Tonks–Girardeau gas of ultracold atoms in an optical lattice. *Nature*, 429(6989):277–281, may 2004.

[127] A. Görlitz, J. M. Vogels, A. E. Leanhardt, C. Raman, T. L. Gustavson, J. R. Abo-Shaeer, A. P. Chikkatur, S. Gupta, S. Inouye, T. Rosenband, and W. Ketterle. Realization of Bose-Einstein Condensates in Lower Dimensions. *Physical Review Letters*, 87(13):130402, sep 2001.

[128] Helene Perrin, Axel Kuhn, Isabelle Bouchoule, and Christophe Salomon. Sideband cooling of neutral atoms in a far-detuned optical lattice. *EPL (Europhysics Letters)*, 42(4):395, 1998.

[129] Meung Ho Seo, Sooyoung Park, and D Cho. Relaxation of atomic temperature anisotropy in a one-dimensional optical lattice enhanced by dynamic control of the aspect ratio. *Physical Review A*, 101(4):043611, 2020.

[130] DS Petrov and GV Shlyapnikov. Interatomic collisions in a tightly confined bose gas. *Physical Review A*, 64(1):012706, 2001.

[131] T Grünzweig, A Hilliard, M McGovern, and MF Andersen. Near-deterministic preparation of a single atom in an optical microtrap. *Nature Physics*, 6(12):951–954, 2010.

[132] MO Brown, T Thiele, C Kiehl, T-W Hsu, and CA Regal. Gray-molasses optical-tweezer loading: controlling collisions for scaling atom-array assembly. *Physical Review X*, 9(1):011057, 2019.

[133] Maxim Olshanii and David Weiss. Producing bose-einstein condensates using optical lattices. *Physical review letters*, 89(9):090404, 2002.

[134] Anthony E. Siegman. *Lasers*. University Science Books, 1986.

[135] Francesco Ferri, Sébastien Garcia, Mohamed Baghdad, Jakob Reichel, and Romain Long. Mapping optical standing-waves of an open-access fabry–perot cavity with a tapered fiber. *Review of Scientific Instruments*, 91(3):033104, 2020.

[136] Ivan Favero, Sebastian Stapfner, David Hunger, Philipp Paulitschke, Jakob Reichel, Heribert Lorenz, Eva M Weig, and Khaled Karrai. Fluctuating nanomechanical system in a high finesse optical microcavity. *Optics express*, 17(15):12813–12820, 2009.

[137] Arnaud Gloppe, Pierre Verlot, Eva Dupont-Ferrier, Alessandro Siria, Philippe Poncharal, Guillaume Bachelier, Pascal Vincent, and Olivier Arcizet. Bidimensional nano-optomechanics and topological backaction in a non-conservative radiation force field. *Nature nanotechnology*, 9(11):920–926, 2014.

[138] Francesco Fogliano, Benjamin Besga, Antoine Reigue, Philip Heringlake, Laure Mercier de Lépinay, Cyril Vaneph, Jakob Reichel, Benjamin Pigeau, and Olivier Arcizet. Mapping the cavity optomechanical interaction with subwavelength-sized ultrasensitive nanomechanical force sensors. *Physical Review X*, 11(2):021009, 2021.

[139] JD Thompson, BM Zwickl, AM Jayich, Florian Marquardt, SM Girvin, and JGE Harris. Strong dispersive coupling of a high-finesse cavity to a micromechanical membrane. *Nature*, 452(7183):72–75, 2008.

[140] Aline Vochezer, Tobias Kampschulte, Klemens Hammerer, and Philipp Treutlein. Light-mediated collective atomic motion in an optical lattice coupled to a membrane. *Physical review letters*, 120(7):073602, 2018.

[141] Gerhard R Guthöhrlein, Matthias Keller, Kazuhiro Hayasaka, Wolfgang Lange, and Herbert Walther. A single ion as a nanoscopic probe of an optical field. *Nature*, 414(6859):49–51, 2001.

[142] Sebastien Garcia, Francesco Ferri, Jakob Reichel, and Romain Long. Overlapping two standing waves in a microcavity for a multi-atom photon interface. *Optics express*, 28(10):15515–15528, 2020.

[143] Ziyuan Li, GE Stedman, and HR Bilger. Asymmetric response profile of a scanning fabry-perot interferometer. *Optics communications*, 100(1-4):240–246, 1993.

[144] Jérôme Poirson, Fabien Bretenaker, Marc Vallet, and Albert Le Floch. Analytical and experimental study of ringing effects in a fabry–perot cavity. application to the measurement of high finesses. *JOSA B*, 14(11):2811–2817, 1997.

[145] Sebastian Weber, Christoph Tresp, Henri Menke, Alban Urvoy, Ofer Firstenberg, Hans Peter Büchler, and Sebastian Hofferberth. Tutorial: Calculation of Rydberg interaction potentials. *J. Phys. B: At. Mol. Opt. Phys.*, 50(13):133001, 2017.

[146] Maximilian Harlander, Michael Brownnutt, Wolfgang Hänsel, and Rainer Blatt. Trapped-ion probing of light-induced charging effects on dielectrics. *New Journal of Physics*, 12(9):093035, 2010.

[147] Shannon X Wang, Guang Hao Low, Nathan S Lachenmyer, Yufei Ge, Peter F Herskind, and Isaac L Chuang. Laser-induced charging of microfabricated ion traps. *Journal of Applied Physics*, 110(10):104901, 2011.

[148] DTC Allcock, TP Harty, HA Janacek, NM Linke, CJ Ballance, AM Steane, DM Lucas, RL Jarecki, SD Habermehl, MG Blain, et al. Heating rate and electrode charging measurements in a scalable, microfabricated, surface-electrode ion trap. *Applied Physics B*, 107(4):913–919, 2012.

[149] Arne Härter, Artjom Krükow, Andreas Brunner, and J Hecker Denschlag. Long-term drifts of stray electric fields in a paul trap. *Applied Physics B*, 114(1):275–281, 2014.

[150] Jérôme Lodewyck, Michal Zawada, Luca Lorini, Mikhail Gurov, and Pierre Lemonde. Observation and cancellation of a perturbing dc stark shift in strontium optical lattice clocks. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 59(3):411–415, 2012.

[151] SE Pollack, MD Turner, S Schlamminger, CA Hagedorn, and JH Gundlach. Charge management for gravitational-wave observatories using uv leds. *Physical Review D*, 81(2):021101, 2010.

[152] A Kastberg, William D Phillips, SL Rolston, RJC Spreeuw, and Poul S Jessen. Adiabatic cooling of cesium to 700 nk in an optical lattice. *Physical review letters*, 74(9):1542, 1995.

[153] Robert A Pease. Understand capacitor sookage to optimize analog systems. *END*, 1982.

[154] Wolfgang Ketterle and Martin W Zwierlein. Making, probing and understanding ultracold fermi gases. *La Rivista del Nuovo Cimento*, 31(5):247–422, 2008.

[155] Gal Ness, Anastasiya Vainbaum, Constantine Shkedrov, Yanay Florshaim, and Yoav Sagi. Single-exposure absorption imaging of ultracold atoms using deep learning. *Physical Review Applied*, 14(1):014011, 2020.

[156] Matthew Turk and Alex Pentland. Eigenfaces for recognition. *Journal of cognitive neuroscience*, 3(1):71–86, 1991.

[157] Xiaolin Li, Min Ke, Bo Yan, and Yuzhu Wang. Reduction of interference fringes in absorption imaging of cold atom cloud using eigenface method. *Chinese Optics Letters*, 5(3):128–130, 2007.

[158] Zachary Vendeiro. absorption_image_processor. https://github.com/zakv/absorption_image_processor.

[159] Zachary Vendeiro. Absorptionimageprocessing_matlab. https://github.com/zakv/AbsorptionImageProcessing.