ATOMIC TRANSITION-BASED PRECISION AND ACCURACY IN MICRO- AND NANOSCALE SYSTEMS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Norimasa Yoshimizu

February 2010

© 2010 Norimasa Yoshimizu

ALL RIGHTS RESERVED

ATOMIC TRANSITION-BASED PRECISION AND ACCURACY IN MICRO- AND NANOSCALE SYSTEMS

Norimasa Yoshimizu, Ph.D.

Cornell University 2010

This thesis describes work on using atomic transition-based stabilization to yield high precision, high stability systems for micro- and nano-scale systems. After discussing the fundamentals of atomic precision and control loops, compact low power radio frequency plasmas are described for use in passive rubidium atomic clocks. A self-powered implementation is then described, using a radioactive isotope as the power source. Next, an atomic beat clock is described as a low power atomic frequency source. The ultimate precision in laser frequency stability due to quantum phase fluctuations is discussed. Next, a nanometrology optical ruler imaging system is described which uses an atomically frequency-stabilized laser. Simulations and experimental results are described. Finally, scanning probe nanofabrication of highly ordered pyrolytic graphite using an electrochemical etching process is described, where a feedback control system allows for fabrication of a large array.

BIOGRAPHICAL SKETCH

Norimasa Yoshimizu won an ovarian lottery ticket to Norio and Sadako Yoshimizu in São Paulo, Brazil in 1981. After living in Yokohama, Japan; Sylvania, Ohio; Cupertino, California; and Bloomfield Hills, Michigan he received his Bachelor of Science in physics at Harvey Mudd College in Claremont, California in 2003. He subsequently moved to Cornell University in Ithaca, New York to work towards a Doctor of Philosophy in Applied Physics working for Professor Amit Lal.

Otousama, Okaasama, Maria, and Melissa; and kazoku no minna.

ACKNOWLEDGEMENTS

Professor Amit Lal, for his dedication to, love of, and breadth of knowledge of science and engineering; his sleight of hand; his open mind. There's no place like SonicMEMS...

Professor Clifford R Pollock, for his patient advice and uncompromising honesty.

Professor Sunil Bhave, for his uncountably infinite enthusiasm and support.

Professor David Hammer for his helpful advice.

All the members of SonicMEMS who have put up with and helped me: Abhijit Sathaye, Shyi-Herng Kan, Hui Li, Rajesh Duggirala, Kursad Araz, Hang Guo, Abhishek Ramkumar, Steven Tin, Alper Bozkurt, Siva Prasad, Janet Shen, Kwame Amponsah, Larry Lu, Yue Shi: and especially Serhan Ardanuç for his electrical engineering help and Shankar Radhakrishnan for teaching me most everything I know about microfabrication.

All the members of OxideMEMS who have put up with and helped me: Hengky Chandrahalim, Dana Weinstein, Eugene Hwang, and Tiffany Cheng.

My other collaborators on TBN, Professors Chung-Hoon Lee and Jiwoong Park, Bryan Hicks, and Mark Levendorf.

The APES, to whom I always look up to as a scientist and engineer, and became wonderful friends: Mike, Leon, Jacob, Alex & Judy, Mark, Yoshi, Chris, Jen, Lena, Sarah,...

Justin, Scott, Kevin, Andrew, and Makoto.

David, Julie, and John Hoch; Steve, Kati, Daniel, and Anna Gale; and Mrs. Seiko.

The technical, administrative, and custodial staff of Phillips Hall and the CNF.

Charles Martin Hall and Paul Héroult, and others well-known or otherwise whose work allowed me to make an epsilon of returned contribution.

The taxpayers of the United States.

TABLE OF CONTENTS

Biographical Sketch iii
Dedication iv
Acknowledgements v
Table of Contents vii
List of Tables ix
List of Figures x

1 Introduction **1**

2 Precision and accuracy in atomic systems **6**

2.1 Quantization in atoms 6

2.1.1 Loss of precision in atomic systems 8

2.1.2 Saturation spectroscopy 11

2.1.3 Hole burning experiments 19

2.1.4 Hole burning with low power VCSELs 26

3 Self wavelength calibrated optical sources for atomic clocks **32**

3.1 Compact, low power radio frequency rubidium plasma for passive rubidium atomic clocks 32

3.1.1 Spectral response of compact, low power radio frequency rubidium plasmas 34

3.2 Self-powered Self-Calibrated Photon Source 43

3.2.1 Alternative sources and applications 51

3.3 Focusing by thin magnetic plates: design by evolutionary algorithms 53

4 Low-power Beat Clock **59**

5 Nano Optical Ruler Imaging System for nanometrology 67
5.1 Image processing 75
5.1.1 Reducing noise in analogue to digital conversion 75
5.2 Aperture array design 84
5.2.1 Diffraction as a Fourier transform 84
5.2.2 Quasiperiodic patterns 86
5.3 Image registration 93
5.4 Empirical results 106
6 Precision patterning of graphite 113
6.1 Graphite etching process 115
6.2 Precision writing of a line array 120
6.3 Conclusion 125
A Control loops 126
A.1 Control loop stability 126
A.1.1 PID control 131
B Nonperiodic aperture arrays 133
Bibliography 138

LIST OF TABLES

3.1 Coil types for RF plasma drive of glass Rb cells. N refers to number of winds, d to the diameter or side length, L to the length along the glass tube, and OIF to the optimum ignition frequency – the frequency of ignition with lowest applied power. 35

3.2 Identification of spectral lines see in Figure 3.2 [56, 55]. 37

3.3 Spectral lines identified between 810 nm and 925 nm[55]. 1: Due to proximity of adjacent lines, identification was also based on published line intensities. 50

LIST OF FIGURES

2.1 Expected calculated transmission spectrum for our VCSEL setup, as calculated in the text. The saturated spectrum shows the three spectral holes, corresponding to the three resonances that were assumed. Two of the spectral holes overlap. 18

2.2 Experimental setup for saturation spectroscopy using external cavity laser. ND filters act as beam splitters, but are selected to allow different probe and pump powers. 19

2.3 Saturation spectroscopy of ^{85}Rb with a 42 mW pump beam and a 200 μW probe beam, using setup in Fig. 2.2. As in Fig. 2.1, two of the holes overlap each other due to their broad linewidths. . . . 20

2.4 Similar data as in Fig. 2.3, but at 9 mW probe beam with a 210 μW probe beam; the structure is for the excitation from the F=3 ground state. The structure is better resolved, and the cross-over resonances are visible. The theoretical frequency splittings are from [7]. The experimental frequencies are found by a linear fit of the spectral hole frequencies, finding general agreement. 22

2.5 Same setup as in Fig. 2.4, but for excitation from the F=2 ground state. The F=2 and F=3 ground state hyperfine splitting in ^{85}Rb is 3.036 GHz. The cross over resonances are not visible here. 23

2.6 Demonstration of low power saturation spectroscopy, especially relevant for the use of low power semiconductor lasers. The pump beam power is decreased to 190 μW, and the probe beam is reduced to 240 nW. Curve shows absorption from the F=3 ground state. Due to the lower beam powers, the atomic vapor temperature is also reduced. As in Fig. 2.4, the cross over resonances are visible. 24

2.7 The spectral hole can be used to effectively narrow a laser linewidth as well. In this Figure, the pump probe is 1.9 mW but the probe beam is only 480 nW. The probe beam is weak. It is almost completely absorbed by the atomic vapor, except at the narrow 6 MHz spectral hole. This spectrum is captured by a Fabry-Perot cavity (Thor Labs SA200). 25

2.8 Experimental setup for hole burning with VCSELs. A linear polarizer and quarter wave plate act as a directional filter, so that laser power does not reflect act to the VCSEL. Due to the low cavity power of the VCSEL, their oscillation is especially sensitive to extra noise. 27

2.9 Experimental setup for hole burning with VCSELs, excitation from F=2 ground state. These data only show a single spectral hole. This is due to the > 100 MHz linewidth of the VCSEL, so that the observed transmission is a convolution of the two and appears to be one single spectral hole. 28

2.10 Similar data as in Fig. 2.9, but from the F=3 ground state. Just as in Fig. 2.9, the fine structure appears to be a single spectral hole due to the large linewidth of the VCSEL. 29

2.11 In Fig. 2.8, the linear polarizer and quarter wave plate were used to prevent feedback to the VCSEL to prevent chaotic oscillations in the lasing cavity. By using a piezo-mounted mirror, the feedback can be used to narrow the VCSEL linewidth. Here, the linewidth has been narrowed from > 100 MHz to 15 MHz, measured using a Fabry Perot cavity. 30

2.12 Setup for narrowing VCSEL linewidth. The piezo-mounted mirror is necessary to carefully adjust the phase of the feedback to the VCSEL. The ND filter adjusts the feedback power. 31

3.1 Images of miniature rubidium plasma and the bulb. a. rubidium plasma excited by an RF coil, the purplish color showing the optical emissions from excited rubidium atoms. b. The glass bulbs were filled with rubidium and some xenon as a buffer gas, custom made by the glass blowers at FEI, Inc. 34

3.2 Spectra taken of the RF excited miniature bulbs of rubidium gas. The bulbs are filled with both rubidium and xenon. The dominance of one set of excitation over the other is seen as a function of the frequency of the RF drive. At 298 MHz, the rubidium lines are dominant in the spectrum. As the RF drive frequency increases past 300 MHz, the xenon spectra begins to dominate until the rubidium lines are not visible at all. 36

3.3 Ratio of the intensities of the spectral lines as a function of the RF drive frequency (see Fig. 3.2). The red line shows the ratio of the 882 nm to the 823 nm, dominant lines of xenon. This ratio is nearly flat; at lower frequencies, the xenon lines are barely visible which causes the ratio to be variable. The ratio of the 780 nm to the 795 nm lines are shown in blue, the lines of rubidium. 38

3.4 Ratio of the line intensities of 823 nm to 780 nm, i.e. ratio of xenon to rubidium (see Fig. 3.2). Near 300 to 350 MHz, there is a sharp transition from rubidium domination to xenon. 39

3.5 Spectra of the miniature bulb rubidium plasma as a function of the RF power. At lower powers at 180 MHz, both the rubidium and a weak set of xenon line are visible. At higher powers, the xenon lines get weaker and the rubidium spectral lines dominate the spectrum. 40

3.6 Data similar to Fig. 3.4: ratio of same-atom spectral lines, but now as a function of RF drive power. See Fig. 3.5). As seen in the frequency dependence, the ratio of the lines of rubidium in red and the ratio of the lines of xenon in blue are fairly constant over the range of applied RF power. 41

3.7 Data similar to Fig. 3.3: ratio of 780 nm (rubidium) to 823 nm (xenon), but now as a function of RF drive power (see Fig. 3.5. At low powers, the xenon spectral lines are visible with respect to the rubidium lines, but as the power is increased the rubidium lines are dominant as the line intensities of xenon decrease. 42

3.8 Observed xenon scintillation contour plot of direct images at (a) 20 kPa and at (b) 100 kPa. The ^{63}Ni is located at the bottom of each plot. They are qualitatively similar, but the higher xenon density at 100 kPa increases the xenon-electron interactions per unit distance, leading to a larger gradient of the optical intensity in (b). 45

3.9 Optical intensity profile at 100 kPa, as indicated by sectional arrows in Fig. 3.8. Line shows result from Spencer's theory using isotropic angular distribution of β decay. 48

3.10 Spectra at 20, 50, and 80 kPa. Insets show less prominent peaks between 890 nm and 920 nm. Six peaks are labeled by the letters at the top of the figure. 49

3.11 Excitation of CO_2 by ^{63}Ni. 52

3.12 Excitation of GaAs by ^{63}Ni, and spectrum. The broad spectrum stops short of the band gap of 870 nm (at 300K). 53

3.13 Geometry of thin film magnet focusing of ^{63}Ni beta decay. A rotationally symmetric curve is cut into the ferromagnet, which is magnetized along the direction of propagation. 55

3.14 Results of a Monte Carlo nonrelativistic finite difference time domain model of ^{63}Ni beta decay to simulate the spatial focusing using rotationally symmetric shapes cut into thin ferromagnetic films. All four runs show qualitatively similar results. First, there are about 25 to 30 iterations where the fitness is three percent. Then there is a sudden increase in the fitness of an agent, who by the rules of propagation generates a progeny. Efficiency can be as high as 11.3%. There is some variation as the agents continue to mutate and mix, but the overall fitness does not seem to increase. 56

4.1 Beat frequency source setup. A VCSEL is locked onto a ^{85}Rb transition using a proportional-integral feedback loop. The laser is then mixed with another VCSEL at the other hyperfine split frequency, and the beat frequency is detected by a photodetector. 63

4.2 Allan deviation data. The red points with error bars are experimental data, and the blue line is calculated by Eq. 4.10. Numbers along top axis denote sample size. Around τ=1 sec, the Allan deviation flattens out due to flicker noise at the mixer. Inset shows the power spectrum of the beat signal at the hyperfine frequency. Black dots show data; the red line shows fit to a pseudo Voigt profile (cf., [71]). 64

4.3 Comparison of normalized Allan deviations, i.e. Allan deviation divided by the phase noise limit Allan deviation in Eq. 4.10. Sources and laser power, linewidth, and locking limit are **a**: 3 mW, 21 kHz, sub-Doppler limited and cavity stabilized[6]; **b**: 5 mW, 30 MHz, Doppler limited[5]; **c**: 3 mW, 300 kHz, Doppler limited [1]; **d**: this work, 100 μW, 100 MHz, Doppler limited; **e**: unknown, 0.5 Hz, cavity stabilized[43]; **f**: 0.2 MHz and 3 MHz, 7 mW and 4.5 mW, sub-Doppler limited[78]; **g**: 3mW, 300 kHz, sub-Doppler limited[1]. It is noted that that sub-Doppler or cavity stabilization can yield lower normalized Allan deviations (**e**, **f**, **g**) over Doppler limited stabilization, as expected. Inset: While it is clear from the normalized Allan deviation that a sub-Doppler limited system (**f**) yielded better results than a Doppler limited system (**c**), their actual Allan deviations are similar due to more inherent phase noise in **f**. 66

5.1 Schematic of NORIS. Section 1 shows the atomically stabilized laser. An external cavity laser is stabilized to the hyperfine structure of ^{85}Rb, to a stability of a few parts per billion by saturation spectroscopy. In section 2, the laser beam is propagated and its beam width is increased slightly by allowing it to traverse a path length. Section 3 shows the optical transfer wafer, which is used to generate a diffraction pattern. The wafer has a suspended thin film metal aperture array, which diffracts and forms the optical ruler. The optical ruler projects above the manufacturing wafer, or any region of space where the nanometrology is to take place. The aperture array thin film is temperature controlled by a thermo-electric cooler. A camera mounted on an actuator is used to image the optical ruler, which allows the system to detect its position. Here, the camera is shown attached to a scanning tunneling tip. Section 4 shows the digital signal processing required to run the system and control the position. Cross correlation methods must be used to detect the position: a fully programmable gate array is shown, which would be a fast method of calculating cross correlations. . . 71

5.2 Optimizing NORIS by considering the mechanical frequency responses. The pneumatically lifted optical table acts as a low pass filter; PVC curtains minimize other, external, acoustic sources of mechanical noise from entering the system. The optomechanics are designed to have high mechanical resonance frequencies. In this Figure, the table specification of a 30 Hz corner frequency for the table is used, and a 250 Hz corner frequency of optical components has been calculated in finite element method simulations of harmonic modes. The resulting pass band shows the best frequency spectrum to actuate NORIS: the system would be unaffected by low frequency noise introduced through the table, but its movements would not excite the high frequency modes in the optomechanics. 74

5.3 Figure shows the intensity ADC calculated at different preamplifications of the input signal. The change in amplification causes different intensities to be measurent. Taking measurements at one amplifiation biases an ADC measurement due to the discretization error. 77

5.4 Figure 5.3 showed the different ADC intensities that were measured; by taking the combined information, the system can converge to the actual analog intensity value. The Figure shows the calculated intensity by combining the information by sweeping the gain from 256 to a setting shown in the abscissa. Even after ten or eleven gain settings are used, the intensity is very nearly close to its actual value, much improved over the value calculated from just one gain setting (at 256). 78

5.5 Calculated intensity error. A Gaussian function is sampled, then its curve is sampled by an ADC. Teal graph shows the calculated curve by taking a number of averages with a 400-level ADC. As the curve approaches the low intensity pixels toward the tails of the Gaussian curve, the quantization noise is too large and the error is large. The gain sweeping method described here captures these parts of the curve: this method has less noise for much more of the curve, reproducing the curve with more precision having taken the same number of images as just averaging them. 79

5.6 Average variance per pixel comparing averaged versus gain swept ADC images of Gaussian curve, as also shown in Fig. 5.5. The gain swept method has about half as much variance as the averaged images. The variance actually decreases as a function of increased pixel noise, a phenomenon akin to stochastic resonance. 80

5.7 Effective ADC sampling rate using gain sweep method, where the gain is swept from the maximum value down to the starting sweep gain value. When the starting sweep gain is half as much as the starting sweep gain, all the gains are factors of higher gains so that no new information is contributed to calculating the image: the effective ADC sampling rate is seen to flatten out. Improvements between 2 and 2.5 orders of magnitudes are seen. 81

5.8 Variance per pixel as a function of starting sweep gain value. At high starting gain values, the variance decreases as the starting gain value is decreased resulting in more information gained about the image. The variance flattens out: as the starting gain is lowered, the noise eventually overwhelms the quantization noise and the variance increases again. 82

5.9 Schematic of NORIS with quasiperiodic pattern. SEM shows holes in silicon-on-insulator silicon patterned as a Penrose tiling, a periodic pattern (bar shows 100 μm). Image shows resulting diffraction pattern optical ruler (bar shows 600 μm). 88

5.10 a. Calculated images of optical ruler, generated by 780 nm Fraunhoffer diffraction from the quasiperiodic, 9662, 3μm diameter holes distributed over 2 mm, sampled at a distance of 25.4mm. Note high density of features across the whole image, as was confirmed in the real image in Fig. 5.9. b. Same conditions as a., but of a periodic square lattice aperture array. 89

5.11 NORIS precision compared between a periodic and quasiperiodic aperture array. The periodic array performs up to 31 times worse than the quasiperiodic array. The sampled portion of the periodic array is particularly high in features so that the periodic array precision will in general be much worse than shown here. 90

5.12 Left half shows Penrose tiling using the thin and thick rhombuses (P3). The thin rhombus has angles 36 and 144 degrees, whereas the thick rhombuses have angles of 72 and 108 degrees. Cutting these rhombuses in half to generate two congruent triangles, these triangles have edges in ratios of the Golden ratio. Right side shows the resulting vertices used to design the aperture array. Circle shows center point: this is the only place in the pattern with five-fold rotational symmetry. 91

5.13 Prototypes of optical diffraction rulers can be made by fabricating a photomask with the desired aperture pattern. The micrograph on the left shows a back-illuminated photomask of a Penrose tiling aperture pattern. Image on right shows resulting diffraction pattern. 92

5.14 Two process fl6ws for making a wafer with an optically accessible diffraction aperture array. Left shows single side polished wafer with LPCVD nitride, thin film of Al etched by RIE for precise patterning, and backside patterning and KOH through-wafer etch to expose the thin film pattern to optical access. Right side shows a silicon-on-insulator wafer. LPCVD nitride is grown, then etched from device side in RIE; subsequent thermal oxide grows only on top device side due to nitride on back side. Due to thin thickness of oxide, precise electron beam lithography can pattern the oxide. Pattern is transferred into device layer Silicon by chlorine chemistry reactive ion etch. Backside nitride is pattern, then through-wafer KOH etched to reach the buried oxide. Then oxide is released by wet buffered oxide etch, then a thin film of Cr/Au is evaporated from the top side to act as the diffractive thin film metal: where holes were etched in the device layer Si, no metal film is evaporated. For large area aperture arrays, a simple process flow of gold thin-film metal evaporated onto a fused silica wafer can be used. 94

5.15 NORIS precision as function of pixel width L_{pix}, in x direction. As in 5.17, the precision is flat until a critical value of L_{pix}=11μm where the precision diverges up to 60% of the pixel size. 97

5.16 NORIS precision as function of pixel width L_{pix}, in y direction. As in 5.17, the precision is flat until a critical value of L_{pix}=11μm where the precision diverges up to 60% of the pixel size. 98

5.17 NORIS precision as a function of N_{pix}, the number of pixels per side, in x direction. The error can reach as high as one third of the pixel size at low N_{pix}. At a critical N_{pix}, the error precipitously drops and decreases less than 50 nm for N_{pix} = 80 to 124. The critical N_{pix} increases as noise is added. 99

5.18 NORIS precision as a function of N_{pix}, the number of pixels per side, in y direction. The error can reach as high as one third of the pixel size at low N_{pix}. At a critical N_{pix}, the error precipitously drops and decreases less than 50 nm for N_{pix} = 80 to 124. The critical N_{pix} increases as noise is added. 100

5.19 NORIS precision as a function of the number of pixels, but with 5% noise added, in x direction. In the noise-free cases in the Figures 5.17 and 5.18, the transition to good precision occurred around 80 pixels, but adding the 5% of pixel noise to about 86 pixels. In addition, the precision before the transition is much less stable and has higher errors as well. 101

5.20 NORIS precision as a function of the number of pixels, but with 5% noise added, in x direction. In the noise-free cases in the Figures 5.17 and 5.18, the transition to good precision occurred around 80 pixels, but adding the 5% of pixel noise to about 86 pixels. In addition, the precision before the transition is much less stable and has higher errors as well. 102

5.21 NORIS precision as a function of N, the percentage of noise per mean pixel intensity. The error is independent of offset, which is a result of the high density of features in the diffraction pattern of the quasiperiodic aperture array throughout the optical ruler. . . . 103

5.22 NORIS precision as a function of N, the percentage of noise per mean pixel intensity. The error is independent of offset, which is a result of the high density of features in the diffraction pattern of the quasiperiodic aperture array throughout the optical ruler. . . . 104

5.23 NORIS precision as a function of the displacement. The Fourier cross-correlation method assumes periodic images. However, when the imager sees a displaced image, the part of the image that disppears to one side does not appear on the other side. This introduces edge effects, and the result is seen here as a sinusoidal error in positioning. 105

5.24 Long-term stability measurement of NORIS in x and y directions of an idle stage. The drift is less than 38 nm, with standard deviations of 27.1 nm and 23.8 nm. 108

5.25 Precision of a long travel stepper motor, corrected for angular misalignment. This data also demonstrates lack of half-wavelength ambiguity in position measurement and the sub-pixel positioning which allows for a continuous resolution over long distances. 109

5.26 Hysteresis of a long travel stepper motor. The position is measured after the stage traverse 10, 30, and 50 mm then returns to its origin. 110

5.27 Comparison of NORIS measured position to high precision capacitive sensor stage position, showing sub-50nm precision (mean offset 29nm = 1/124th of pixel size), a precision of 5×10^{-7} over four inches. The slight oscillatory behaviour is due to the signal processing. 111

5.28 Comparison of NORIS measured position to high precision capacitive sensor stage position, showing sub-50nm precision (mean offset 29nm = 1/124th of pixel size), a precision of 5×10^{-7} over four inches. The slight oscillatory behaviour is due to the signal processing. 112

6.1 In custom scanning tunneling microscope fabricated with NORIS optics, a few layer graphene sheet was etched with a small hole by applying a voltage to the STM tip. After etching, the STM was scanned to view the hole; a subsequent image shows drift in the image. 114

6.2 When the sample bias voltage is positive with respect to the scanning probe tip, the essentially irreversible reaction in Equation 6.1 etches the graphite. The opposite polarity results in no etching of the graphite as the generation of methane requires the carbon to react with four hydrogen ions and four free electrons. The polarity-dependent etching of HOPG allows for electrochemical cleaning of the probe and possibly graphite annealing. 116

6.3 Etch profiles at several sample bias voltages. Note the constant etch rates, especially at +7 V and +8 V sample bias, which imply a constant activity coefficient. The current is also measured as the tip etches the carbon. At a bias of +5 V, the hydrogen ion and free electron currents are not measured by the scanning probe tip. However at a bias of +6 V, the tip and substrate capture some of the current. 117

6.4 (a) Tapping mode image of large hole (scale bars show 50 nm) showing some crevice structure. After imaging, three contact mode scans were made of the hole. Subsequent tapping mode image shows removal of water. (b) Tapping mode image of a smooth line (scale bar shows 100 nm). 118

6.5 Schematic summary of modes of operation for AFM graphite nanolithography: (a) Etching occurs at positive sample bias voltages with a meniscus formed between the tip and the sample via an electrochemical etch chemistry. While etching, carbon deposits form on the AFM tip. (b) Tip cleaning of carbon deposits occurs at negative sample bias voltages. (c) Image scans are at zero bias voltage. 119

6.6 Temperature dependence of HOPG etching with V_H=10.5 V, V_L=9 V. Top shows decrease in cutting rate as the temperature of the HOPG is raised. A decrease in cutting rate can be offset by increasing the sample voltage. Bottom shows improved lithography resolution as temperature of HOPG is raised. Increased temperature leads to increased volatility due to evaporation at the edges of the meniscus, resulting in a smaller meniscus resulting in a smaller etch feature size but also much slower etch rates. 121

6.7 Etch profiles using etch control depicted in inset with $V_H = +11.5$ V and $V_L = +11$ V; a three second tip cleaning at -10 V sample bias precedes all the etches. (a) An optimal etching profile, where the etching depth and deceleration are smooth. (b) The etch begins very slowly for 2.8 seconds. Arrows indicate pauses during etching, possibly due to disappearance and reappearance of the meniscus. (c) Very slow etch with several etching pauses. Even with the same parameters as (a) and (b), the etch time is twice as the nominal etch in (a). Inset shows control feedback for precision etching of graphite. The applied bias starts at a large V_H, rapidly initiating the cut and etching most of the desired depth. The applied bias decreases linearly with etching depth, until at some percentage of the desired etch depth the applied bias is set to a low V_L, typically a volt below V_H. The V_L yields a well-controlled etch depth and width by slowing down the etch rate. 123

6.8 AFM image shows large array of 18.3±3 nm deep etched lines in HOPG. Lines are randomly selected to be 136 nm and 183 nm, yielding 6 nm (4.4 %) and 5 nm (2.7 %) precision, respectively. Figure below shows sample cross sections of the short and long lines. 124

A.1 Schematic of feedback controlled system. The user inputs a desired setpoint. At the summing junction, the system state is subtracted from the desired setpoint. The error signal is fed into a control algorithm, which calculates a response to the system to correct for the error. The signal is applied to the system to try to move its state to the desired setpoint. 127

A.2 Schematic of a loop. The loop transfer function is denoted L(s), and the negative feedback is summed back into the loop. The blue line indicates the point of the loop where it is theoretically broken and the stability can be calculated. As described in the text, the poles can be calculated to deduce the loop stability. 128

A.3 A control loop for a MEMS oscillator. The MEMS oscillator is modeled by a serial RLC tank and a feedthrough capacitance. . . . 129

A.4 Nyquist plots. a. Nyquist plot for MEMS oscillator in Fig. A.3, with G=5, R=2, L=4, C=2, C_1=10. The Nyquist loop circles -1+0i twice, i.e. N=2. Z=N+P=2, so that the loop is unstable. b. Nyquist plot for G=5, R=3, L=1, C=5, C_1=10. The Nyquist loop no longer circles -1+0i, so that N=0 and therefore has no unstable poles. The transition from instability to stability (no oscillation) is a reflection of the decrease in the Q=1/R$\sqrt{L/C}$ of the resonator. . 130

B.1 Fibonacci spiral design, where $r \approx \sqrt{\theta}$. 134

B.2 Arrays of Golden spirals. Holes were placed approximately along the Golden spiral, $r = a \exp b\theta$, where $b = \ln \phi/(\pi/2)$ where ϕ is the Golden ratio. 135

B.3 Radial structure of a circular pattern. Circles of apertures are placed radially away from the center point. 136

B.4 Arrays of spiral of circles. Circle of apertures, whose radius decreases slowly around the circumference, are placed in an array. . . 137

CHAPTER 1
INTRODUCTION

Standards of measurements has a broad range of impact including business, technology, and science. Measurement standards enables of intra- and international trade. In the United States, the National Institute of Standards and Technology – an agency of the Department of Commerce – develops and maintains a set of national standards that the private sector can use for its weights and measurements. It has been estimated that in 1998, weights and measurement regulation impacted some $4.5 trillion of the $8.51 trillion U.S. GDP [33]. The Bureau International des Poids et Mesures creates standards and metrology on the international scale, maintaining the Système International d'Unités (SI). Historically, improvements in metrology and standards have paved the way for tremendous improvements in both the quality and quantity of industrial output. The development of interchangeable parts and mass production in the nineteenth and early twentieth century were catalyzed by precise dissemination of the mil (0.001 inch) using gauge blocks and vernier calipers [26]. The current explosion in computing, sensing, and actuation devices based on microfabrication is enabled by precise metrology at the micron and sub-micron scales due to mechanical actuators and especially laser-based interferometric systems. Fundamental science benefits from improvements in metrology as well. For example, it has been proposed that the fine structure constant $\alpha = e^2/\hbar c 4\pi\epsilon_0 \approx 1/137$ has relatively changed 4.5×10^{-8} over the last two billion years [37], a change of 7.1347×10^{-25} per second. The value of the fine structure constant is of great interest in physics, since it is a nondimensional constant and therefore the same value in all systems of units, and also appears in many physical relationships such as atomic spectra. Its measurement, however, requires

experiments using standardized units: their high precision and definition allows precise measurements of the fine structure constant based on empirical data. It is also expected that sensitive predictions of General Relativity will be measured as well.

In micro- and nanoscale systems, precision in length and time are of primary concern, although the complete SI base units now includes mass (kilograms), electric current (amperes), thermodynamic temperature (Kelvins), amounts of substance (moles), and luminous intensity (candela) in addition to meters and kilograms. The definition of the meter traces back to 1799 as 10^{-7} of a quadrant of the earth. In 1889, platinum-iridium bar meter-length standards were used. In such cases, the standard is traceably transferred around the world while the actual standard remains permanently placed at one location. The current definition of the meter is the distance traveled by light, in vacuum, in 1/299792458 seconds. The definition of time was astronomical for most of its history. It was first defined as 86,400th of the mean solar day, then based on the tropical year 1900. It was only in 1967 that the second was defined as the duration of 9,192,631,770 periods of the radiation of the hyperfine splitting in the ground state of the cesium 133 atom, later confirmed to be at rest at 0 K.

These two examples stress the importance of several aspects of standards. Standards are only useful insofar as they can be disseminated or recalibrated in labs around the world. Therefore, standards are, at best, as good as their definitions. The meter definition based on the quadrant of the earth was off inherently off by 0.2 mm due to a miscalculation of the reference ellipsoid shape of the earth as it rotates. In addition, every level of dissemination will lead to a decrease in precision.

The use of prototypes can also lead to time dependent loss of standards precision: prototypes of the kilogram standard have drifted some 50 μg over a century. It is not known which samples have gained or lost mass, or by how much, and new standards for the kilogram are being developed.

The ultimate standards, then, are based on atomic systems. Atoms are both ubiquitous and identical: for example, the cesium isotope used in the definition of time is the only naturally occurring isotope of cesium. In addition, atoms are governed by quantum mechanics yielding several quantized physical properties which are exploited for standards. Packaging options are diverse with atomic systems as well, ranging from microfabricated gas cells to table-top systems.

While standards are important for disseminating accuracy, the high precision of systems are required. Precision is the amount of variation in a measurement. A ruler that can disseminate a meter as 1.05 m $\pm$ 250 μm has a lower precision than a vernier calipers which disseminate a meter as 1.1 m $\pm$ 5 μm. However, the accuracy of the ruler is higher because the measurement of dissemination of 1.05 m is closer to the desired value of 1 m than the dissemination of 1.1 m by the calipers.

As pointed out by Schattenburg, historically the rise of highly productive massive production grew with and as a result of new capabilities in metrology. The industrial revolution in the nineteenth and early twentieth centuries were a result of precision metrology on the order of the mil=0.001 inches ($\approx$25μm). Household appliances and Ford automobiles were mass produced with interchangeable parts machined with single-mil precision and disseminated by vernier calipers. The semi-

conductor revolution since the second half of the twentieth century is enabled by precision lithography down at the micron scale (and now down to tens of nanometers) and thin films grown to even higher precision and quality. In addition, the introduction of atomic resonance physics in the fifties and sixties spurred technologies in high precision timing: atomic fountain clocks today would lose one second over thirty million years. Schattenburg claims a new nanotechnology revolution starting in the twenty-first century (although apparently lasting only fifty years), where higher precision lithography and bottom-up self assembly will result in nanometer precision to enable a fountain of nanotechnology benefits.

However, while technologies are able to fabricate at smaller and smaller sizes that is not to say that metrology has actually kept up. Consider that in some two hundred years the highest fabrication resolution has gone from 25μm down to tens of nanometers, a reduction of three orders of magnitude. Yet the *precision* in manufacturing is worse. Thin film processing includes variations in actual versus expected etch rates, aspect ratios, etc. A large amount of process engineering is required when a few nanometers results in ten percent change in W/L ratios. In addition, metrology has not kept pace with the progress in resolution. For example, one could not break down a photolithography step into two separate masks and still get the same device characteristics.

The reason, of course, is that while micro- and nanotechnology allows us to manipulate things at smaller and smaller dimensions, it is not possible to do so well over very large areas. Sub-micron manufacturing is based on calibrated or post-testing-adjusted manufacturing (i.e., modifying the photomask based on device performance rather than the reverse). In addition, the 45 nm CMOS technology

node does not place two transistors at opposite ends of the chip within 45 nm, but rather fabricates devices that have features approximately 45 nm in size. As a result, nanofabrication precision is actually on the order of percents or parts of percents, rather than 45 nm over a wafer- or chip-size. Nanofabrication is done with local precision attached to feature-insensitive interconnects, rather than actual wafer-scale precision.

The aim of this work is to develop the same long- and short-scale precision that is available in atomic timing technologies for atomic length metrology for micro- and nanotechnology. In atomic clocks, a local oscillator provides high precision for integration times up to minutes or a few hours. A physics package can lower the precision and accuracy of the clock down to 10^{-12} or 10^{-14} at an hour or more integration because of the unambiguous frequency stability of the atomic physics package. The high precision in the optical timing period corresponds to a high level of precision in its wavelength $\lambda = c/\nu$, where c has been defined to be exactly 299 792 458 meters per second, which can be used to generate a precise, long term stable distance standard.

CHAPTER 2

PRECISION AND ACCURACY IN ATOMIC SYSTEMS

Using atomic systems for metrology alleviates the problems of calibration traceability and standards dissemination. Atomic systems are a plentiful source of perfectly identical calibration-free prototypes that are easily accessible around the world. It is perhaps more appropriate to regard the atom itself as the "atomic clock".

In this chapter the quantum mechanics and interactions with the atom are useful for using the atom as a standard reference are discussed, and also the effects that result in a loss of precision and accuracy.

2.1 Quantization in atoms

This section reviews generally the derivation of the states of the canonical hydrogen atom[66]. The Hamiltonian is,

$$\begin{aligned} \hat{H} &= \frac{\hat{\mathbf{p}}^2}{2\mu} - \frac{Ze^2}{|\hat{\mathbf{r}}|} \\ \frac{\hat{\mathbf{p}}^2}{2\mu} &= \frac{\hat{p}^2_{radial}}{2\mu} + \frac{\hat{\mathbf{L}}^2}{2\mu r^2}, \end{aligned} \tag{2.1}$$

where $\hat{\mathbf{p}}$ is the linear momentum operator, μ is the effective mass, Z is the atomic number, e is the electron charge, and the linear momentum operator is decomposed in terms of a radial momentum operator and an orbital momentum operator. Solutions are found to the time-independent Schrödinger equation, or energy eigenvalue problem, by separating the wave function into its radial and angular parts:

$< r,\theta,\phi >= R(r)Y_{l,m}(\theta,\phi)$. The radial part of the energy eigenvalue problem yields the principal quantum number n=1+l+n_r, where l is the azimuthal quantum number and n_r is the radial quantum number, and subsequently quantization of the bound state energy,

$$E_n = -\frac{\mu c^2 Z^2 \alpha^2}{2n^2}, \tag{2.2}$$

where c is the speed of light and α is the fine structure constant. The quantization in Eq. 2.2 is degenerate with respect to l, the so-called accidental degeneracy. The degeneracy is lifted by the fine structure produced by the spin-orbit coupling of the electron spin and the proton. The perturbative spin-orbit Hamiltonian (including the Thomas factor) is,

$$\hat{H}_{SO} = \frac{Ze^2}{2m_e^2 c^2 r^3}\hat{\mathbf{L}} \cdot \hat{\mathbf{I}}. \tag{2.3}$$

This fine structure, for example, leads to the D1- and D2- lines of alkali atoms. Often, the spin-orbit coupling is stronger than the hyperfine coupling (described next), so that the angular momenta are coupled as the sum of the orbital and spin momenta (also called L-S or Russell-Saunders coupling, as opposed to jj-coupling). In hydrogenic atoms, the quantum number j of the total angular momentum operator $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$ is found to be equal to $l+1/2$ and $l-1/2$, thereby splitting the degeneracy of the orbital angular momentum, so that $E_{SO} = \frac{m_e c^2 Z^4 \alpha^4}{4n^3 l(l+1/2)(l+1)}(\times l, \times -(l+1))$ for $(j = l+1/2, j = l-1/2)$. The first-order and fine structure yielded resonances at X-ray and down to optical frequencies. However it is the hyperfine structure allows RF probing of the atoms thereby enabling atomic precision in electronics. The hyperfine Hamiltonian is due to two terms,

$$\hat{H}_{hfs} = A_{hfs}\hat{I} \cdot \hat{J} + B_{hfs}\frac{3(\hat{I} \cdot \hat{J})^2 + \frac{3}{2}(\hat{I} \cdot \hat{J}) - I(I+1)J(J+1)}{2I(2I-1)J(2J-1)} \tag{2.4}$$

so that

$$\Delta E_{hfs} = \frac{1}{2}A_{hfs} + KB_{hfs}\frac{\frac{3}{2}K(K+1) - 2I(I+1)J(J+1)}{4I(2I-1)J(2J-1)}, \tag{2.5}$$

where the two terms correspond to magnetic dipole and electric quadrupole interactions; the constants A_{hfs} and B_{hfs} can found from empirical data and $K = F(F+1) - I(I+1) - J(J+1)$ [3, 64].

2.1.1 Loss of precision in atomic systems

Three sources of loss of precision in atomic systems are discussed: power broadening, collision or pressure broadening, and Doppler broadening. The latter two are the more dominant sources of line broadening.

Power or saturation broadening is caused by the saturation of the absorption rate of the atoms. Consider an atom with a ground state and an excited state coupled by an electromagnetic field at resonance [58]. Let the ground state and excited states be 1 and 2, respectively. Then, W is the stimulated transition rate between 1 and 2 and $w_{12}/w_{21} = \exp(-\hbar\omega_0/kT)$ be the ratio of the $1 \to 2$ to the $2 \to 1$ relaxation rates, stemming from the Boltzmann distribution at thermal equilibrium. With the equilibrium populations,

$$\frac{dN_1}{dt} = -\frac{dN_2}{dt} = -(W + w_{12})N_1 + (W + w_{21})N_2, \tag{2.6}$$

and the thermal equilibrium population,

$$\Delta N_0 \equiv N_{10} - N_{20} = \frac{w_{21} - w_{12}}{w_{21} + w_{12}}N, \tag{2.7}$$

the population difference is,

$$\begin{aligned}\frac{d}{dt}\Delta N &= \frac{d}{dt}(N_1 - N_2) = -2W\Delta N - (w_{12} + w_{21}) \\ &\times \left(\Delta N - \frac{w_{21} - w12}{w_{12} + w_{21}} N\right) = -2W\Delta N - \frac{\Delta N - \Delta N_0}{T_1},\end{aligned} \tag{2.8}$$

where it is defined that $1/T_1 \equiv w_{12} + w_{21}$ to be the relaxation time. Then at steady state, $\frac{d}{dt}\Delta N = 0$ and

$$\Delta N = \Delta N_0 \frac{1}{1 + 2WT_1}, \tag{2.9}$$

whose form implies,

$$\alpha \propto \frac{1}{1 + (I/I_{sat})}, \tag{2.10}$$

where α is the atomic absorption coefficient. Recalling that the atomic lineshape is Lorentzian, the frequency dependence can be added by adding the expected Lorentzian form into the last equation. The result is that the Lorentzian linewidth is modified by the saturation term, resulting in a linewidth of the form,

$$\Delta\omega_{pb} = \sqrt{1 + I/I_{sat}}\,\Delta\omega_{Lz}, \tag{2.11}$$

where the power broadened linewidth ω_{pb} is increased over the original Lorentzian linewidth ω_{Lz}. Note that the atomic linewidth is not itself changed. Rather, the saturation of the absorption near the middle of the atomic line at higher powers creates the appearance of linewidth broadening because of the decreased absorption in at line center and normal absorption towards the tails.

Pressure or collision broadening is due to the interactions between atoms in gaseous form, which perturbs their otherwise isolated interaction with the probing electromagnetic field. A semiclassical theory [15] yields pressure broadening estimates that may be accurate to 10% [38]. Resonance self-broadening is the interaction between atoms of the same species only where either the upper or lower

energy levels has an electric dipole transition to the ground state.. Its Lorentzian full width at half maximum is $\Delta\lambda_{self} \approx 8.6 \times 10^{-30}(g_i/g_k)^{1/2}\lambda^2\lambda_r f_r N_i$; λ is the observed wavelength, λ_r and f_r are the wavelength and oscillator strength of the resonance line (to the ground state); g_k and g_i are the statistical weights of the upper and lower states (resonance lines), respectively; and N_i is the number density of the ground state. Wavelengths are in Angstroms and densities are in per cubic centimeters.

In addition, Van der Waals interactions occur between the dipole moment of an excited atom with the induced dipole moment of atoms in the ground state. For hydrogen-like atoms such as the alkali atoms typically used here, the linewidth is approximately $\Delta\lambda \approx 3.0 \times 10^{16}\lambda^2 C_6^{2/5}(T/m)^{3/10}N$, where λ is the wavelength, T is the temperature, m is the mass, N is the density; and C_6 is the interaction coefficient $\approx C_k - C_i$, where $C_{i(k)} = 1.64 \times 10^{-13}(3I_H/4E*)[I_H/(I - E_{i(k)})]^2$, where I_H is the ionization energy of hydrogen, I is the ionization energy of the radiating atom, $E*$ is the energy of the first excited level of the perturbing atom, and $E_{i(k)}$ are the energies of the upper and lower levels, respectively. Wavelengths are in Angstroms, energies in cm^{-1}, temperatures in Kelvin, and densities are per cubic centimeters.

Doppler broadening is an inhomogeneous broadening of the absorption linewidth of an ensemble of atoms. Because of their kinetic energy, the population is distributed in velocities as described by the Maxwell-Boltzmann distribution. In the rest frame of any one atom, the wavelength of the probe beam will be Doppler shifted away from the set wavelength of the laser beam in the laboratory rest frame. At room temperature, the ensemble of atoms will be Doppler broadened by

hundreds of MHz. Doppler broadening is by far the largest source of broadening, but can be removed or decreased by introducing a buffer gas (Dicke narrowing) or by saturation spectroscopy.

2.1.2 Saturation spectroscopy

Saturation spectroscopy or hole burning is considered as a means of linewidth narrowing for improved frequency stability by eliminating Doppler broadening. The analysis is semiclassical, similar to the Bloch equations. By eliminating Doppler broadening, it is possible to see the lifetime-broadened linewidths of alkali atoms.

In saturation spectroscopy or hole burning, a pump beam is used to excite or burn 'holes' of excited atoms in its path. A weaker, counterpropagating beam (at the same wavelength as the pump beam) probes the same population. An atom with zero net velocity along the beam axes will absorb both the pump and probe beams at resonance, so that the probe beam will experience almost no absorption because the pump beam has excited the atoms. An atom with some velocity along the axes, however, cannot absorb both the pump and probe beams because it is blue-shifted to one beam but red-shifted to the other: it is no longer in resonance with both wavelengths. Therefore, if beams are off resonance, atoms that have the correct Doppler shift will not be excited by the pump beam, but will absorb the probe beam. This saturation or hole burning leads to a small dip in the absorption spectrum. This small dip enables a tighter lock on the individual hyperfine states if a featuer narrower than Doppler broadening is required . This effect can also be used to linewidth narrow a broad laser. If a laser linewidth is

less than the Doppler broadened linewidth of the absorption cell, only the holes of the absorption spectrum will be passed through the cell.

In our analysis, the $5S_{1/2}$ is the ground state and $5P_{3/2}$ $F = 2, 3, 4$ are the excited states of ^{85}Rb. The Hamiltonian is the sum of the atomic energy states and the interaction between the atom and the electromagnetic radiation,

$$H = H_{atom} + H_{em}. \tag{2.12}$$

H_{em} is diagonal with respect to the individual atom states, and the elements are of the form $\hbar\omega_i$. H_{em} couples the ground and excited states through the atomic dipole moments. The terms of H_{em} are of the form of a pump beam and a counterpropagating probe beam,

$$H_I = P\epsilon_+ \cos(\omega t - kz) + P\epsilon_- \cos(\omega t + kz), \tag{2.13}$$

where ϵ_+ is for the pump beam and ϵ_- is for the probe beam. The total Hamiltonian, then, is,

$$\mathbf{H} = H_I \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & hf_{g2e2}/H_I & 0 & 0 \\ 1 & 0 & hf_{g2e3}/H_I & 0 \\ 1 & 0 & 0 & hf_{g2e4}/H_I \end{pmatrix}, \tag{2.14}$$

where the atomic states are defined by,

$$\begin{pmatrix} g_2(5S_{1/2}) \\ e_2(5P_{3/2}, F = 2) \\ e_3(5P_{3/2}, F = 3) \\ e_4(5P_{3/2}, F = 4) \end{pmatrix}. \tag{2.15}$$

The density matrix formalism is used. The density matrix ρ, where the diagonal elements are the state populations and the off-diagonal terms are the coherence between states, evolves as,

$$i\hbar\frac{\partial\rho}{\partial t} = [H, \rho]. \tag{2.16}$$

The density matrix elements are defined so that the off-diagonal elements are in the rotating frame ($\rho_{ij} = \tilde{\rho_{ij}} e^{-i\omega t}, i \neq j$), removing all time dependence, and the difference of the diagonal elements are used ($\sigma_{ig2} = \rho_{ii} - \rho_{g2g2}$). Then,

$$\begin{aligned}
\dot{\sigma}_{e2g2} &= -2\alpha i f^{+}(k,z)\ \mathrm{Im}\left\{2\rho_{\tilde{g2e2}} + \rho_{g2e3} + \rho_{g2e4}\right\} - \gamma\left(\sigma_{e2g2} + 1\right) && (2.17)\\
\dot{\sigma}_{e3g2} &= -2\alpha i f^{+}(k,z)\ \mathrm{Im}\left\{\rho_{\tilde{g2e2}} + 2\rho_{g2e3} + \rho_{g2e4}\right\} - \gamma\left(\sigma_{e3g2} + 1\right) && (2.18)\\
\dot{\sigma}_{e4g2} &= -2\alpha i f^{+}(k,z)\ \mathrm{Im}\left\{\rho_{\tilde{g2e2}} + \rho_{g2e3} + 2\rho_{g2e4}\right\} - \gamma\left(\sigma_{e4g2} + 1\right) && (2.19)\\
\dot{\tilde{\rho}}_{g2e2} &= \tilde{\rho}_{g2e2}\left(i\Delta\omega_{oe2} - \gamma/2\right) - i\alpha f^{-}(k,z)\sigma_{e2g2} && (2.20)\\
\dot{\tilde{\rho}}_{g2e3} &= \tilde{\rho}_{g2e3}\left(i\Delta\omega_{oe3} - \gamma/2\right) - i\alpha f^{-}(k,z)\sigma_{e3g2} && (2.21)\\
\dot{\tilde{\rho}}_{g2e4} &= \tilde{\rho}_{g2e4}\left(i\Delta\omega_{oe4} - \gamma/2\right) - i\alpha f^{-}(k,z)\sigma_{e4g2}, && (2.22)\\
& && (2.23)
\end{aligned}$$

where

$$f^{+}(k,z) = \epsilon_{+}e^{ikz} + \epsilon_{-}e^{-ikz}, \tag{2.24}$$

$$f^{-}(k,z) = \epsilon_{+}e^{-ikz} + \epsilon_{-}e^{ikz}, \tag{2.25}$$

the $\Delta\omega$ are between the ground state and the indicated excited state, γ is a phenomenological excited state decay rate, $\alpha = P/2\hbar$ where P is the atomic electric dipole moment, and the $\dot{} = v\ \partial/\partial x$, since time dependence has been removed. For modes of e^{nikz} for integers n, the functions f^{+} and f^{-} couple the modes of ρ and σ. It is seen that the nth modes of ρ couple to the $n \pm 1$ modes of σ, and the nth modes of σ couple to the $n \pm 1$ modes of ρ. Then all the modes of ρ are even

and all the modes of σ are odd, or vice versa. Equations 2.17, 2.18, and 2.19 have a constant term on the right hand side ($-\gamma$), so it must be that the σ's are even modes and the ρ's are odd modes of e^{nikz}. The equations above are solved for the 0th order of σ and the ± 1 orders of ρ, resulting in fifteen equations. The linear system to be solved is,

$$\mathbb{M}\,\mathbb{X} = \mathbb{B}, \tag{2.26}$$

where,

$$\mathbb{M} = \left(\begin{array}{ccccc}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & \frac{\gamma}{2} & 0 \\
0 & 0 & 0 & 0 & \frac{\gamma}{2} \\
0 & 0 & 0 & 0 & 0 \\
\alpha\epsilon_{minus} & 0 & 0 & kv - \Delta\omega_{oe2} & 0 \\
0 & \alpha\epsilon_{minus} & 0 & 0 & kv - \Delta\omega_{oe3} \\
0 & 0 & \alpha\epsilon_{minus} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\alpha\epsilon_{plus} & 0 & 0 & 0 & 0 \\
0 & \alpha\epsilon_{plus} & 0 & 0 & 0 \\
0 & 0 & \alpha\epsilon_{plus} & 0 & 0
\end{array} \right.$$

$$
\begin{array}{cccccc}
 & 0 & \frac{2\alpha\epsilon_{minus}}{\gamma} & \frac{\alpha\epsilon_{minus}}{\gamma} & \frac{\alpha\epsilon_{minus}}{\gamma} & 0 \\
 & 0 & \frac{\alpha\epsilon_{minus}}{\gamma} & \frac{2\alpha\epsilon_{minus}}{\gamma} & \frac{\alpha\epsilon_{minus}}{\gamma} & 0 \\
 & 0 & \frac{\alpha\epsilon_{minus}}{\gamma} & \frac{\alpha\epsilon_{minus}}{\gamma} & \frac{2\alpha\epsilon_{minus}}{\gamma} & 0 \\
 & 0 & -kv+\Delta\omega_{oe2} & 0 & 0 & 0 \\
 & 0 & 0 & -kv+\Delta\omega_{oe3} & 0 & 0 \\
 & \frac{\gamma}{2} & 0 & 0 & -kv+\Delta\omega_{oe4} & 0 \\
 & 0 & \frac{\gamma}{2} & 0 & 0 & 0 \\
\cdots & 0 & 0 & \frac{\gamma}{2} & 0 & 0 \\
 & kv-\Delta\omega_{oe4} & 0 & 0 & \frac{\gamma}{2} & 0 \\
 & 0 & 0 & 0 & 0 & \frac{\gamma}{2} \\
 & 0 & 0 & 0 & 0 & 0 \\
 & 0 & 0 & 0 & 0 & 0 \\
 & 0 & 0 & 0 & 0 & -kv-\Delta\omega_{oe2} \\
 & 0 & 0 & 0 & 0 & 0 \\
 & 0 & 0 & 0 & 0 & 0
\end{array}
$$

$$
\cdots \begin{pmatrix}
0 & 0 & \frac{2\alpha\epsilon_{plus}}{\gamma} & \frac{\alpha\epsilon_{plus}}{\gamma} & \frac{\alpha\epsilon_{plus}}{\gamma} \\
0 & 0 & \frac{\alpha\epsilon_{plus}}{\gamma} & \frac{2\alpha\epsilon_{plus}}{\gamma} & \frac{\alpha\epsilon_{plus}}{\gamma} \\
0 & 0 & \frac{\alpha\epsilon_{plus}}{\gamma} & \frac{\alpha\epsilon_{plus}}{\gamma} & \frac{2\alpha\epsilon_{plus}}{\gamma} \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & kv + \Delta\omega_{oe2} & 0 & 0 \\
\frac{\gamma}{2} & 0 & 0 & kv + \Delta\omega_{oe3} & 0 \\
0 & \frac{\gamma}{2} & 0 & 0 & kv + \Delta\omega_{oe4} \\
0 & 0 & \frac{\gamma}{2} & 0 & 0 \\
-kv - \Delta\omega_{oe3} & 0 & 0 & \frac{\gamma}{2} & 0 \\
0 & -kv - \Delta\omega_{oe4} & 0 & 0 & \frac{\gamma}{2}
\end{pmatrix} \quad (2.27)
$$

$$
\mathbb{X} = \begin{pmatrix} \sigma_{e2}^{(0)} \\ \sigma_{e3}^{(0)} \\ \sigma_{e4}^{(0)} \\ \mathrm{Re}\{\rho_{e2}^{(1)}\} \\ \mathrm{Re}\{\rho_{e3}^{(1)}\} \\ \mathrm{Re}\{\rho_{e4}^{(1)}\} \\ \mathrm{Im}\{\rho_{e2}^{(1)}\} \\ \mathrm{Im}\{\rho_{e3}^{(1)}\} \\ \mathrm{Im}\{\rho_{e4}^{(1)}\} \\ \mathrm{Re}\{\rho_{e2}^{(-1)}\} \\ \mathrm{Re}\{\rho_{e3}^{(-1)}\} \\ \mathrm{Re}\{\rho_{e4}^{(-1)}\} \\ \mathrm{Im}\{\rho_{e2}^{(-1)}\} \\ \mathrm{Im}\{\rho_{e3}^{(-1)}\} \\ \mathrm{Im}\{\rho_{e4}^{(-1)}\} \end{pmatrix}, \ \mathbb{B} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{2.28}
$$

The solution desired is the sum of the real parts of the $\rho^{(-1)}$'s, where the e^{-ikz} matches the spatial distribution of the probe beam. The current solution returned by Mathematica is long and will need to be simplified, for example, by eliminating negligible terms. In addition, a Maxwell's distribution will need to be added to account for the distribution of particle velocities.

The absorption is then given by,

$$
\left\langle E \frac{d}{dt} P \right\rangle, \tag{2.29}
$$

where

$$
P = Trace\left(\rho \hat{P}\right), \tag{2.30}
$$

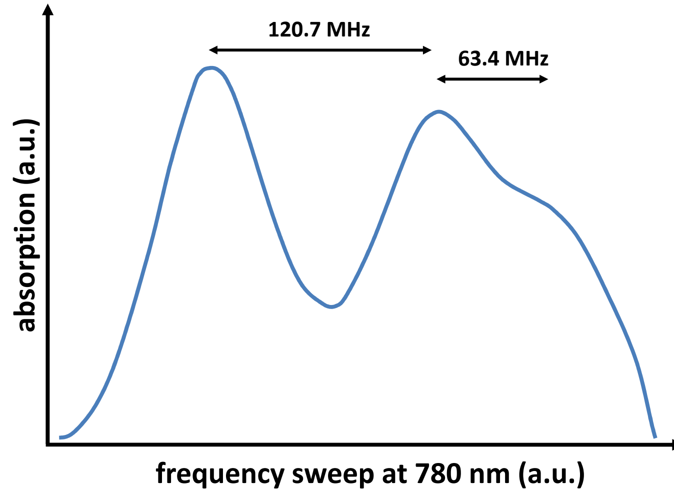


Figure 2.1: Expected calculated transmission spectrum for our VCSEL setup, as calculated in the text. The saturated spectrum shows the three spectral holes, corresponding to the three resonances that were assumed. Two of the spectral holes overlap.

where ρ is the density matrix and $\hat{P}$ is the polarization operator: the result is a sum over the non-diagonal elements of the density matrix.

The resulting calculated absorption spectrum is shown in Fig. 2.1, where the width of the Figure is 6 GHz. As derived in the previous section, there are three main resonances from the ground state to the three excited states, two of which overlap. Note that this model shows several sources of precision noise. First, the propagating terms in Equation 2.13 will yield Doppler broadening; the finite lifetimes of the excited states cause a Lorentzian linewidth in the frequency domain; and modeling of excitation and decay rates will yield some power broadening as well. The sources of saturated absorption broadening serves as a frequency stability baseline when stabilizing lasers to a saturated absorption line. In this case, the broadening results in the overlap between two the saturation absorption lines. As

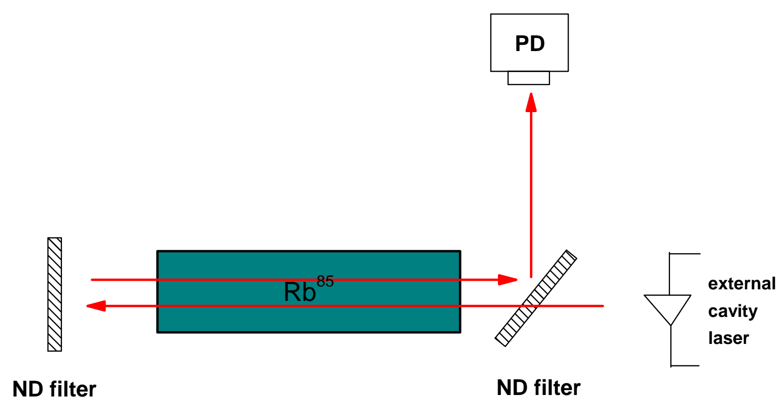


Figure 2.2: Experimental setup for saturation spectroscopy using external cavity laser. ND filters act as beam splitters, but are selected to allow different probe and pump powers.

will be seen, in this method the natural linewidths are the minimum linewidths, which can be achieved by tuning the setup.

2.1.3 Hole burning experiments

Hole burning experiments have been done with a Littrow-configuration external cavity laser with a 7 MHz linewidth (Toptica Photonics DL100). A setup similar to that of McIntyre, et al. [11] is used (see Fig. 2.2). Fig. 2.3 shows hole burning with a 42 mW pump beam and a 220 μW probe.

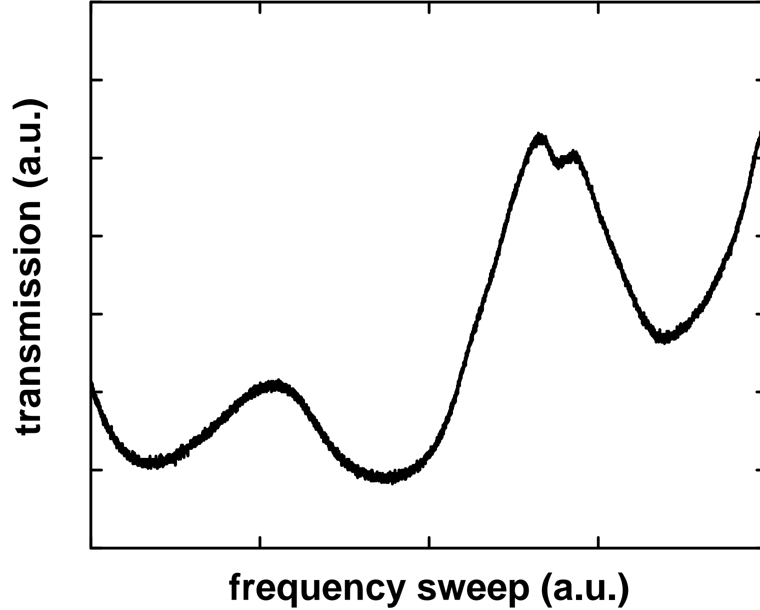


Figure 2.3: Saturation spectroscopy of ^{85}Rb with a 42 mW pump beam and a 200 μW probe beam, using setup in Fig. 2.2. As in Fig. 2.1, two of the holes overlap each other due to their broad linewidths.

Note that, as in Fig. 2.1, two of the saturation absorption lines overlap. The lines are broad due to the power broadening. The power broadening can be reduced by reducing the power of the pump and probe beams. By doing so, the fine structure resolving the different F states are visible, thereby resulting in lifetime-limited absorption lines that will lead to more precise laser frequency stabilization.

In order to study hole burning of ^{85}Rb at low powers, a linear polarizer is included between the laser and the first ND filter. By rotating the polarizer, it is possible to vary continuously the laser power that is used for the pump / probe beams; the external cavity laser is linear polarized to begin with. First the absorption from the F=2 ground state is studied, which has less absorption than that from the F=3 ground state. The wavelength is at 7800 Å, which is the transition between the $5^2S_{1/2}$ and the $5^2P_{3/2}$ states. Fig. 2.4 shows hole burning

with a 9 mW probe beam with a 210 μW probe beam at the photo detector. The numbers in the topside portion show the hyperfine structure in frequency. The theoretical numbers are extracted from [7], and the fit numbers are found as a linear fit to the linearly increasing frequency sweep of the cavity laser. The numbers show general agreement. The ground state F=3 state couples to the excited F=2, 3, and 4 states due to dipole selection rules. In this case, it is possible to see all the crossover resonances. They are labeled by "CO" in Figure 2.4.

The saturation absorption spectrum from the F=2 ground state is shown in Fig. 2.5. The selection rule states that the F=3 ground state will couple only with the F=1, 2, and 3 excited states. All three resonances are visible in the Figure. The crossover resonances are not apparent, although there is enough spectral resolution to do so. The transitions from the F=2 ground state has consistently been void of crossover resonances.

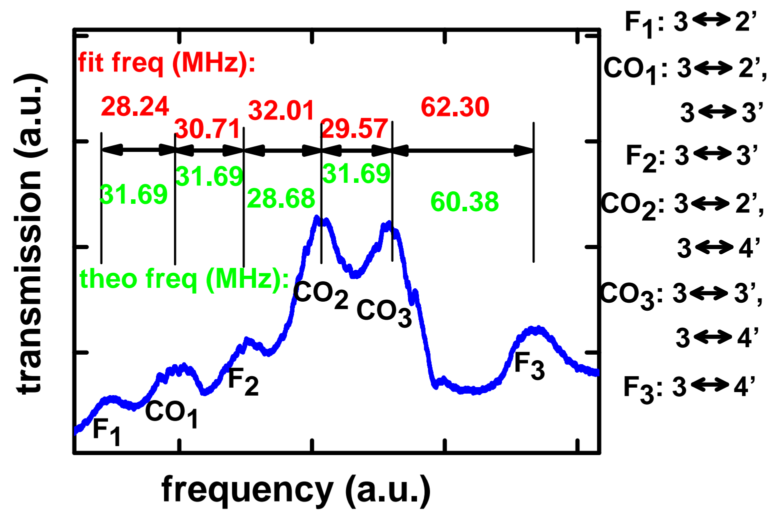


Figure 2.4: Similar data as in Fig. 2.3, but at 9 mW probe beam with a 210 μW probe beam; the structure is for the excitation from the F=3 ground state. The structure is better resolved, and the cross-over resonances are visible. The theoretical frequency splittings are from [7]. The experimental frequencies are found by a linear fit of the spectral hole frequencies, finding general agreement.

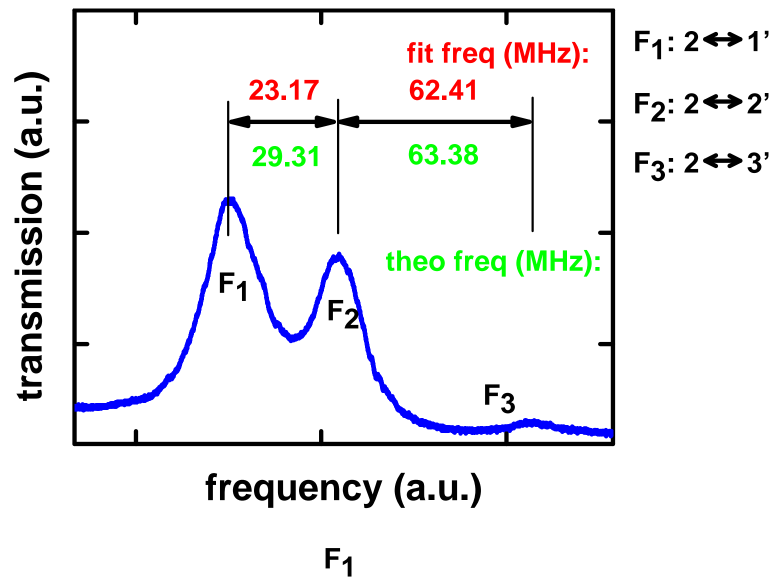


Figure 2.5: Same setup as in Fig. 2.4, but for excitation from the F=2 ground state. The F=2 and F=3 ground state hyperfine splitting in ^{85}Rb is 3.036 GHz. The cross over resonances are not visible here.

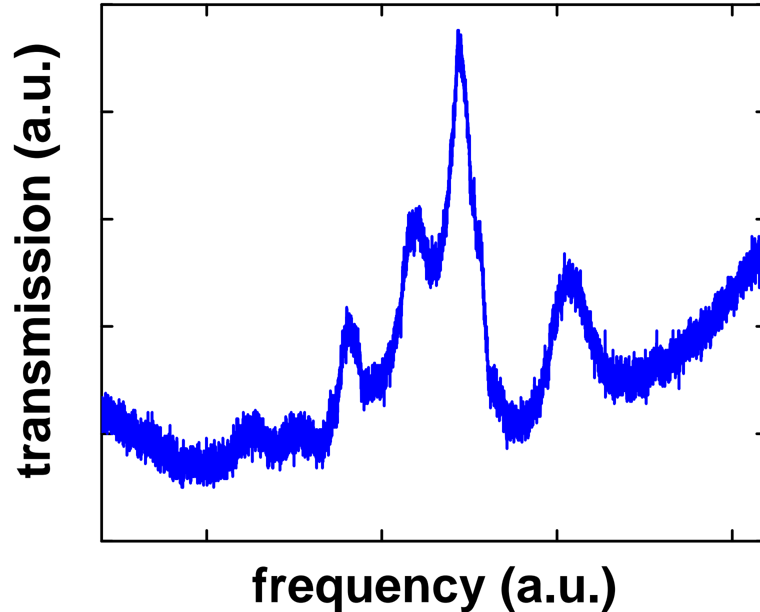


Figure 2.6: Demonstration of low power saturation spectroscopy, especially relevant for the use of low power semiconductor lasers. The pump beam power is decreased to 190 μW, and the probe beam is reduced to 240 nW. Curve shows absorption from the F=3 ground state. Due to the lower beam powers, the atomic vapor temperature is also reduced. As in Fig. 2.4, the cross over resonances are visible.

The data in Figures 2.4 and 2.5 use a very strong pump beam, almost 50 times larger than the probe beam. The ratio of the pump beam intensity to the probe beam intensity can be decreased at a cost, most prominently as the visibility of the fine structure will decrease. At lower pump beam intensities, the percentage of the atom population in the excited state will decrease, and those that are not excited cannot undergo saturation spectroscopy and contribute to the spectral hole burning.

In Figure 2.6, the pump beam is decreased significantly from 9 mW to only 190 μW, resulting in a probe beam that is 240 nW. This is the absorption from

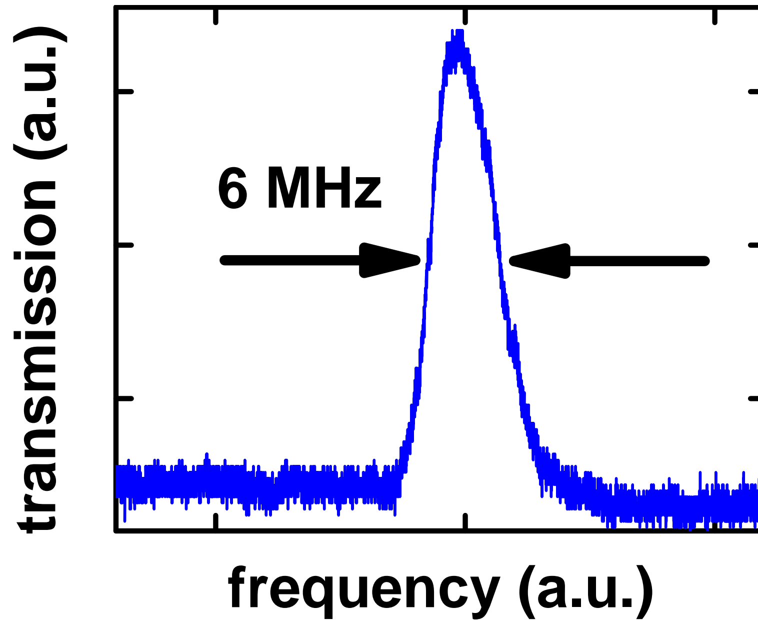


Figure 2.7: The spectral hole can be used to effectively narrow a laser linewidth as well. In this Figure, the pump probe is 1.9 mW but the probe beam is only 480 nW. The probe beam is weak. It is almost completely absorbed by the atomic vapor, except at the narrow 6 MHz spectral hole. This spectrum is captured by a Fabry-Perot cavity (Thor Labs SA200).

the F=3 ground state. In order to accommodate the lower pump and probe beam intensities, it is also necessary to lower the population of probed atoms. This is done by reducing the atomic vapor heater temperature from 60 $°C$ to 40 $°C$. As in Figure 2.4, the fine structure and cross over resonances are seen.

Saturation spectroscopy can be used to frequency stabilize a laser, that is to increase the accuracy of the laser center frequency to a constant of nature (the atomic resonances). In Figure 2.7 it is shown that this tool as well can be used to actually narrow the laser frequency as well. The same set up as in 2.2 is used, with a cell temperature of 60 $°C$. However, the pump to probe intensities are greatly skewed so that the pump probe is 1.9 mW, but the probe beam is only

480 nW (a ratio of almost 4000). The probe beam is so weak that it is completely absorbed by the low population of atoms that were not excited by the pump probe. However, the spectral holes still exist and the wavelengths at the spectral holes pass through easily with high signal to noise (compared to the low 480 nW probe power). Outside the spectral hole, the probe beam is much too weak to make its way through the atomic vapor. The result is that the portion of the laser beam that is passed through the spectral hole acts as a narrowed, atomically accurate laser of 6 MHz width. Note that a feedback loop is not as important here. Whatever portion of the laser overlaps the spectral hole, that will create the desired transmitted laser line. This appears to be similar in concept to atomic line filters. One could imagine, therefore, using a poor high power but broadband diode laser and using a similar setup to generate a narrow, stabilized wavelength.

2.1.4 Hole burning with low power VCSELs

Finally, saturation spectroscopy is demonstrated with very low power vertical cavity surface emitting lasers (VCSELs). Holes were burned using 200 μW VCSELs. The setup is shown in Fig. 2.8. The quarter wave plate and linear polarizers are required to protect the VCSEL from the reflected power. Because of their low photon cavity density, the VCSEL is much more sensitive to the reflected power which can cause chaotic behaviour in its lasing behaviour. Such a condition causes both the frequency and the intensity of the laser to randomly change rapidly. In addition, the system is extremely sensitive to mechanical noise, for example, if the laser beam return mirror is vibrated.

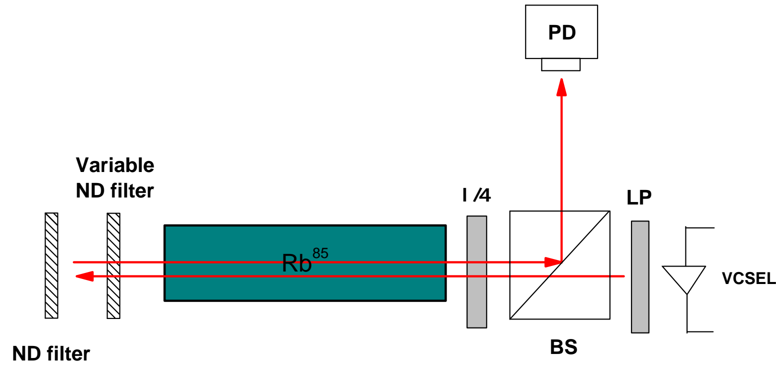


Figure 2.8: Experimental setup for hole burning with VCSELs. A linear polarizer and quarter wave plate act as a directional filter, so that laser power does not reflect act to the VCSEL. Due to the low cavity power of the VCSEL, their oscillation is especially sensitive to extra noise.

Fig. 2.9 shows hole burning in the D2-line from the F=2 ground state, and Fig. 2.10 shows hole burning from the F=3 ground state. Rather than picking up the three lines as in Fig. 2.3, a single peak in the hole transmission spectrum is found. This is due to the large (>100 MHz) linewidth of the VCSEL, which causes the absorption spectrum to be convoluted with all three holes of the hyperfine lines, resulting in a single, large (>100 MHz) linewidth hole. It should be noted that the actual spectrum will still contain three distinct burnt holes.

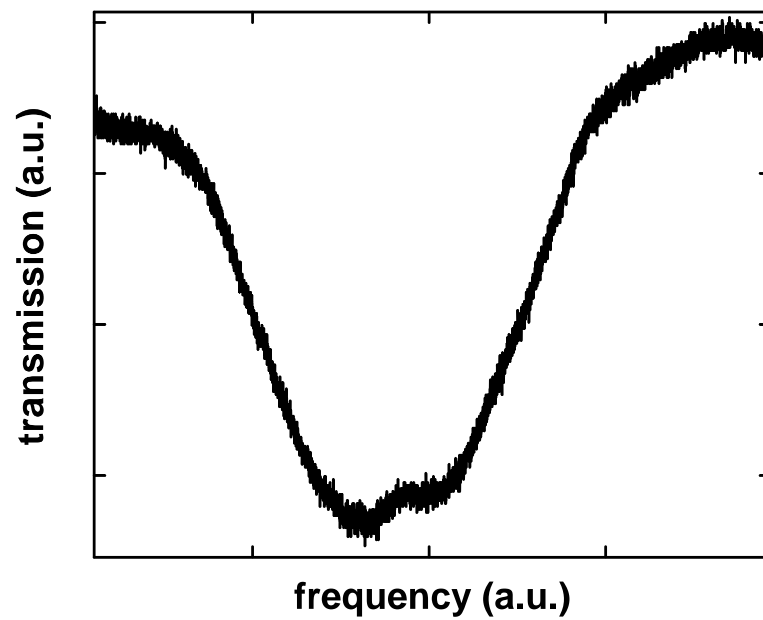


Figure 2.9: Experimental setup for hole burning with VCSELs, excitation from F=2 ground state. These data only show a single spectral hole. This is due to the $>$ 100 MHz linewidth of the VCSEL, so that the observed transmission is a convolution of the two and appears to be one single spectral hole.

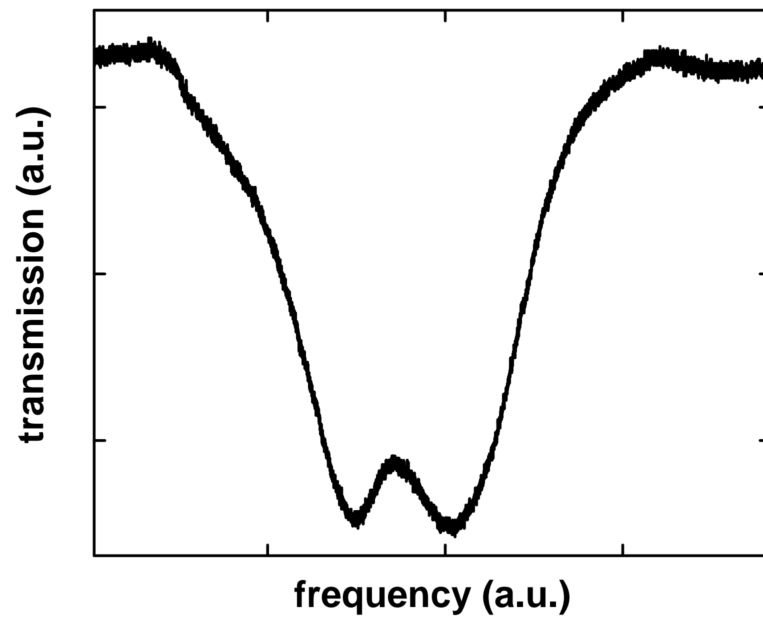


Figure 2.10: Similar data as in Fig. 2.9, but from the F=3 ground state. Just as in Fig. 2.9, the fine structure appears to be a single spectral hole due to the large linewidth of the VCSEL.

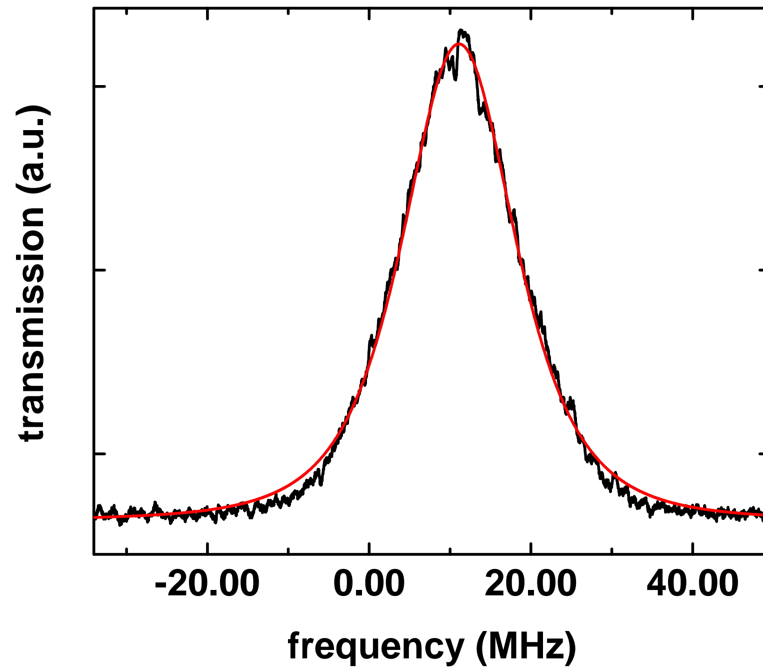


Figure 2.11: In Fig. 2.8, the linear polarizer and quarter wave plate were used to prevent feedback to the VCSEL to prevent chaotic oscillations in the lasing cavity. By using a piezo-mounted mirror, the feedback can be used to narrow the VCSEL linewidth. Here, the linewidth has been narrowed from > 100 MHz to 15 MHz, measured using a Fabry Perot cavity.

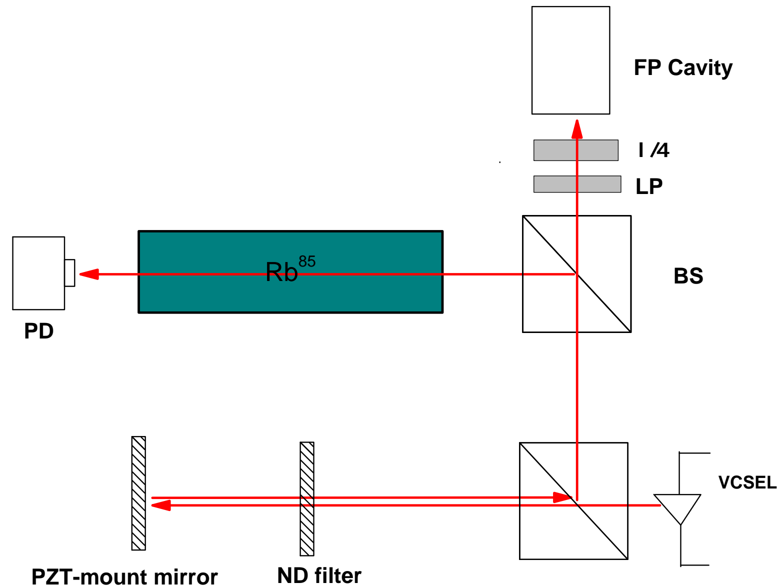


Figure 2.12: Setup for narrowing VCSEL linewidth. The piezo-mounted mirror is necessary to carefully adjust the phase of the feedback to the VCSEL. The ND filter adjusts the feedback power.

Finally, although uncontrolled feedback into the VCSEL can cause chaotic behaviour in its lasing and therefore frequency characteristics, controlled feedback can result in reduced line width (i.e., modified Schawlow-Townes linewidth). The same strategy is used in external cavity lasers. The linewidth of the VCSEL has been measured to be $108MHz$; this has been observed at least as low as 15 MHz using feedback. The linewidth is measured by using a scanning Fabry Perot cavity. See Figs. 2.11 and 2.12.

CHAPTER 3
SELF WAVELENGTH CALIBRATED OPTICAL SOURCES FOR ATOMIC CLOCKS

In the next two sections, optical sources that are self calibrated are discussed. In the first case, a plasma of rubidium atoms are excited by inductive coupling. The rubidium atoms emit photons that are charcteristic of their quantum mechanically defined transitions, resulting in spectra that are wavelength calibrated. In the second case, xenon atoms are excited by the energetic electrons of the β-decay of the radioisotope ^{63}Ni. Again, the atoms emit photons that are quantum mechanically defined and the resulting spectra are wavelength calibrated to these transitions. In the latter case, the photoemission is self powered by the radioactive thin film ^{63}Ni.

3.1 Compact, low power radio frequency rubidium plasma for passive rubidium atomic clocks

The work in this chapter describes the scaling down of large, high power (>10 W) Rb plasmas for Rb-based passive atomic vapor clocks. The clock starts with a Rb^{87} lamp filtered by a Rb^{85} cell. The lamp consists of Rb^{87} gas which radiates its D2- and D1- lines (5 $2P_{3/2}$ - 5 $2S_{1/2}$ @ 780 nm and 5 $2P_{1/2}$ - 5 $2S_{1/2}$ @ 794.7 nm, respectively). The hyperfine structure in the upper states are masked by the Doppler broadening, but the splitting in the ground states (F=2 and 1) results in both the D2- and D1-lines emitting a pair of hyperfine split wavelengths: in Rb^{87}, the ground state is split by 6.835 GHz. One of these lines, specifically the transition

to the upper F=2 hyperfine state, can be filtered by a Rb^{85} cell. A lucky coincidence occurs whereby the upper hyperfine state, at F=3, of Rb^{85} nearly lines up with the F=2 hyperfine state of Rb^{87}. Since the upper states are close and Doppler broadened, the F=3 hyperfine state of Rb^{85} state can absorb light emitted by a transition to the F=2 state of Rb^{87}. On the D2-line the two hyperfine states are separated by 1020 GHz, and on the D1-line they are separated by 780 MHz. The Rb^{85} filter cell can incorporate argon buffer gas to increase the Doppler broadening and shift the two hyperfine states together, increasing the wavelength overlap and therefore the absorption. At this point, the light from the Rb^{87} lamp consists only of radiation from the upper state to the F=1 hyperfine state of the ground state. When this radiation illuminates a cell of Rb^{87}, the F=1 state is depopulated and the atoms are population inverted so that the upper hyperfine F=2 state is populated. A buffer gas, such as argon or xenon, is also typically used. Then by applying an RF resonant field at the ground state hyperfine splitting frequency, the population inversion is modulated and the transmission of the Rb^{87} lamp radiation can be monitored to detect atomic resonance.

In a typical macroscale vapor clock, the Rb^{87} lamp will consume much power. A miniature, low power atomic clock will require a reduction in size and power of the lamp. These experiments used natural Rb gas cells prototyped by FEI, Inc to be much smaller than their counterparts in their production lines. These cells were typically only about one to two centimeters in length and about one millimeter or one centimeter in diameter, containing both rubidium and a buffer gas of xenon. Coils of different shapes were made by hand using 0.5 mm diameter copper magnetic wire. The RF was applied to the coil, a 1 Ω resistor, and a matching capacitor in series. The frequency at which a minimum applied amplitude

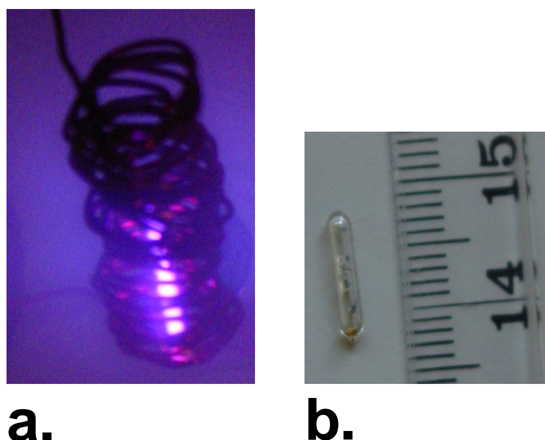


Figure 3.1: Images of miniature rubidium plasma and the bulb. a. rubidium plasma excited by an RF coil, the purplish color showing the optical emissions from excited rubidium atoms. b. The glass bulbs were filled with rubidium and some xenon as a buffer gas, custom made by the glass blowers at FEI, Inc.

was required to ignite the plasmas was found, and also the frequency of maximum observed optical emission. The lowest power for sustaining a plasma was 48 mW using coil type #5.

3.1.1 Spectral response of compact, low power radio frequency rubidium plasmas

Once the plasma is initiated, the power and/or frequency can be tuned to change the output spectrum of the miniature rubidium plasma lamp. Due to the presence of xenon buffer gas, the xenon may emit energy as well although in practice it is desired that only the rubidium emits light. Higher coil frequencies cause both

Table 3.1: Coil types for RF plasma drive of glass Rb cells. N refers to number of winds, d to the diameter or side length, L to the length along the glass tube, and OIF to the optimum ignition frequency – the frequency of ignition with lowest applied power.

	transverse shape	N	d (mm)	L (mm)	OIF (MHz)
1	circular	23	10	20	410
2	square	30	6	18	440
3	square	19	10	16	380
4	quad circle	4	2	2	450
		7	12	7	
		8	2	3	
		12	12	14	
5	split circle	7	2	4.5	250
6	split circle	5	2	2.6	290
7	split circle	2	2	1	290

xenon and rubidium emission lines to appear. By lowering the frequency further, the xenon emission lines get weaker until they suddenly disappear. The effect is seen in Figure 3.2: as the applied frequency decreases, the xenon lines get weaker until they disappear, and only the D1- and D2- lines of rubidium are strong. The Figure labels the strongest emission lines, and the corresponding transition and wavelengths are numbered and tabulated in Table 3.2 as well[56, 55]. At higher frequencies it is possible to excite several resonances in the xenon atoms, and the rubidium lines are not visible. As the frequency of operation goes below 320 MHz, the resonances switch to excite only the rubidium atoms and the xenon emission lines are no longer visible. The spectra in Fig. 3.2 have been normalized, but it is clear that as the resonance switch from xenon to rubidium the spectral lines of

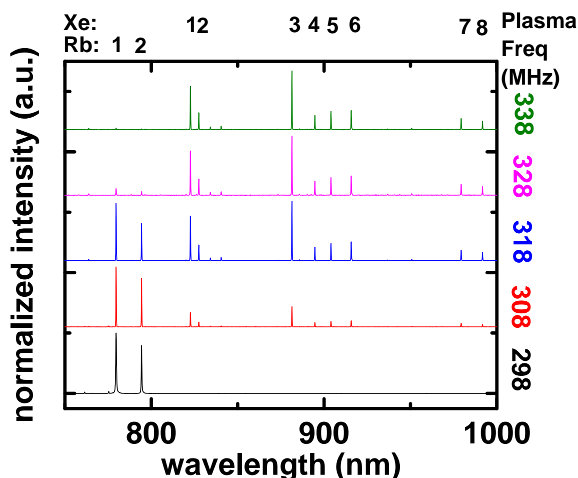


Figure 3.2: Spectra taken of the RF excited miniature bulbs of rubidium gas. The bulbs are filled with both rubidium and xenon. The dominance of one set of excitation over the other is seen as a function of the frequency of the RF drive. At 298 MHz, the rubidium lines are dominant in the spectrum. As the RF drive frequency increases past 300 MHz, the xenon spectra begins to dominate until the rubidium lines are not visible at all.

xenon are brighter as well.

The switch from xenon to rubidium resonances can be seen by measuring the ratios of the resonance lines of the two atoms. Several ratios are measured. First is the ratio between the 780 nm and 823 nm lines. These lines are labeled Rb 1 and Xe 1 in Table 3.2, respectively. They are two of the strongest resonances in rubidium and xenon. The ratio of these two lines will indicate the ratio of excitation of the two atoms in the plasma bulb.

First the ratio of lines as a function of the plasma excitation frequency is studied. Figure 3.3 shows a plot the ratios of the 780 nm and 795 nm lines in rubidium (Rb 1 and Rb2), and of the 882 nm and 823 nm lines of xenon (Xe 3 and Xe 1). The ratio of these lines are fairly consistent. This indicates that the changes

Table 3.2: Identification of spectral lines see in Figure 3.2 [56, 55].

Spectral Lines of Rb I, Xe I in plasma bulb (Å)	Transition
Rb 1 7800	$5^2S_{1/2}$ - $5^2P_{3/2}$
Rb 2 7948	$5^2S_{1/2}$ - $5^2P_{1/2}$
Xe 1 8231	$6s[3/2]_2$ - $6p[3/2]_2$
Xe 2 8280	$6s[3/2]_1$ - $6p[1/2]_0$
Xe 3 8819	$6s[3/2]_2$ - $6p[5/2]_3$
Xe 4 8952	$6s[3/2]_1$ - $6p[3/2]_2$
Xe 5 9045	$6s[3/2]_2$ - $6p[5/2]_2$
Xe 6 9163	$6s[3/2]_1$ - $6p[3/2]_1$
Xe 7 9800	$6s[3/2]_2$ - $6p[1/2]_1$
Xe 8 9923	$6s[3/2]_1$ - $6p[5/2]_2$

in the frequency does not affect the states to which the atoms are excited, so that when they decay and emit their resonances the ratios remain fairly constant. This data covers the range from 35 MHz to 500 MHz. Figure 3.4 shows the ratio of the 823 nm (Xe 1) and 780 nm (Rb 1) lines over the same frequency range. From 318 MHz to 378 MHz, the plasma bulb shows a transition from emitting the rubidium lines to emitting the xenon lines. The ratio changes from 0.8 to 62 over a range of 60 MHz. It would be most sensitive to frequency variations when one set of lines is just visible, for example when the ratio is near 43 (avoiding the region where the ratio is insensitive to the frequency). Conservatively assuming 25% sensitivity to the ratio yields a 5.6 kHz accuracy in the frequency, or an accuracy of 5.6×10^{-6} at 1 GHz.

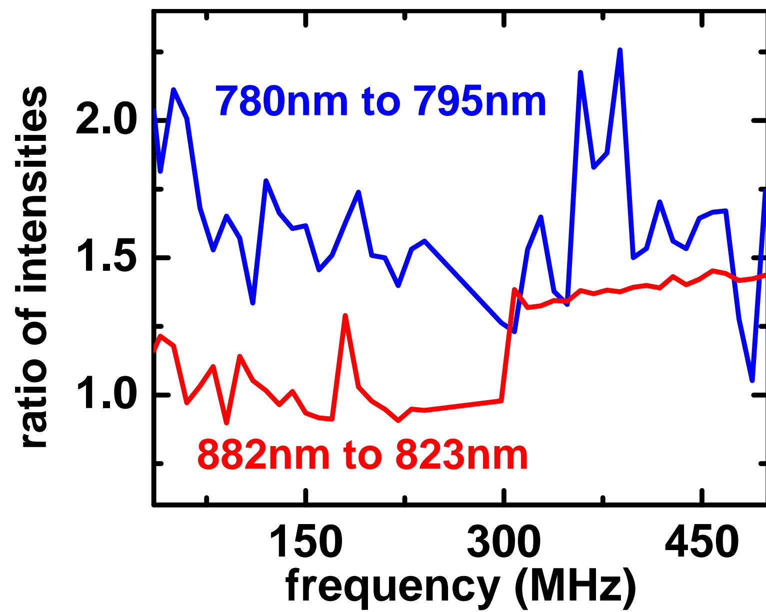


Figure 3.3: Ratio of the intensities of the spectral lines as a function of the RF drive frequency (see Fig. 3.2). The red line shows the ratio of the 882 nm to the 823 nm, dominant lines of xenon. This ratio is nearly flat; at lower frequencies, the xenon lines are barely visible which causes the ratio to be variable. The ratio of the 780 nm to the 795 nm lines are shown in blue, the lines of rubidium.

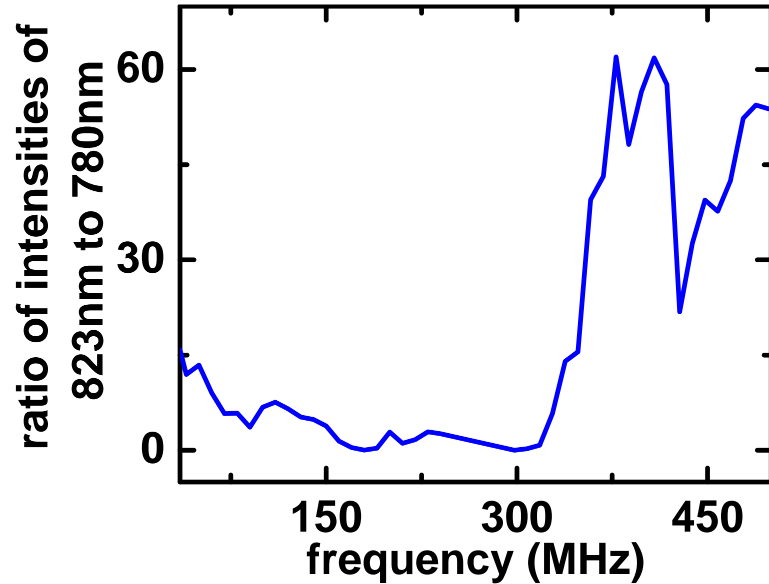


Figure 3.4: Ratio of the line intensities of 823 nm to 780 nm, i.e. ratio of xenon to rubidium (see Fig. 3.2). Near 300 to 350 MHz, there is a sharp transition from rubidium domination to xenon.

The ratio of the lines are also studied as a function of the applied power. In these data, the plasma is driven at 180 MHz and the applied power is swept from -6 dBm to 8 dBm. As in the data above, there is a transition as the plasma introduces xenon or rubidium lines as a function of the applied power. Figure 3.5 shows the spectra at -6 dBm, -3 dBm and 0 dBm, the region where a transition occurs whereby the xenon emission lines are visible then decrease in strength. At low powers the rubidium lines and a weak seat of xenon lines are visible. As the power is increased by 6 dBm, the xenon lines disappear and the rubidium lines dominate the spectrum. The same lines are visible as identified in Table 3.2.

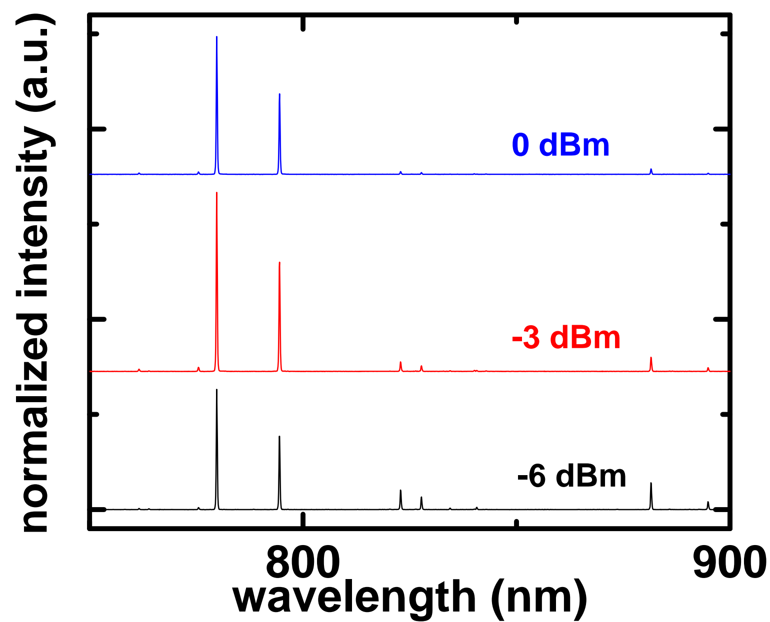


Figure 3.5: Spectra of the miniature bulb rubidium plasma as a function of the RF power. At lower powers at 180 MHz, both the rubidium and a weak set of xenon line are visible. At higher powers, the xenon lines get weaker and the rubidium spectral lines dominate the spectrum.

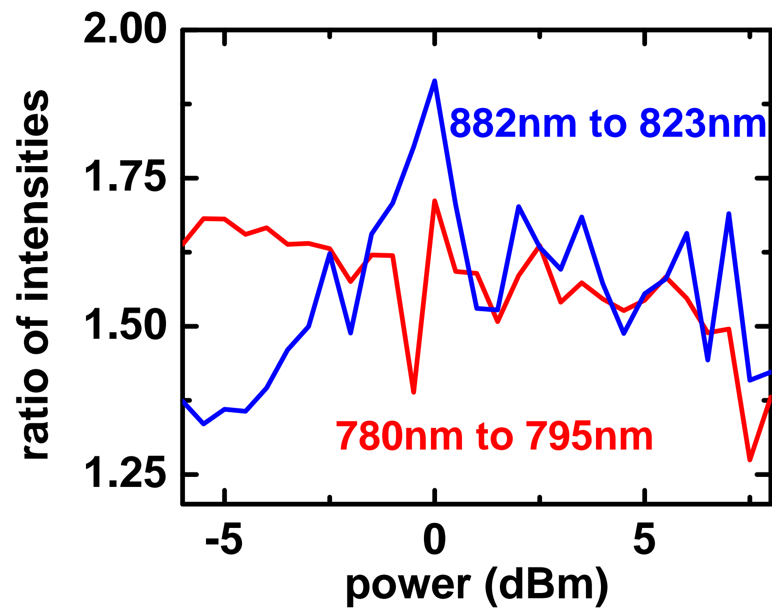


Figure 3.6: Data similar to Fig. 3.4: ratio of same-atom spectral lines, but now as a function of RF drive power. See Fig. 3.5). As seen in the frequency dependence, the ratio of the lines of rubidium in red and the ratio of the lines of xenon in blue are fairly constant over the range of applied RF power.

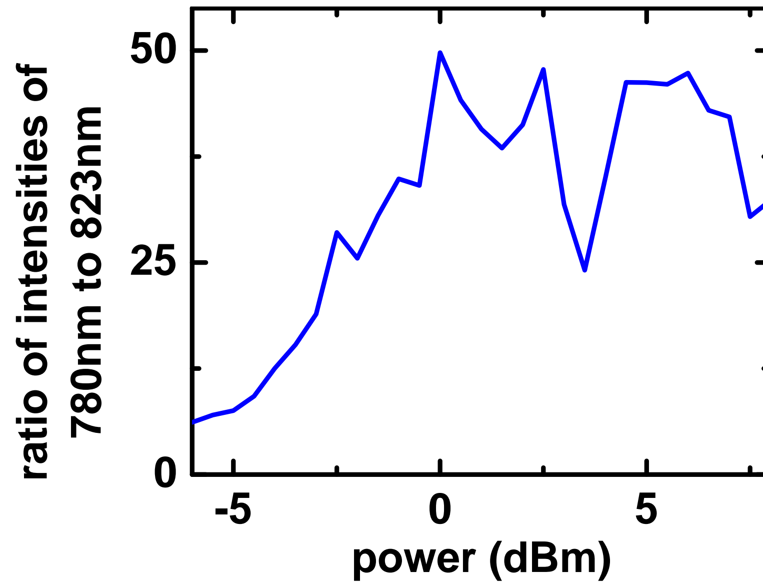


Figure 3.7: Data similar to Fig. 3.3: ratio of 780 nm (rubidium) to 823 nm (xenon), but now as a function of RF drive power (see Fig. 3.5. At low powers, the xenon spectral lines are visible with respect to the rubidium lines, but as the power is increased the rubidium lines are dominant as the line intensities of xenon decrease.

Data shows the dependence of the line strengths on the applied power. As before, the ratio of the 780 nm and 795 nm lines of rubidium, and the 882 nm and 823 nm lines of xenon are shown in Figure 3.6. As compared to the frequency dependence in Figure 3.3, the ratio of the lines per atom have less variation, as the frequency is kept constant. Figure 3.7 shows the power dependence of the ratio of the 780 nm to the 823 nm (Rb 1 to Xe 1) lines. As in the frequency data above, there is a transition of the ratio. At lower powers the xenon lines are visible along with the rubidium lines; the xenon lines are about 16% of the rubidium line strengths. As the applied power is increased, the xenon lines start to disappear. Starting from about 0 dBm and on to higher powers, the ratio of the 780 nm to 823 nm (Rb 1 to Xe 1) lines reaches a maximum of 50.

3.2 Self-powered Self-Calibrated Photon Source

This is section is a reprint of a manuscript printed in Applied Physics Letters [77].

While the plasma lamps are very low power, applications for portable atomic clocks or remote sensing require much lower power consumptions than are required to drive the plasma. Therefore, self-calibrated photon sources powered by radioactive thin films were studied.

Radioactive materials are attractive candidates for power in micro scale systems due to their small size, integratability, and high power densities. Electrical power generation has been demonstrated [40]; this work demonstrates a photon source powered by the beta decay from Ni^{63} with an endpoint energy of 67 keV. Simply, xenon gas is exposed to the radioactive nickel. The beta rays excite the xenon atoms, which then emit their decay lines. Since these lines are atomically precise, the photons are at certain characteristic wavelengths. Hence, this system can be used as a self-powered, self-calibrated photon source.

Electromechanical and betavoltaic power generation using radioactive thin films have been studied as power sources for microelectromechanical systems[36, 40, 17]. Radioactive power is attractive at the micro-scale because of the potential for high energy density and long life using isotopes with half-lives as long as a hundreds of years, which could eliminate frequent replacement in a portable device. Certain metallic radioactive sources are easy to integrate into a device because these metals, specifically ^{63}Ni, can be conveniently plated. ^{63}Ni also has a very long half-life of approximately 100 years and produces only low-energy β decay, making

it relatively easy to shield. In addition to generating electromechanical power, it is feasible to use radioactive materials to provide a low-intensity self-powered photon source[25] for micro-scale devices. It should be noted that some commercially available products such as emergency exit signs use radioactive materials combined with phosphors and scintillators for self-powered illumination. Directly excited by β decay from radioactive ^{63}Ni, gases and scintillators can produce optical emission as they relax. For gases, this is a nearly direct conversion of the radioactive power to optical output, and the sharp nature of atomic spectra can be used as a zero-power wavelength standard for application in calibrating microsystems. In this letter, the near-IR scintillation of xenon gas due to excitation by the β decay of ^{63}Ni was studied. Intensity measurements of xenon IR emission due to ^{241}Am α decay have been reported by others[8, 73].

The radioactive source used for these experiments was a plate of aluminum with dimensions 5 mm by 5 mm, onto which a 4 mm by 4 mm area was electrolessly plated with 37 MBq of radioactive ^{63}Ni. The ^{63}Ni source was placed inside a windowed chamber which was pumped down to approximately 10^{-4} Pa. After the chamber was sealed off from the pump, research-grade natural xenon was bled into the chamber. For direct imaging, the optical emission was observed by a cooled, avalanche-gain charge-coupled device (CCD) camera (Photometrics Cascade 512B). The CCD array has a working range from 350 nm to 1000 nm. The images were taken such that the normal vector of the source plate is parallel to the image plane, so that the images are viewed from the side.

Intensity contour plots of direct images of the optical emission taken at 20 kPa and 100 kPa are shown in Fig. 3.8. The grayscale implies the intensity. The

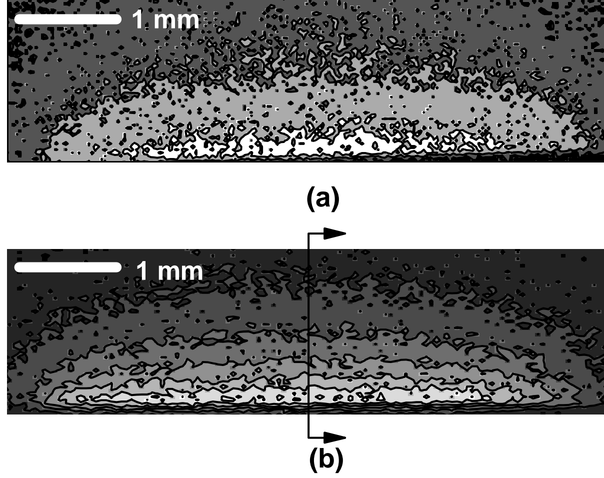


Figure 3.8: Observed xenon scintillation contour plot of direct images at (a) 20 kPa and at (b) 100 kPa. The ^{63}Ni is located at the bottom of each plot. They are qualitatively similar, but the higher xenon density at 100 kPa increases the xenon-electron interactions per unit distance, leading to a larger gradient of the optical intensity in (b).

total emission, including that not collected by the CCD, is approximately 2×10^5 photons/sec at 20 kPa and 1×10^5 photons/sec at 100 kPa. These numbers were calculated using the quantum efficiency of the CCD provided by the manufacturer (60 %), the gains associated with the electronics of the CCD camera (160), and extrapolation to 4π of solid angle. This is a low intensity of 10^{-10} W/m^2 at a distance of 5 mm for 4π of solid angle, and can be increased by using a larger amount of ^{63}Ni or a radioactive material with higher specific activity.

Spencer's theory of electron penetration in an infinite medium[62] calculates the electron flux and energy deposition for an infinite plane source of monoenergetic electrons using Lewis' equation of electron transport[39] and a function-fitting method. Summarizing the analysis begins with Spencer's form of Lewis' equation

of electron transport,

$$
\begin{aligned}
-\frac{\partial I}{\partial r} + \cos\theta\frac{\partial I}{\partial z} \quad &= \int d\Omega' N\sigma(r,\Theta)\{I(r,\theta',z) - I(r,\theta,z)\} \\
&+ (4\pi)^{-1}\delta(z)\delta(r-r_0). \qquad (3.1)
\end{aligned}
$$

The electron flux between θ and $\theta + d\theta$ with residual range between r and $r + dr$ at a point z away from the electron source is $2\pi I(r,\theta,z)\sin\theta d\theta dr$. In the integral, $\sigma(r,\Theta)$ is the scattering cross section and N is the number of atoms per unit mass of the scattering medium. The last term defines the electron source, here as an isotropic source; r_0 is the residual range of the scattering medium at the source energy. Equation (3.1) is expanded in spherical harmonics and spatial moments of the electron flux are defined by,

$$I_{ln}(t) = \int_{-1}^{1} dx x^n I_l(t,x), \qquad (3.2)$$

where t and x are the scaled residual range and z, respectively. After calculating the spatial moments of the electron flux using Eq. (19) of Ref.[62], the spatial moments of the energy dissipation curve are calculated by,

$$J_n(t) = \int_{0}^{1} dt(dT/dt)I_{0n}(t,x), \qquad (3.3)$$

where dT/dt is the stopping power of the scattering medium. The integration is simplified by approximating dT/dt by a sum of powers in t, as in Eq. (23) of Ref. [62]. Then the energy dissipation curve is approximated by a sum of a fitting function,

$$
\begin{aligned}
J(x) \approx &\quad \sum_i a_i F(\beta_i,x)^\gamma, \qquad (3.4) \\
F(\beta,x)^\gamma = &\quad \beta^{-1}\left(1-\frac{x}{\beta}\right)^\gamma \exp\left\{\frac{-Ax}{\beta-x}\right\},\ 0 \le x \le \beta \\
= &\quad 0,\ x > \beta, \qquad (3.5)
\end{aligned}
$$

where $F(\beta, x)^{\gamma}$ is a suggested fitting function given in Eq. (33) of Ref. [62] and A is calculated using Eq. (34) of Ref. [62]. The coefficients a_i and parameters β_i are calculated by solving the equations of spatial moments of the energy dissipation curve,

$$J_n = \sum_i a_i F_n(\beta_i). \tag{3.6}$$

The electron energy dissipation curve calculated from Spencer's theory was compared to the optical emission data at 100 kPa. Even though the spectrum of the radioactively-pumped xenon has several peaks, their energies are less than 7 % different. Then the energy dissipated by the electrons and the number of photons emitted by the xenon should correlate, and it is expected that the energy dissipation curve will sufficiently describe the optical intensity. Four terms of the fitting function in Eq. (3.5) were used with γ=1. Stopping-power and range data for xenon were acquired from the NIST ESTAR database, and slight changes to the power series approximation of dT/dt were made accordingly. The differential elastic scattering cross-sections were calculated by using the tables of Bullard and Massey[10], where Born's approximation was applied to the Thomas-Fermi theory of atoms; past experiments show good agreement at electron energies as low as 800 eV[4].

It is also necessary to sum over the spectrum of energies of the β decay. The ^{63}Ni atoms undergo β^- decay, whereby a neutron decays into an electron and an antineutrino. The decay energy is distributed as kinetic energy between the electron and antineutrino, resulting in a broad spectrum of energies for the ejected β particle. The endpoint energy of ^{63}Ni, the maximum energy of the β decay, is 67 keV and the spectrum peaks around 14 keV, the statistical mode. The

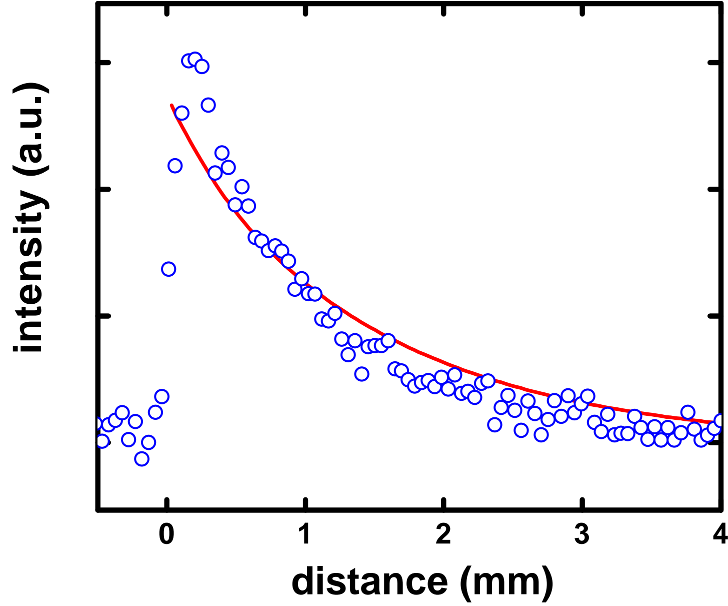


Figure 3.9: Optical intensity profile at 100 kPa, as indicated by sectional arrows in Fig. 3.8. Line shows result from Spencer's theory using isotropic angular distribution of β decay.

theoretical energy spectrum is described by the Fermi theory of β decay[34, 18], and experiments have shown good agreement for ^{63}Ni[2, 20]. Other experiments have shown disagreement at lower energies, even differing among similar samples[32, 29]. For the calculations presented here, the theoretical allowed β spectrum was summed in steps of 5 keV up to 50 keV, above which negligible change in the energy dissipation curve was seen.

Fig. 3.9 shows the profile of the optical emission at 100 kPa as a function of distance from the ^{63}Ni face, extracted from the direct image in Fig. 3.8. The curve calculated using Spencer's theory shows moderate agreement with the data. It is interesting to note the decrease in total optical intensity as the pressure was raised from 20 kPa to 100 kPa. The data show a decrease by a factor of two, while Spencer's theory predicts a factor of three.

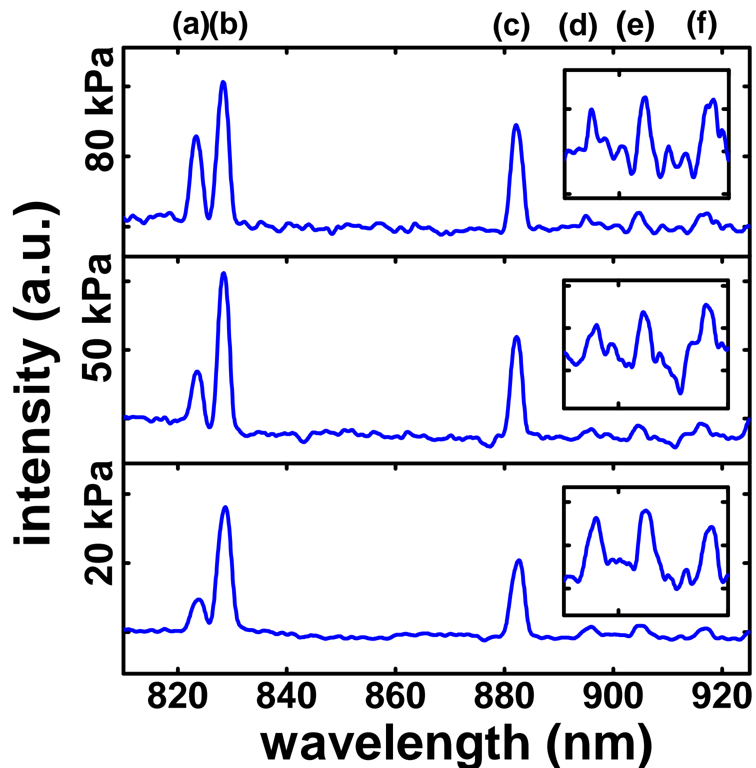


Figure 3.10: Spectra at 20, 50, and 80 kPa. Insets show less prominent peaks between 890 nm and 920 nm. Six peaks are labeled by the letters at the top of the figure.

Spectral data were taken using a 0.300 m Czerny-Turner monochromator using a grating with a blaze wavelength of 750 nm. The spectrum was detected by the CCD camera. Due to the low intensity of the radioactively-pumped xenon, the setup aimed to maximize signal strength at the cost of spectral resolution. As in the direct images, the emission was viewed from the side: emission through a 4 mm-thick sample of radioactively-pumped xenon was captured by the CCD, compressed into a slit-like image less than 2 mm wide. Using a simple compound microscope system, f/1.18 lenses collected about 0.18 π of solid angle. A large 1 mm slit width was used to increase the optical throughput. Finally, the low dark current of the CCD allowed for eleven-minute integration times.

Table 3.3: Spectral lines identified between 810 nm and 925 nm[55]. 1: Due to proximity of adjacent lines, identification was also based on published line intensities.

Spectral Line of Xe I		Transition
(a)	823.16336	$6s[3/2]_2$ - $6p[3/2]_2$
(b)	828.01162	$6s[3/2]_1$ - $6p[1/2]_0$
(c)	881.94106	$6s[3/2]_2$ - $6p[5/2]_3$
(d)	895.22509[1]	$6s[3/2]_1$ - $6p[3/2]_2$
(e)	904.54466	$6s[3/2]_2$ - $6p[5/2]_2$
(f)	916.56520[1]	$6s[3/2]_1$ - $6p[3/2]_2$

The optical emission that is viewed by the CCD camera is spatially asymmetric and broad, as seen in Fig. 3.9. This results in a spectrum with broad peaks. However, each peak of the xenon spectrum is nearly monochromatic and will disperse equally in the monochromator. Then the shape of the optical emission should be the same for all the spectral lines. The raw data was cross-correlated with the zeroeth-order image of the optical emission,

$$\xi(\lambda) = \int_{-\infty}^{\infty} S(\lambda + \lambda')Z(\lambda')d\lambda', \tag{3.7}$$

where S is the raw spectrum, Z is the zeroeth-order image, and λ is the wavelength. This sharpened the peaks and corrected a small offset in their wavelengths, and increased the signal-to-noise ratio from five, helping bring out several weaker spectral lines.

The processed spectrum from 810 nm to 925 nm for a range of pressures is shown in Fig. 3.10, each spectrum representing two sets of data. The maximum resolution of the system is approximately 0.5 nm and the Doppler broadening is expected to

be less than 0.001 nm, but the data reflect a much lower 2.8 nm resolution. The three large peaks were found to be the only prominent features within the working range of the CCD. Three weaker lines were also found in this region and have been included, for completeness, in the identification of the spectral lines in Table 3.3. The data suggest that despite electron energies up to tens of keV, very little of the xenon atoms ionize. The strong lines of singly-ionized xenon are in the visible region of the spectrum, around 500 to 700 nm. Those identified in Table 3.3 are for neutral xenon. These lines also indicate the formation of metastable states. The $6s[3/2]_2(^3P_2)$ state is a highly metastable state, relaxing to the ground state through a magnetic quadrupole decay. Depending on the isotope, the lifetime of this state can be between several or upwards of forty-three seconds, as has been measured by Walhout, Witte, and Rolston using a magneto-optical trap[72].

3.2.1 Alternative sources and applications

It is also possible to excite other sources for different applications. For example, detection of carbon typically relies on its IR vibrational and rotational spectra. In Fig. 3.11, the presence of carbon was detected by its excitation by ^{63}Ni. The lowest wavelengths of the simplest vibrational bands start at several micrometers, which is beyond the range of the CCD imager used. It is likely that UV to visible wavelengths are being detected by the CCD camera. Optical emission in sub-IR has been observed previously [19, 60], attributed to transitions between higher order vibrational states. The detectable image may also be a result of resonances in ionized carbon dioxide CO_2^+, which forms as a result of the energetic beta decay from the ^{63}Ni [67]. Therefore, it is found that radioisotopes can be used to detect

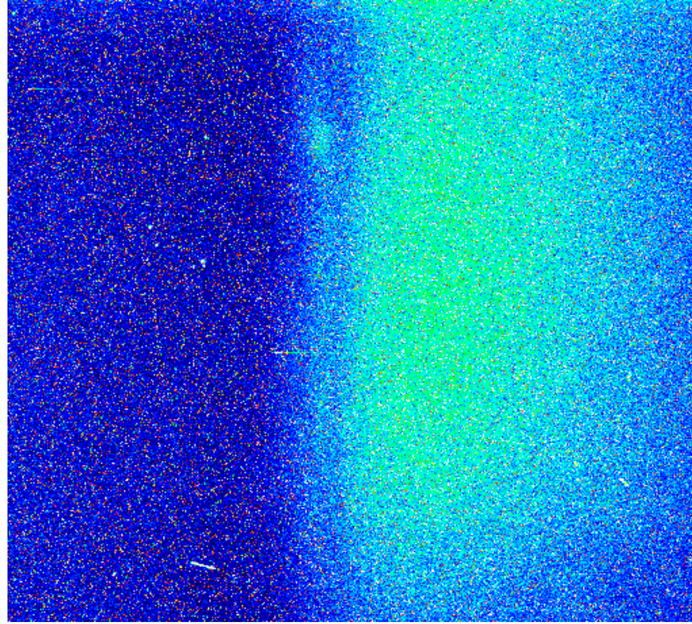


Figure 3.11: Excitation of CO_2 by ^{63}Ni.

more complex molecules, in addition to simpler atomic gases. Detection of CO_2, for example, could be for a low power greenhouse remote gas sensor.

In addition, Fig. 3.12 shows that GaAs can be excited by ^{63}Ni. The high energy ^{63}Ni electrons excite the direct band gap GaAs. The band gap limits to mininum energy, or highest wavelength, photon that is detected. The photon spectrum does not extend higher than 870 nm, the GaAs band gap at room temperature. This suggest that excitation of semiconductor materials for optical applications is possible using ^{63}Ni as well. For instance, radioisotopes could be used to power stimulated emission optical amplifiers in long term applications, such as intercontinental underwater optical communication channels that are not easily accesible.

Therefore, it is seen that there are a number of possible applications of ^{63}Ni excited materials. Such possibilities include optical sources, gas sensors, and optical

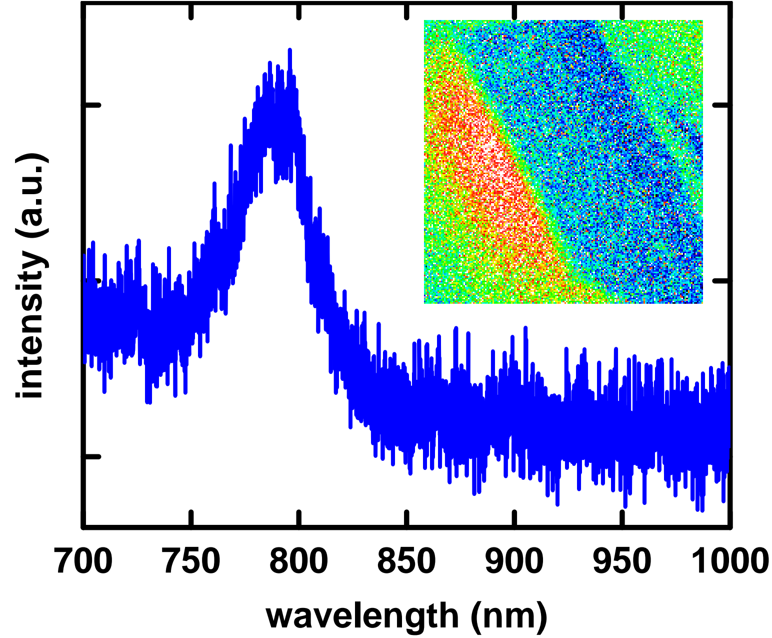


Figure 3.12: Excitation of GaAs by ^{63}Ni, and spectrum. The broad spectrum stops short of the band gap of 870 nm (at 300K).

amplifiers and devices.

3.3 Focusing by thin magnetic plates: design by evolutionary algorithms

The beta decay from radioactive ^{63}Ni, by its nature, covers much of its phase space. When a ^{63}Ni decays, one of its nuclear neutrons undergoes a β^- decay,

$$n \rightarrow p^+ + e^- + \bar{\nu}_e + KE, \tag{3.8}$$

where the neutron decays into a proton, electron, antineutrino, and kinetic energy, respectively. This processes must conserve both energy and momenta (from the rest frame of the neutron and proton). Without the antineutrino, the energies of

the electron would be separately fixed (α decay, for example, has a nearly monoenergetic spectrum). However, the presence of the antineutrino underspecifies the conservation of energy and momenta whereby the daughter particles can have a range of kinetic energies. For ^{63}Ni, the energy spectrum for the β particle is mostly decided by the statistical part, dictated by the equal probability of energy distribution between the electron and the antineutrino. There is a slight correction due to the Fermi function, caused by the electrostatic interaction between the charged particles as the electron moves away. The beta decay in ^{63}Ni has an endpoint energy of 67 keV, that is to say there is 67 keV energy that is shared between the electron and antineutrino. In addition, there is no preferential direction to the weak force-mediated beta decay (although there is a weak magnetic field dependence in direction). Thus, as found in the theoretical work in the previous section, the beta decay is very nearly isotropic. As a result the beta decay covers a large portion of phase space. While Louiville's theorem dictates that it is not possible to unambiguously decrease the phase space distribution, it is possible to narrow the distribution in its spatial distribution. In an optical application, for example, it may be desired to focus the beam spatially to generate a large amount of light at a small point to generate a very bright source. Here, a monte carlo nonrelativistic finite difference time domain model of ^{63}Ni beta decay is used to simulate the spatial focusing using rotationally symmetric shapes cut into thin ferromagnetic films. By using evolution algorithms, an increase in the focusing up to 11.3% of the total beta decay into a 500μm radius spot is possible.

The magnetization of the 2.5mm-thick ferromagnetic film is assumed to be perpendicular to its surface (through its thickness direction). A rotationally symmetric hole is cut into the magnet, which serves as a focusing tunnel or lens. The

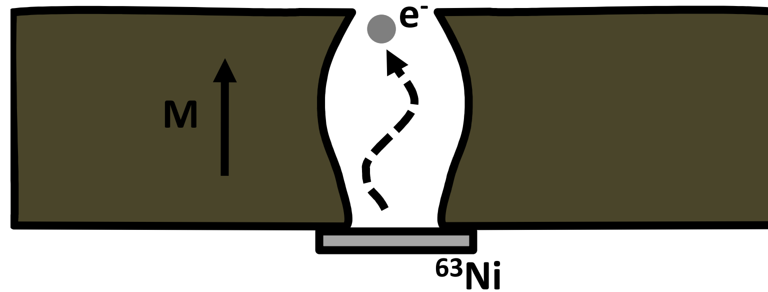


Figure 3.13: Geometry of thin film magnet focusing of ^{63}Ni beta decay. A rotationally symmetric curve is cut into the ferromagnet, which is magnetized along the direction of propagation.

shape of the hole is evolved to maximize the fitness, here defined to be the number of electrons that escape within the 500μm at the end of the hole.

At each step of the evolution, the magnetic field of each agent is calculated by first calculating the surface currents on the wall of the hole due to its magnetization, its corresponding vector potential field $\vec{A}$, then numerically calculating the elliptical integrals to calculate the corresponding magnetic field. The electron trajectory is calculated in cylindrical coordinates and using a rotating coordinate system that follows the electron as its rotational angle changes. The starting point of an electron is randomly positioned, its energy is chosen from its theoretical energy spectrum, and its initial vector is chosen to be isotropically oriented.

In these simulations, five agents were evolved simultaneously. The shapes are randomly designed initially, each described by ten parameters. At each step, the agents who have a fitness greater than the average fitness survive to the next step. Then, each agent has a probability of 1/2-1/2cos πF of surviving, where F is the fitness of the agent. Next, agents are propagated by a probability equal to (1/2-1/2cos $\pi F1$)$\times$(1/2-1/2cos $\pi F2$), where F1 and F2 are the fitnesses of the

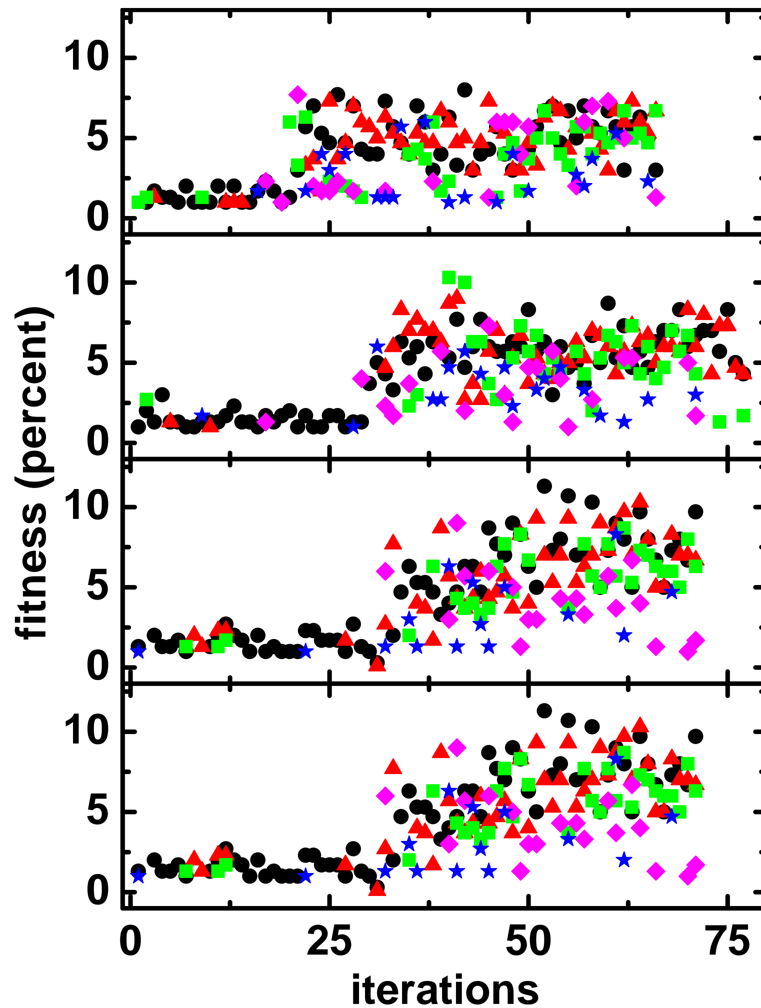


Figure 3.14: Results of a Monte Carlo nonrelativistic finite difference time domain model of ^{63}Ni beta decay to simulate the spatial focusing using rotationally symmetric shapes cut into thin ferromagnetic films. All four runs show qualitatively similar results. First, there are about 25 to 30 iterations where the fitness is three percent. Then there is a sudden increase in the fitness of an agent, who by the rules of propagation generates a progeny. Efficiency can be as high as 11.3%. There is some variation as the agents continue to mutate and mix, but the overall fitness does not seem to increase.

two agents. If they propagate, the shape of the progeny is determined by using the average of the ten parameters that describe the magnet shape. New agents are then generated if required. Finally, there is all agents undergo a 1% chance of mutating, where one of the ten shape parameters is changed randomly. Random numbers for all parts of the simulations were generated by the MT19937 Mersenne Twister (with a period of $2\hat{1}9937$-1) developed by Makoto Matsumoto and Takuji Nishimura.

The results from four runs are shown in Fig. 3.14. Agents are first tracked when their efficiency reaches above 1%, which may take up to 30 or 50 steps of evolution. All four runs show qualitatively similar results. First, there are about 25 to 30 iterations where the fitness is three percent. Then there is a sudden increase in the fitness of an agent, who by the rules of propagation generates a progeny. Often, this agent will overwhelm the rest of the population and the whole population jumps to fitness values up to 11.3%. There is some variation among the agents after this happens, probably as a result of a high amount of mutation and mixing among the agents. While the simulation continues to run, the populations do not seem to be able to further improve their fitness.

It is found, therefore, that passive ferromagnetic thin films can be used to enhance the current density of the β-decay of radioactive thin films. Rather than the electrons isotropically leaving the thin film, a magnetic thin film helps capture some of these electrons by effectively focusing them onto a small spot. In the case of radioactive istropic sources which occupy a large region of phase space, it is difficult to analytically design an appropriate magnetic thin film structure, requiring effectively a numerical guessing method. Here, this was implemented by an evo-

lutionary algorithm. This focusing enhances the power output of radioactive thin films which, although they have high energy densities, have low power densities. The passive focusing allows the distribution in the phase space to be skewed to increase the number of electrons that pass through the target area, although the resulting cost is a larger spread in electron energies as they pass through this target region. The increase in electron density is especially important when using low count radioactive sources, which may easily be down to only thousands of electron decays per second.

CHAPTER 4

LOW-POWER BEAT CLOCK

In this section, the quantum phase noise limit of semiconductor lasers is first presented, then derive the minimum Allan deviation as a result. Next, a modest performance low-power beat clock is presented and, based on this result, show that an atomic lock to a Doppler broadened atomic resonance can achieve enough frequency stability to reach the quantum phase noise limit of the Allan deviation. The results are used to compare semiconductor laser Allan deviation data from the literature.

First, the Allan deviation limit due to phase fluctuations as a function only of the laser linewidth is derived. This result was incorrectly derived by Yamamoto [75], differing by a factor of π. Ohtsu et al. estimated the Allan deviation limit but as a function of a number of parameters [47], some of which are not necessarily known in literature or in product specifications. The analysis can be used to compare semiconductor laser hetrodyne systems. For example, harmonic frequency chains and optical frequency combs bridge the optical and microwave frequency domains and allow precise measurements of optical frequencies, and measurements of laser stabilization often involve a hetrodyne stage [57, 68, 14, 76].

The phase fluctuations in the optical field of a VCSEL caused by spontaneous emission events are first considered [27]. Each event i at time t_i with random phase θ_i at an average spontaneous emission rate R_{sp} causes a phase fluctuation in the lasing field of $\Delta\phi_i = I^{-1/2}\left(\sin\theta_i - \alpha\cos\theta_i\right)$ so that the phase change at some time

t' is,

$$\Delta\phi(t') = \frac{1}{\sqrt{I}} \sum_{t_i \in (0,t')} \left(\sin\theta_i - \alpha\, \cos\theta_i\right), \tag{4.1}$$

where I is the number of photons in the cavity and α is the linewidth enhancement factor. Relaxation oscillations are ignored, which die out on the order of nanoseconds [28, 69]: this transient behaviour is inconsequential when calculating the phase over a sampling time $\tau \gg$ nanoseconds. The relaxation oscillations add a peak in the spectral density of frequency fluctuations near the relaxation resonant frequency ω_R [69], but from Eq. 4.9 it is clear that it contributes little to the Allan deviation unless the sampling time $\tau \ll 1/\omega_R \approx$ nanoseconds. The autocorrelation is then,

$$R_{\phi EE}\left(\Delta\phi(t_1)\Delta\phi(t_2)\right) = \frac{1}{I}\, E\,[\sum_{t_i \in (0,t_1)} (\sin\theta_i - \alpha\, \cos\theta_i) \times \sum_{t_i \in (0,t_2)} (\sin\theta_i - \alpha\, \cos\theta_i\,)], \tag{4.2}$$

where $E\,[\,]$ is the expected value. If $t_2 > t_1$,

$$\begin{aligned} & R_{\phi EE}\left(\Delta\phi(t_1)\Delta\phi(t_2)\right) \\ &= \frac{1}{I}\, E\,[\sum_{t_i \in (0,t_1)} (\sin\theta_i - \alpha\, \cos\theta_i)^2 + \sum_{t_i \in (0,t_1)} (\sin\theta_i - \alpha\, \cos\theta_i) \qquad (4.3) \\ &\times \sum_{t_i \in (t_1,t_2)} (\sin\theta_i - \alpha\, \cos\theta_i\,)] \\ &= \frac{1}{I}\, E\,[\sum_{t_i \in (0,t_1)} (\sin\theta_i - \alpha\, \cos\theta_i)^2 + \sum_{t_i \in (0,t_1)} \sum_{t_j \in (t_1,t_2)} (\sin\theta_i - \alpha\, \cos\theta_i) \qquad (4.4) \\ &\times (\sin\theta_j - \alpha\, \cos\theta_j\,)] \end{aligned}$$

(4.5)

The $\sin^2\theta_i$ and $\cos^2\theta_i$ terms have expected values of one half, and all other terms are zero. The summation yields the average number of spontaneous emission events

$R_{sp}\, t_1$ so that,

$$R_{\phi EE}\left(\Delta\phi(t_1)\Delta\phi(t_2)\right) = \frac{R_{sp}\, t_1}{2I}\left(1+\alpha^2\right) \qquad \text{for} \quad t_2 > t_1. \tag{4.6}$$

Similarly,

$$R_{\phi EE}\left(\Delta\phi(t_1)\Delta\phi(t_2)\right) = \frac{R_{sp}\, t_2}{2I}\left(1+\alpha^2\right) \qquad \text{for} \quad t_1 > t_2. \tag{4.7}$$

Now, with $\omega(t') = \frac{\partial\phi(t')}{\partial t'}$, $R_{\phi EE}(\Delta\omega(t_1)\Delta\omega(t_2)) = \frac{\partial^2}{\partial t_1 \partial t_2} R_{\phi EE}(\Delta\phi(t_1)\Delta\phi(t_2)) = R_{sp}(2I)^{-1}(1+\alpha^2)\delta(t_2 - t_1)$. Then the Wiener-Khinchin theorem yields the power spectrum of frequency fluctuations $S_{\Delta\omega} = R_{sp}(2I)^{-1}\,(1+\alpha^2) = \Delta\omega$, where the last equality comes from Henry's form of the modified Schalow-Townes formula [27]. Our result differs from the derivation of Yamamoto [75] by a factor of π, but agrees with the result of Vahala and Yariv [70] who calculated the field equations and included a Langevin noise source. In addition, $y = \Delta\omega/\omega_0$ so that $S_y = S_{\Delta\omega}/\omega_0^2 = \Delta\omega/\omega_0^2$.

So far only the autocorrelation of a single laser has been considered. Since the beat frequency is being measured, the phase of the intensity field is of interest. For two lasers with a beat of ϵ and random phases $\Delta\phi_1$ and $\Delta\phi_2$ the intensity field is,

$$\begin{aligned} E^* E &= \left[\exp\left(i\omega t\right)\exp\left(i\Delta\phi_2\right) + \exp\left(i(\omega+\epsilon)t\right)\exp\left(i\Delta\phi_1\right)\right] \\ &\times \left[\exp\left(-i\omega t\right)\exp\left(-i\Delta\phi_2\right) + \exp\left(-i(\omega+\epsilon)t\right)\exp\left(-i\Delta\phi_1\right)\right] \\ &= 2 + 2\cos\left(\epsilon t + \Delta\phi_1 - \Delta\phi_2\right), \end{aligned} \tag{4.8}$$

so that $\Delta\phi = \Delta\phi_1 - \Delta\phi_2$. But $R(\Delta\phi(t)\Delta\phi(t+\tau)) = E\left[(\Delta\phi_1(t) - \Delta\phi_2(t))(\Delta\phi_1(t+\tau) - \Delta\phi_2(t+\tau))\right] = E\left[\Delta\phi_1(t)\Delta\phi_1(t+\tau)\right] + E\left[\Delta\phi_2(t)\Delta\phi_2(t+\tau)\right] - E\left[\Delta\phi_1(t)\Delta\phi_2(t+\tau)\right] - E\left[\Delta\phi_2(t)\Delta\phi_1(t+\tau)\right]$. The last two expected values are both zero, since ϕ_1 and ϕ_2 are uncorrelated. Each of the first two terms is the autocorrelation of any single

laser, so the autocorrelation of the intensity is just twice the autocorrelation of the electric field of a single laser, $R_{\phi II} = 2R_{\phi EE}$, and $S_{\Delta\omega}$ must be doubled. That is to say that the electric field is being used as the frequency standard, except the presence of two laser cavities leads to twice as many spontaneous emissions. Simply doubling the spontaneous emission rate, $R_{sp} \rightarrow 2R_{sp}$, leads to the same factor of two to appear in our calculations.

Finally $\sigma_y^2(\tau) = \int_0^\infty S_y(f)\, 2\left[\sin^4(\pi\tau f)\right](\pi\tau f)^{-2} df$ [12, 54] so that,

$$\begin{aligned} \sigma_{y,sp}^2(\tau) &= 2\int_0^\infty \frac{\Delta\omega}{\omega_0^2}\frac{2\sin^4(\pi\tau f)}{(\pi\tau f)^2} df \\ &= \frac{\Delta\omega}{\omega_0^2}\frac{1}{\tau} = \frac{\Delta f}{{f_0}^2}\frac{1}{2\pi\tau}. \end{aligned} \tag{4.9}$$

Therefore the Allan deviation of the beat frequency due to the spontaneous emissions in the lasing cavities is,

$$\sigma_y(\tau) = \frac{1}{f_0}\sqrt{\frac{\Delta f}{2\pi\tau}} \tag{4.10}$$

The beat frequency source setup is shown in Fig. 4.1. Two VCSELs at the top and bottom of the Figure are locked to two D2 hyperfine lines. A single-mode 780nm VCSEL with a linewidth ≈ 100 MHz (U-L-M Photonics ULM780-01-TN-S46FOP) is temperature stabilized and injection current modulated sinusoidally with a frequency dither of 200 MHz. The frequency modulation amplitudes are matched by comparing the hyperfine splittings measured by the two VCSELs and by minimizing the change in the beat frequency. The injection current to the VCSELs is approximately 1.5 mA and stabilized with 150 μF shunt capacitors,

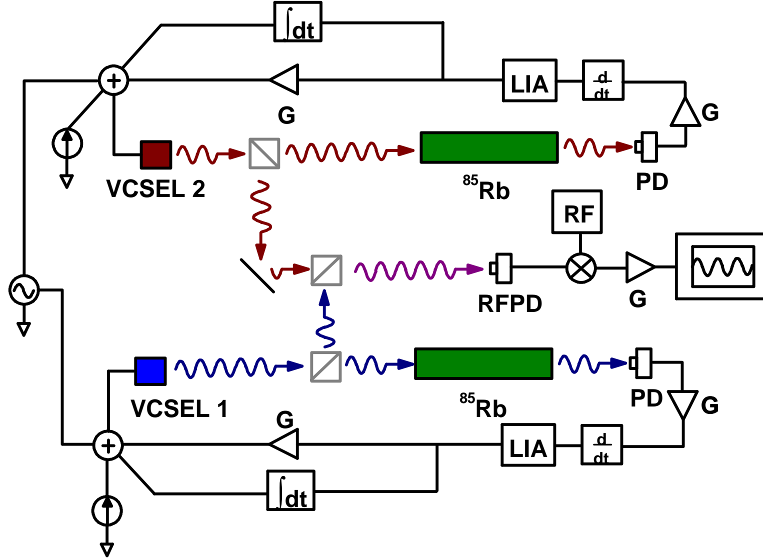


Figure 4.1: Beat frequency source setup. A VCSEL is locked onto a ^{85}Rb transition using a proportional-integral feedback loop. The laser is then mixed with another VCSEL at the other hyperfine split frequency, and the beat frequency is detected by a photodetector.

with a resulting optical power ≈ 100 μW at the beat frequency photodetector. The signal generated at the photodetector is approximately -62 dBm peak (at 1 MHz bandwidth). The laser absorption through a ^{85}Rb gas cell (55 ^{o}C) is measured, and the signal is transimpedance amplified and its derivative is fed to a lock-in amplifier with a time constant of 100 μs. Proportional and integral error signals are fed back to a summing junction at the VCSEL, resulting in a lock of the VCSEL to a D2 hyperfine absorption line of ^{85}Rb.

A beam splitter diverts one laser beam to be mixed with another VCSEL with a similar setup, generating the beat frequency. The linear polarizations of the two VCSELs are aligned. The beat frequency is detected by a photodetector (Electro-Optics Technology ET-4000A). To accommodate the time interval counter and high gain amplifier, it was then necessary to mix the signal with a 2.4 GHz radio frequency source down to less than 1 GHz. The intermediate frequency was then

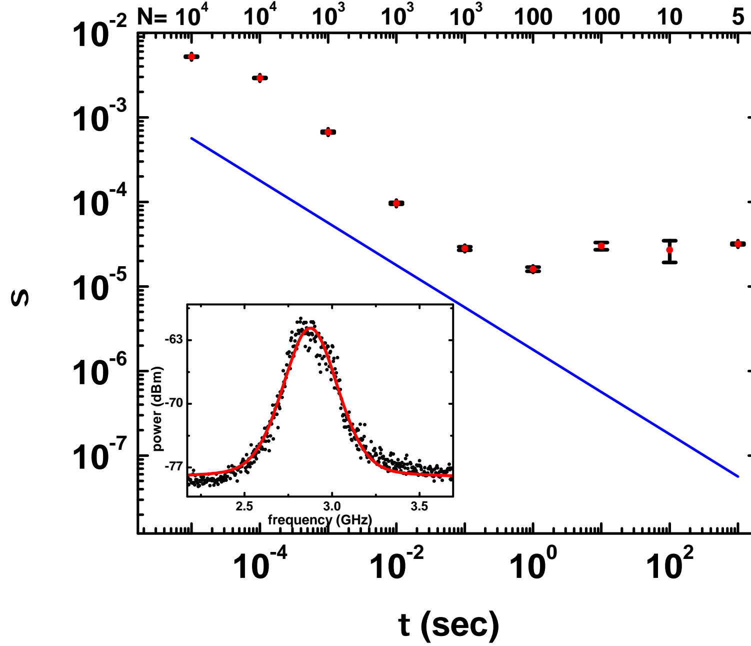


Figure 4.2: Allan deviation data. The red points with error bars are experimental data, and the blue line is calculated by Eq. 4.10. Numbers along top axis denote sample size. Around τ=1 sec, the Allan deviation flattens out due to flicker noise at the mixer. Inset shows the power spectrum of the beat signal at the hyperfine frequency. Black dots show data; the red line shows fit to a pseudo Voigt profile (cf., [71]).

amplified by a 60dB amplifier (MITEQ AM-1646).

The VCSELs are locked to the two hyperfine-split D2 ($5\,^2S_{1/2}$ - $5\,^2P_{3/2}$) Doppler-broadened transitions of ^{85}Rb. There is hyperfine structure in both the ground and excited states, and as a result the beat frequency could be between 2.822 and 3.100 GHz [3] depending on the absorption strengths of the different transitions. The measured beat frequency was 2.849 GHz. The beat between the F = 4,3 and F = 2,2 transitions is 2.852 GHz, which are the strongest hyperfine transitions [7].

Measured Allan deviations and Eq. 4.10 are shown in Fig. 4.2. Note that f_0

= 2.85 GHz since the microwave, not the optical, frequency is being used as the oscillator frequency. It is found that contributions to the frequency deviation due to the control circuit, e.g. current shot noise in the VCSELs, are negligible and not included.The Allan deviation suffers for integration times below milliseconds, which are time scales shorter than the time constant of the lock-in amplification. The Allan deviation continues to follow Eq. 4.10 until it flattens out around $\tau \approx$ 1 sec. This is a result of flicker noise at the output of the mixer. It is noted that similar flicker noise starting at $\tau \approx 1$ sec can be duplicated by replacing the beat frequency photodetector with a low-power RF source, but is removed as the power is increased.

Reported Allan deviations of lasers can reach as low as 10^{-15} or lower due to the high quality factor and large resonance frequencies. In addition, linewidths can be narrowed with higher power and larger cavities [27], yielding improved Allan deviation at the cost of power and size. Using Eq. 4.10, the performance of the frequency stability of hetrodyne measurements of semiconductor lasers can be compared taking into account such factors. For example, micro- and nano-mechanical resonator performance is often characterized by the fQ product; here, the semiconductor optical resonator performance is characterized by its $\sigma\mathrm{f}/\sqrt{\Delta f}$. In Fig. 4.3, the Allan deviations of several published works are compared[6, 5, 1, 43, 78], including the work here in Fig. 4.2. However, the Allan deviations are normalized using Eq. 4.10. As expected, the saturation sub-Doppler limited and cavity-locked systems perform better than Doppler-limited or cavity-free systems. Overall, a direct comparison of the Allan deviations shows a factor of 10^7 among the data sets shown here. However, the phase noise limit metric reveals that performances vary by about 10^4. In particular, the comparison of data sets c and f

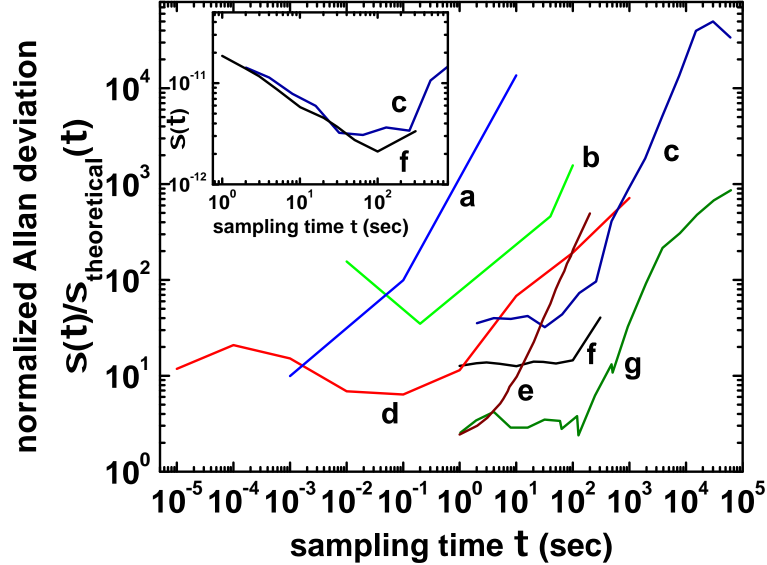


Figure 4.3: Comparison of normalized Allan deviations, i.e. Allan deviation divided by the phase noise limit Allan deviation in Eq. 4.10. Sources and laser power, linewidth, and locking limit are **a**: 3 mW, 21 kHz, sub-Doppler limited and cavity stabilized[6]; **b**: 5 mW, 30 MHz, Doppler limited[5]; **c**: 3 mW, 300 kHz, Doppler limited [1]; **d**: this work, 100 μW, 100 MHz, Doppler limited; **e**: unknown, 0.5 Hz, cavity stabilized[43]; **f**: 0.2 MHz and 3 MHz, 7 mW and 4.5 mW, sub-Doppler limited[78]; **g**: 3mW, 300 kHz, sub-Doppler limited[1]. It is noted that that sub-Doppler or cavity stabilization can yield lower normalized Allan deviations (**e**, **f**, **g**) over Doppler limited stabilization, as expected. Inset: While it is clear from the normalized Allan deviation that a sub-Doppler limited system (**f**) yielded better results than a Doppler limited system (**c**), their actual Allan deviations are similar due to more inherent phase noise in **f**.

are highlighted in the inset of Fig. 4.3. Set c[1] is a Doppler-limited system, while f[78] is a saturation, sub-Doppler limited system. A direct comparison of their Allan deviations in the inset appears nearly identical. However, a comparison of their Allan deviations normalized by Eq. 4.10 shows that given the quality of their semiconductor lasers and feedback systems, the saturation, sub-Doppler limited system[78] is in fact performing better.

CHAPTER 5

NANO OPTICAL RULER IMAGING SYSTEM FOR NANOMETROLOGY

The work so far described has developed wavelength and timing precision based on atomic stability. The work here describes a new type of nanometrology, especially suitable for scanning probe microscopes. The nanometrology is for positioning. The SI unit of length is based on the electromagnetic radiation. It is defined to be the distance traveled by light (in vacuum) in 1/299792458 second: that is, the speed of light has been defined to be 299792458 meters per second, where the second is defined by the ground state hyperfine splitting in Cesium atoms. From an engineer point of view, direct metrology of this can be difficult for very high resolution. A 10 GHz counter, for example, could only resolve 3 cm and a resolution of just 100 μm would require a 3 THz counter. Therefore, the only practical way so far to engineer high precision, accuracy, and resolution metrology is to measure an optical standing wave. The standing wave is a function of its wavelength, which is in turn related to the speed of light constant $c=\nu\lambda$. The wavelength can be well known by atomic resonances. A distance can be measured by confining light into an optical mode, for example in a Fabry-Perot cavity, where resonances and antiresonances occur at geometric distances that are some function of the optical wavelength. A simple example is a Michelson interferometer, where the differential position can be measured as the difference in optical path lengths between the two mirrors. In that case, the optical intensity at the output would go as $1+\cos(2\pi\delta L/\lambda)$, where δL is the optical path length difference. There are several types of nanometrology tools currently available. The simplest are linear optical encoders. Along the travel of a stage (usually a stepper stage to allow for

large travel), an actual ruler of sorts is attached. This track will have apertures cut into it to allow light to pass through. The metrology is done by measuring the transmission through the track; as the stage moves, the number of holes passed is counted as a square wave at the optical detector and the distance is measured from the fabricated distance between holes. More than one aperture track can be cut wit different duty cycles, resulting in distance calculation by detecting multiple channels. These optical encoders are often fabricated down to parts of a micron, limited by fabrication limits, and interpolation of the optical signal may result in a precision down to tens nanometers or more. While optical encoders are cheap and easy, the precision depends entirely on how well the track is fabricated and how well it is stabilized in a dynamic environment. Often these encoders will measure the stage translation to no better than a percent or so over the whole translation motion. On a four inch wafer, this can be as high as millimeter. The common alternative for precision nanometrology of large travel stages (e.g. for writing photo masks) is an interferometric system.

In addition, flexural piezo stages are required for nanoscale work that requires positioning below the often 100 nm or greater resolution of stepper motors. For example, piezoelectric tubes are a necessity in scanning probe microscopes. Metrology for piezo actuators is especially important because they can suffer greatly from hysteresis, creep, and drift. This a result of the transduction of the piezoelectric material, which requires a large electric field and great strain, which results in charge, mechanical creep, etc. The primary method of nanometrology for piezoelectric actuators is a simple calibrated transfer function between the applied voltage and the displacement. This conversion can be calibrated by scanning a known calibration sample. This calibration can work fairly well for small regions of dis-

placement, but do nothing to account for hysteresis, creep, and drift. The result is that even moving a few microns can result in buildup of positioning errors. As an alternative, flexural piezo stage can be fabricated with extra metrology. Capacitive sensors measure the capacitance between a flat surface of the stage and a fixed reference plate. These can provide very high precision, down to less than a nanometer, but capacitance is a highly nonlinear phenomenon and has a very low range. The capacitance, varying inversely to the distance, will have very high sensitivity at small gaps and have quadratically decreasing sensitivity as the gap increases. Strain gauge sensors are low cost solutions and can have precision down to nanometers, but suffer from even smaller ranges (before the strain sensor would yield) and have poor linearity and temperature sensitivity.

This work concentrates primarily on nanometrology for scanning probe microscopes. In addition to the metrology shortcomings described above, there are some practical issues related to the combination of metrology and stage technologies as well. A mechanical problem lines in the large masses that are moved. In a stepper motor, the stage which holds the sample is up to ten inches per side covering at least its maximum travel range. With a significant thickness, the stage will have considerable mass. Systems with XY interferometeric metrology will also require a large reflective block to be mounted along two edges of the stage, further adding mass. The high mass results in a low mechanical bandpass of the stage. This means a stage cannot move quickly around the wafer to different points. A fast interferometer, for example, may be able to run at ten meters per second (16 MHz at 632 nm), but stages do not reach those speeds because the acceleration is not high enough. In wafer-scale nanofabrication or device testing, this can become a heavy cost in time. On a six inch wafer, for example, a possible grid of 10 nm squares

for access by a scanning probe microscope would have 10^{15} addressable positions. In addition, the separate metrology of the large travel stage and the continuous actuation piezoelectric element results in a saturation of the net nanometrology. The actual position of the scanning probe tip over the sample is as poor as both the nanometrology of the piezo head (high precision locally) or the nanometrology of the stepper stage (poor).

The system described here aims to eliminate or reduce many of these problems. The nano optical ruler imaging system (NORIS) generates a precise optical ruler in space using the diffraction pattern from a microfabricated grating generated by an atomically stabilized laser. The schematic is shown in Fig. 5.1. The section 1 shows the stabilized external laser cavity. The laser is atomically stabilized as discussed in section 2.1.2, stabilized to a tens of ppb. Section 2 shows some optical path length, which is used to allow the laser beam to diffract and increase its beam width. Section 3 shows the generation of the diffractive aperture array. A wafer is microfabricated with an aperture array of a reflective thin film. The thin film pattern acts as a diffractive element, creating an optical pattern in space which is used as a nano-precise optical ruler. Its pattern in a two dimensional slice shown as a number of red spots in the plane of the scanning probe. The scanning probe, here shown as a scanning tunneling tip, is attached to a imaging camera. The camera captures an image of the optical ruler, then uses that information to calculate its position using sub-pixel digital image processing techniques, identified as section 4.

It is instructive to find first approximations to noise sources and how they may affect the performance of the system. For example, a 1 nm precision over a six

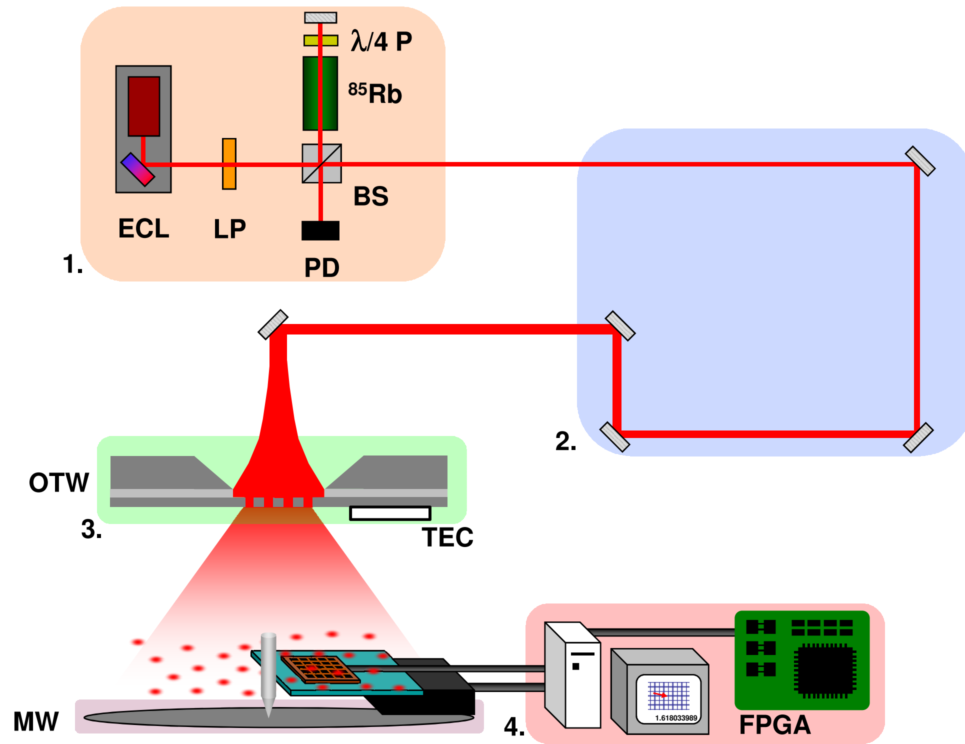


Figure 5.1: Schematic of NORIS. Section 1 shows the atomically stabilized laser. An external cavity laser is stabilized to the hyperfine structure of ^{85}Rb, to a stability of a few parts per billion by saturation spectroscopy. In section 2, the laser beam is propagated and its beam width is increased slightly by allowing it to traverse a path length. Section 3 shows the optical transfer wafer, which is used to generate a diffraction pattern. The wafer has a suspended thin film metal aperture array, which diffracts and forms the optical ruler. The optical ruler projects above the manufacturing wafer, or any region of space where the nanometrology is to take place. The aperture array thin film is temperature controlled by a thermo-electric cooler. A camera mounted on an actuator is used to image the optical ruler, which allows the system to detect its position. Here, the camera is shown attached to a scanning tunneling tip. Section 4 shows the digital signal processing required to run the system and control the position. Cross correlation methods must be used to detect the position: a fully programmable gate array is shown, which would be a fast method of calculating cross correlations.

inch wafer requires a precision of 6×10^{-9}. The precision in section 1 is determined by the linewidth and the atomic locking of the laser. The laser linewidth is 100 kHz, which is a precision $\delta\lambda/\lambda$ of 3×10^{-10}. The D2-line at 780nm is used, which is the transition from the transition $5^2P_{3/2}$ excited state. The natural lifetime is 26.2 ns [9, 59, 53], or a linewidth of 38 MHz. A good lock to the atomic transition can yield a sweep of one tenth of the linewidth, which results in $\delta\lambda/\lambda=1\times10^{-8}$. This could well be reduced yet again by clocking the diffraction image. As learned in 4, while the frequency of the laser may dither up to tens of MHz (in that case, up to 500 MHz), the frequency stability of the laser while it's dithering can still lead to low phase noise. Therefore, this last figure may be reduced up to an order of magnitude or more by clocking the illumination of the diffraction grating only when the laser is at the line center of the saturated absorption line of rubidium.

The gross optomechanics of the system are considered, namely the air lifted table and the many optics of the system, in order to isolate our experiment from the uncontrolled environment. It is best to learn from the work in mechanical isolation for scanning tunneling microscopes, which are able to run experiments with sub Angstrom resolution [49]. The general strategy is to use the vibration isolation of the table to low pass filter mechanical noise, then build the scanning probe microscope (and the optics, etc) with very high mechanical resonances. Mechanical noise is reduced by having minimal overlap between these two pass bands. In systems such as scanning probe microscopes where there is some mechanical movement inherent in the system (and thereby may excite vibrations in the system), the strategy is to build those parts in-between the two passbands, so that it will not be excited by noise passed through the air lifted table, nor excite modes in turn of the optics and related systems thereby diminishing the performance of the

complete system. This idea is seen in Figure 5.2. The table has a low pass corner frequency of 30 Hz, and the optomechanics are estimated from finite element modeling to have resonance frequencies of no less than 250 Hz. This creates a stability window between 30 and 250 Hz where the system can be moved with high stability. A mechanical working frequency of 250 Hz is low compared to, for example, commercial scanning probe microscopes but is a result of the large system with which this works. From this the approximate error can be estimated that will be acquired by the optomechanics, which causes a change in the laser frequency. Note that as long as the diffraction grating is uniformly illuminated, there is no change in the optical ruler image. The frequency error $\Delta f = d\phi/dt \approx F_{noise} I_{width} T^2/mQ$, where F_{noise} is the noise force, m is the system mass, I_{width} is the spectral width of the noise, Q is the quality factor, and T is the transmission of the bandpass. It is estimated that the error produced is $\delta L/L = 10^{-10}$.

Errors from the optical wafer are estimated. First, the spontaneous emission in the laser beam can cause spurious phase noise in the laser wave front which causes a relative phase shift between the apertures resulting in a shifting of the position of features in the optical ruler. By estimating that $sqrt(h\nu v_g \alpha_m / 2P_0) = 10^{-3}$ [27], the error $\delta L/L$ will be less than -z $\delta\phi\lambda$ / x L 2π N, where N is the number of apertures, = 10^{-9}. The wavelength noise calculated above also yields an error in the optical ruler, $\delta L/L$ = -$z^2\delta\lambda$/xλ N = 10^{-9}.

The NORIS system is especially insensitive to variations in temperature at the diffractive aperture. This is due to its small size. Interferometric blocks made of special Zerodur, for example, have thermal expansion coefficients as low as $2\times10^{-8}/^{\circ}C$ but may cost at least a thousand dollars for a sizable, optically flat

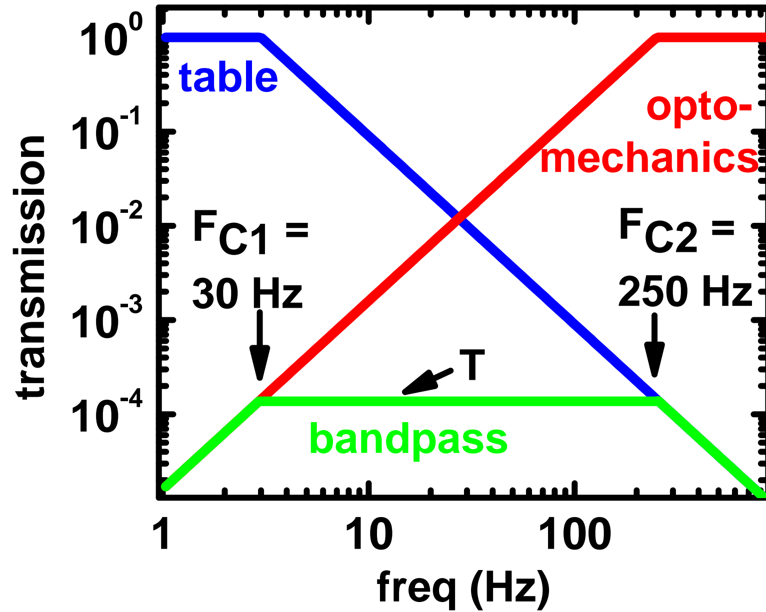


Figure 5.2: Optimizing NORIS by considering the mechanical frequency responses. The pneumatically lifted optical table acts as a low pass filter; PVC curtains minimize other, external, acoustic sources of mechanical noise from entering the system. The optomechanics are designed to have high mechanical resonance frequencies. In this Figure, the table specification of a 30 Hz corner frequency for the table is used, and a 250 Hz corner frequency of optical components has been calculated in finite element method simulations of harmonic modes. The resulting pass band shows the best frequency spectrum to actuate NORIS: the system would be unaffected by low frequency noise introduced through the table, but its movements would not excite the high frequency modes in the optomechanics.

piece. The diffractive aperture can be temperature controlled very easily because of its small size; it therefore has a small thermal mass and high conductance over a very small region. With a thermal coefficient of as much as 10^{-6}, it is estimated that a stability of $\delta L/L{=}10^{-8}$ is possible.

Finally, the CMOS imager samples a large portion of the optical ruler, which results in an increased precision in position calculation. If only in the imager itself

is considered, the noise is uncorrelated and the standard deviation will sum in quadrature and it is estimated that $\delta x/x = (\partial I/\partial x \mathrm{SNR} I x \sqrt{N_{pixels}})^{-1}$=$10^{-8}$.

The resulting precision from the noise sources considered here yields 4×10^{-8}, which is below 10^{-7}=10 nm / 50 mm.

5.1 Image processing

In addition to the optics and optomechanics, the NORIS precision depends on the digital image processing that occurs after taking the image of the optical nanoruler. This section discusses some of the aspects of the image processing.

5.1.1 Reducing noise in analogue to digital conversion

Analogue-to-digital conversion of pixel intensity is ubiquitous to most sensors, but can be detrimental to a precision system like NORIS where an undistorted image of the optical ruler is required for proper image processing. Unfortunately, the ADC conversion is limited by the imager hardware and cannot be changed. In modest CMOS imagers, for example, the ADC conversion may be a single byte, or eight bits. The resulting discretization error is 1/256 or 4×10^{-3}. In order to decrease this source of error, a method is developed that results in decreased discretization noise.

Consider a single pixel with 8-pixel ADC whose nominal intensity measurement

is I_0=161.803 at a gain of 256. The ADC will convert the intensity value to 162. Averaging will still result in a mean value of 162. Note that if more noise can be added to the system so that the intensity value will vary at least by one pixel value (i.e. down to 161 and up to 163), the discretization error can be reduced. Similar to stochastic resonance, the mean value will better approximate the nominal value I_0, since a proportionate number of values will be 161, 162, and 163 due to the noise but averaging towards the mean value of 161.803.

By changing the gain before the ADC, the discretization noise can be reduced. Since most CMOS imagers also use digital control of the analog gain, only integral gain settings will be considered. At a gain of 256, the ADC will again read a intensity of 162. Now, at a gain of 255 the ADC will see an intensity of round(161.803×255/256)=161; then the calculated intensity is 161×256/255, or 161.631. Next at a gain of 254 the ADC will see an intensity of 161 again, but the intensity is now calculated to be 162.268. At different gain settings, the calculated gain will differ slightly depending on the gain used, as seen in Figure 5.3. In order to calculate a precise value for the intensity, these calculated intensities are averaged. The result is seen in Figure 5.4.

An example of one dimensional imaging is performed using this technique. A Gaussian was imaged by a 400-level ADC, then its image was calculated by averaging or processing a number of gain settings as discussed above. Figure 5.5 shows a sampling of the Gaussian. Five percent noise was introduced per pixel, with a 1% error in the preamplification gain. The same number of images are used to calculate both sets. While both methods show the same errors in the middle region of the Gaussian where the intensity is high, the averaging method

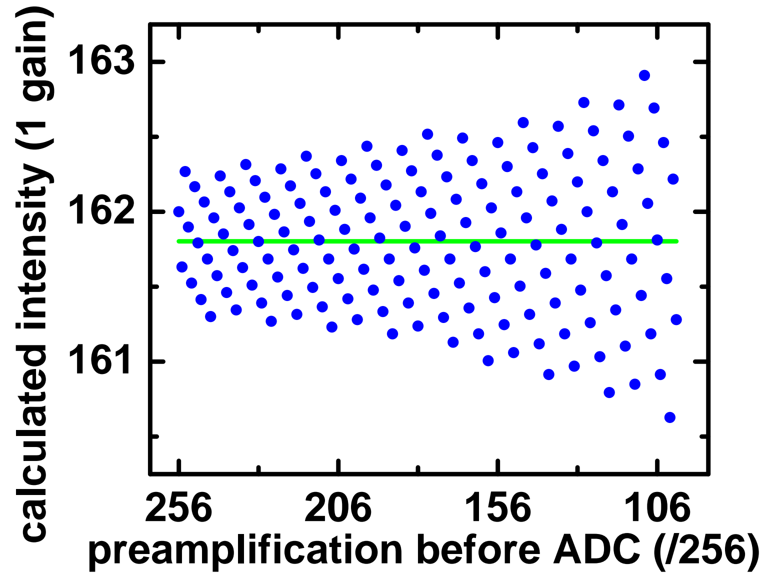


Figure 5.3: Figure shows the intensity ADC calculated at different preamplifications of the input signal. The change in amplification causes different intensities to be measurent. Taking measurements at one amplifiation biases an ADC measurement due to the discretization error.

develops errors about 40% out from the peak of the Gaussian. In this region, the discretization error is worse at low intensities and image averaging, and the result is an overwhelming increase in imaging error. In contrast, the gain sweeping method preserves the Gaussian structure to 80% of its width, at which both methods essentially see the same image. Figure 5.6 shows the effect of noise on the image precision by measuring the average variance of the pixel intensity error. Interestingly, a small amount of noise reduces the pixel error, akin to stochastic resonance mentioned above. Note the greater error in the averaging scheme over the gain sweeping method.

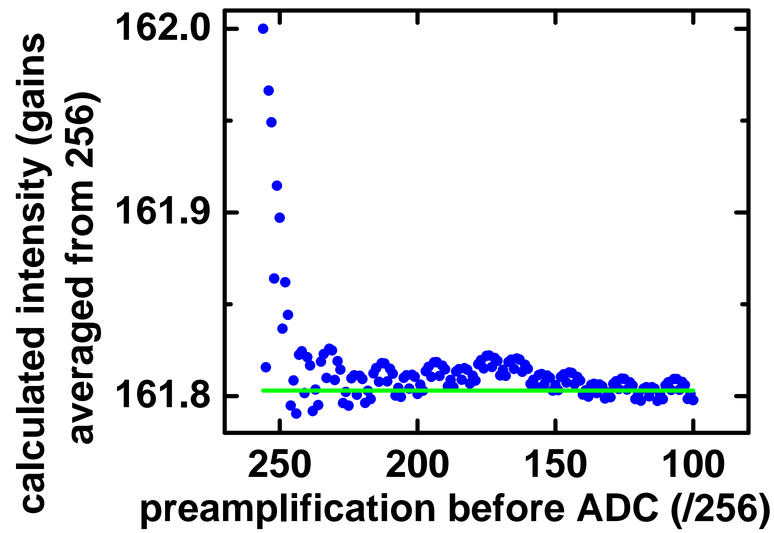


Figure 5.4: Figure 5.3 showed the different ADC intensities that were measured; by taking the combined information, the system can converge to the actual analog intensity value. The Figure shows the calculated intensity by combining the information by sweeping the gain from 256 to a setting shown in the abscissa. Even after ten or eleven gain settings are used, the intensity is very nearly close to its actual value, much improved over the value calculated from just one gain setting (at 256).

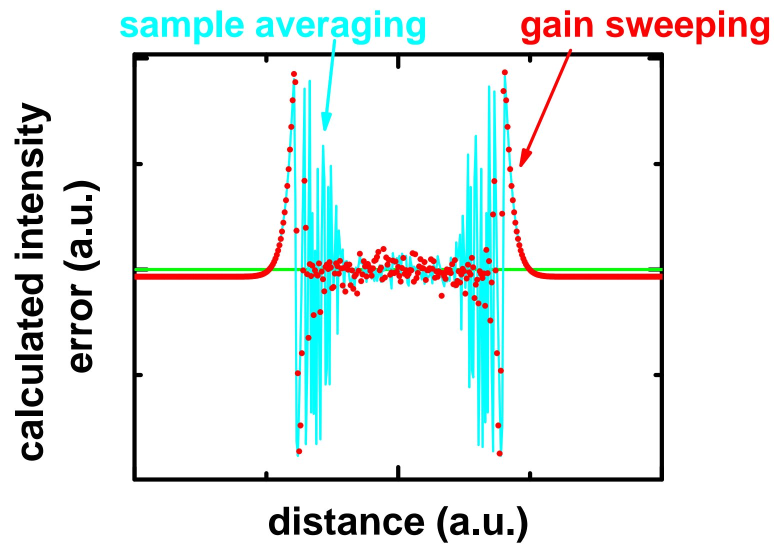


Figure 5.5: Calculated intensity error. A Gaussian function is sampled, then its curve is sampled by an ADC. Teal graph shows the calculated curve by taking a number of averages with a 400-level ADC. As the curve approaches the low intensity pixels toward the tails of the Gaussian curve, the quantization noise is too large and the error is large. The gain sweeping method described here captures these parts of the curve: this method has less noise for much more of the curve, reproducing the curve with more precision having taken the same number of images as just averaging them.

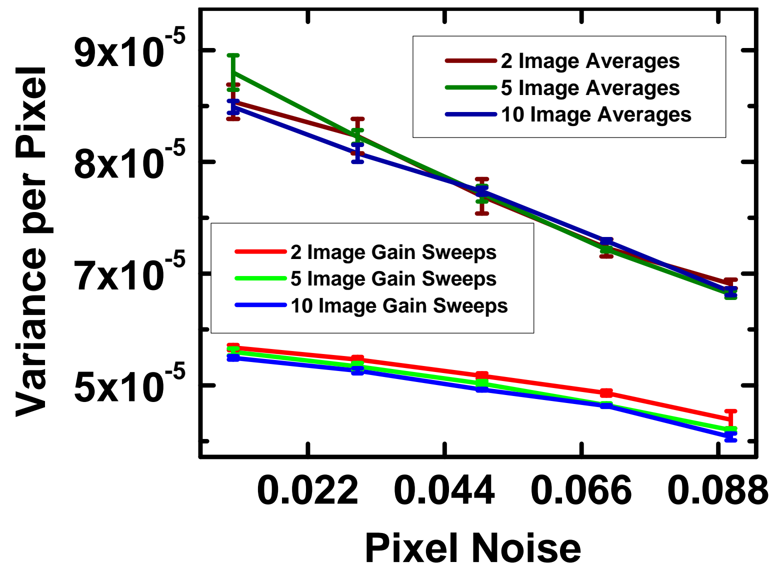


Figure 5.6: Average variance per pixel comparing averaged versus gain swept ADC images of Gaussian curve, as also shown in Fig. 5.5. The gain swept method has about half as much variance as the averaged images. The variance actually decreases as a function of increased pixel noise, a phenomenon akin to stochastic resonance.

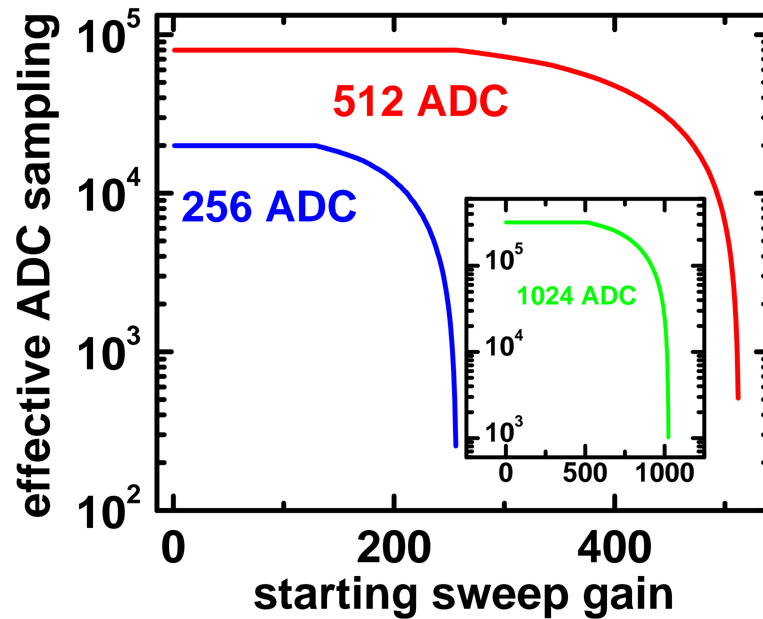


Figure 5.7: Effective ADC sampling rate using gain sweep method, where the gain is swept from the maximum value down to the starting sweep gain value. When the starting sweep gain is half as much as the starting sweep gain, all the gains are factors of higher gains so that no new information is contributed to calculating the image: the effective ADC sampling rate is seen to flatten out. Improvements between 2 and 2.5 orders of magnitudes are seen.

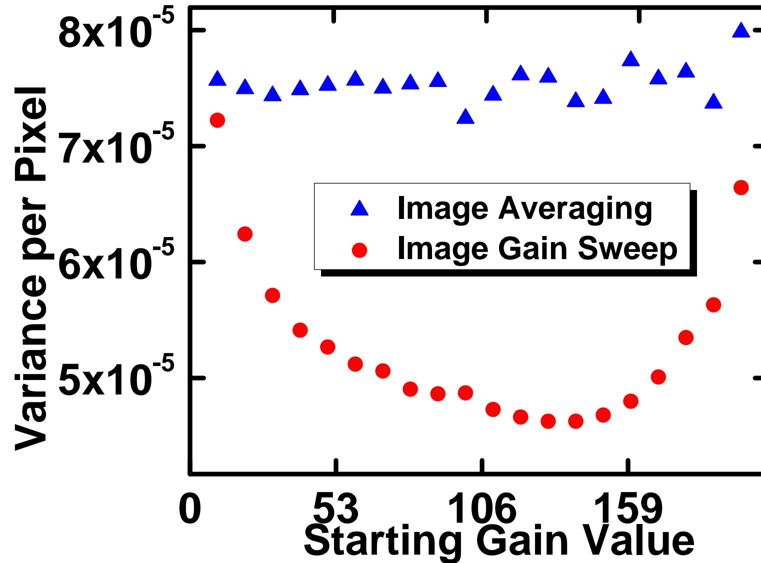


Figure 5.8: Variance per pixel as a function of starting sweep gain value. At high starting gain values, the variance decreases as the starting gain value is decreased resulting in more information gained about the image. The variance flattens out: as the starting gain is lowered, the noise eventually overwhelms the quantization noise and the variance increases again.

Finally, the question of what range of gain settings to use is considered. There are two counter balancing trends to consider. The problem with a small range of gain settings (i.e. gains 240 to 256) is that the resulting data set is small and less information is gained about the intensity value, resulting in a small improvement in the ADC measurement; see 5.4. However, it is not completely advantageous to acquire every possible gain setting. The first problem is that the pixel noise will eventually overwhelm the intensity measured at a low gain. For example, an intensity measured to 1.6 only has to have a 6% intensity noise in order to switch from an ADC of 2 to 1: at such a low gain setting, that actually results in a 50% ADC error. In addition, gains at lower settings have a higher probability of generating extra information which does not increase the measurement results. For

example, the gains with settings of 120 and 64 share the factors 2^3, and as a result eight of their gains will share the same measured intensity (8,16,24,32,40,48,56,64) for 64 and (15,30,45,60,75,90,105,120) for 120. The worst case is one gain is a factor of the other, in which case no new information is acquired. At the least, the possibility of one gain being a factor of the other and resulting in no new information gained can be eliminated by using a gain sweep that extends from the largest gain setting to a gain setting is just greater than half the largest gain setting. Not considering common factors between gain factors and assuming a maximum gain setting of M, the intensity can be resampled at $3\ (M^2+2\ M)/8$ effective ADC levels. The Figure 5.7 plots the effective sampling as a function of the starting gain sweep setting. The algorithm used here acquires data just before the point where no new information is gathered (half the maximum gain setting).

The result these two competing effects is seen in Fig. 5.8. When very low starting gain values are used, the imager can be overwhelmed by noise and there is no advantage over simply averaging many images. When high starting gains values are used, not many gain settings are sampled and little extra information is acquired by the system and there is also no improvement in imaging precision. A medium region shows the best results. Note that the averaging method sees improvement starting at low starting gain levels (i.e. just acquiring and averaging more images), but the improvement is small and negligible compared to the gain sweep method.

5.2 Aperture array design

The design of the aperture array is important to establish the baseline characteristics of the optical ruler.

5.2.1 Diffraction as a Fourier transform

Considering a diffracting surface in the (ξ, η) plane illuminated by an amplitude U(ξ, η), the diffraction pattern at the (x,y) plane at a normal distance z from the diffracting plane is calculated using the Huygens-Fresnel principle[22]. The Huygens-Fresnel principle states that all points in a propagating wavefront are sources of wavefronts themselves. The principle results in,

$$U(x,y) = \frac{z}{i\lambda} \iint d\xi d\eta U(\xi,\eta) \frac{e^{(ikr)}}{r^2}, \tag{5.1}$$

where r=$\sqrt{z^2 + (x-\xi)^2 + (y-\eta)^2}$. In the Fresnel approximation, r is binomially expanded to $r \approx (1 + 1/2(x-\xi)^2/z^2 + 1/2(y-\eta)^2/z^2)$ which yields,

$$U(x,y) = \frac{e^{ikz}}{i\lambda z} e^{i\frac{k}{2z}(x^2+y^2)} \iint d\xi d\eta \left\{ U(\xi,\eta) e^{i\frac{k}{2z}(\xi^2+\eta^2)} \right\} e^{-i\frac{2\pi}{\lambda z}(x\xi+y\eta)}. \tag{5.2}$$

In the Fresnel approximation (near-field), the diffraction can be calculated by the Fourier transform of the diffraction plane amplitude with a phase factor, which

is again multiplied by a phase factor. Furthermore, the Fraunhoffer approximation leads to further simplification where z$\gg$k$(\xi^2 + \eta^2)/2$ resulting in,

$$U(x,y) = \frac{e^{ikz}}{i\lambda z} e^{i\frac{k}{2z}(x^2+y^2)} \iint d\xi d\eta U(\xi,\eta) e^{-i\frac{2\pi}{\lambda z}(x\xi + y\eta)}. \tag{5.3}$$

which is a Fourier transform of the diffracting plane with a phase factor. The phase factors result in scaling of the Fourier transform / diffraction pattern as the imaging plane moves away from the diffraction plane (i.e. as z increases), and leads to decreased intensity of the pattern away from the axis center. However, the features of the diffraction pattern are dominated by the Fourier transform.

The diffraction as a Fourier transform of the aperture array (assuming monochromatic, planar wavefront illumination) means that certain designs should be avoided. Periodic arrays, as are well studied in crystallography, lead to periodic diffraction patterns. Periodicity implies that the pattern will have translational symmetry, which in NORIS would result in non-unique positioning information. Regular interferometeric systems suffer from translational symmetry, where the measured intensity in a cavity arrangement is spatially symmetric at integer parts of the wavelength of light. The symmetry then requires that the metrology system count the fringes as the stage travels (although that is typically not the limiting constraint for speed).

Patterns lacking translational symmetry, then, will require patterns that are not periodic. In opposite contrast to the periodic structures are the amorphous or

In Appendix B, micrographs of aperture arrays developed in chrome photomask (aperture hole $\approx 3\mu$m in diameter) and the corresponding diffraction pattern. As desired, the diffraction pattern optical ruler extends out to several inches in radius, which can be used for positioning across wafer-scale positions. While these show large coverage and apparently some small features, such designs will have some local translational symmetries (i.e. the patterns which are translations of some pattern). The patterns which are formed by a rotation may lack translational symmetry, but tend to lack have gaps in the middle of the optical ruler: a rotated pattern loses aperture density (quadratically) from its center, so that it loses low frequency features and therefore having less optical energy near the center of the optical ruler.

5.2.2 Quasiperiodic patterns

Quasiperiodic or aperiodic patterns appear to be a perfect solution. Quasiperiodicity arises from infinite tilings of the plane which lack translational symmetry. By inventing the first aperiodic tiling set, Roger Berger disproved Hao Wang's conjecture that the so-called Domino problem of whether a set of tilings can fill the plane is undecidable. There are several quasiperiodic tilings, some requiring up to tens of thousands of tiles. Tilings with a deep connection to the physical sciences are the Penrose tilings discovered by Roger Penrose and independently discovered by Roger Ammann. They have been identified in medieval Islamic architecture [41]. The P3 rhombus tilings are one of the most well-know, whose tilings are only two tiles of rhombuses whose constituent triangles have edges of the golden ratio. Quasiperiodic arrays are important for several reasons. Because they lack trans-

lational symmetry their Fourier transform is very dense. Rather than filling the k-space with peaks at the inverse of the periodicity, their quasiperiodicity maps all the points map to different places in the k-space. The result is a set of very sharp peaks at very many places in k-space. This is contrast to amorphous or otherwise nonperiodic structures whose Fourier transforms will, rather than moving energy into these Fourier peaks, will spread out that energy everywhere among the spectrum (showing no structure). However, the quasiperiodic arrays will densely fill the diffraction plane as well which will allow high transmission of laser intensity and therefore increase the optical signal-to-nose ratio at the optical ruler imaging plane.

The advantage of using a quasiperiodic over a periodic pattern can be seen by the simulated data shown here. Using the Fraunhoffer approximation, the precision of the nanometrology was simulated. The sizes of the image is 2 mm per side. The distance between the diffraction and imaging planes is 25.4 mm. The laser wavelength is 780nm, with an incident plane wave with a homogenous intensity across the diffraction plane. The aperture array is a set of 9662 holes of separation constant 10 μm with holes of diameter 3μm. See Figure 5.10 that shows a calculated Penrose tiling diffraction pattern optical ruler, and an optical ruler from a simple square lattice. Note the high density of features in the Penrose tiling diffraction pattern, as opposed to that for the square lattice.

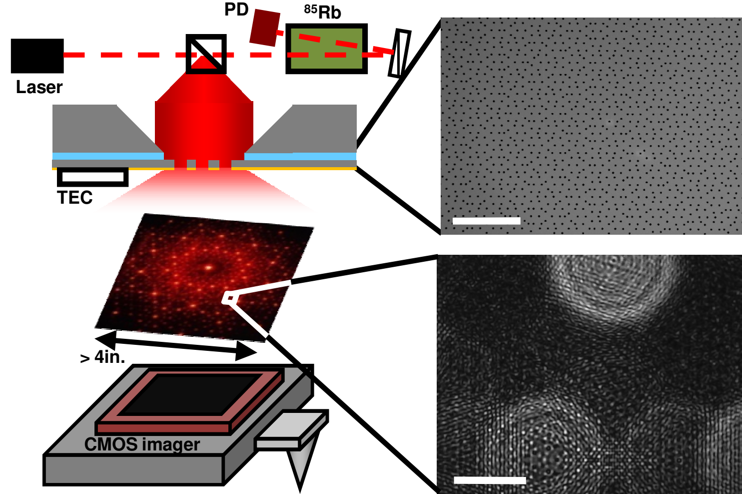


Figure 5.9: Schematic of NORIS with quasiperiodic pattern. SEM shows holes in silicon-on-insulator silicon patterned as a Penrose tiling, a periodic pattern (bar shows 100 μm). Image shows resulting diffraction pattern optical ruler (bar shows 600 μm).

The quasiperiodic pattern generates an optical diffraction pattern that is translationally asymmetric, thus preventing positioning ambiguities in constrast to a periodic pattern. Therefore, unlike other interferometric optical methods large displacements can be measured without a half-wavelength ambiguity. In addition, the quasiperiodic pattern has a Fourier transform that is dense in $\vec{k}$-space, i.e. a diffraction pattern that is dense in real space. The expected performance can be estimated based on the mean square error (MSE) of the image registration of offset $\mathbf{r}$ using the Cramer-Rao bound,

$$\mathrm{MSE}\left(\mathbf{r}\right) \geq J^{-1}\left(\mathbf{r}\right), \tag{5.4}$$

assuming an unbiased estimator, where the Fisher information matrix J is,

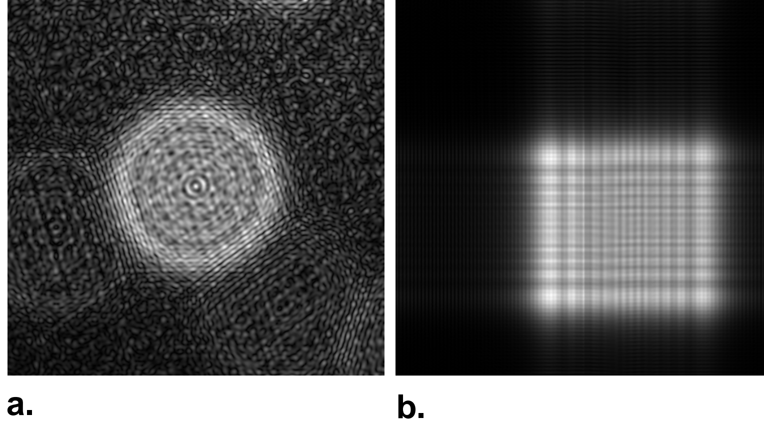


Figure 5.10: a. Calculated images of optical ruler, generated by 780 nm Fraunhoffer diffraction from the quasiperiodic, 9662, 3μm diameter holes distributed over 2 mm, sampled at a distance of 25.4mm. Note high density of features across the whole image, as was confirmed in the real image in Fig. 5.9. b. Same conditions as a., but of a periodic square lattice aperture array.

$$[J(\mathbf{r})]_{ij} = -E\left[\partial^2(\log f)/(\partial \mathbf{r_i} \partial \mathbf{r_j})\right], \tag{5.5}$$

the negative expectation value of the partial derivatives of the log of the likelihood function[51]. Having an optical ruler with dense features across the whole image increases the Fisher information in Equation 5.5 through the partial derivatives and decreases the lower bound of the MSE in Equation 5.4. This is in contrast to an aperiodic structure, which would spread the optical energy across space but without much structure, thereby increasing the MSE. In addition, periodic structures have prominent features only in small regions. In the region of interest studied numerically, it is found that even for the same number of apertures, the quasiperiodic pattern has better MSE estimates. Its square root of the trace of the J^{-1}, as a prediction of the image registration performance, are five times better than that for the periodic pattern.

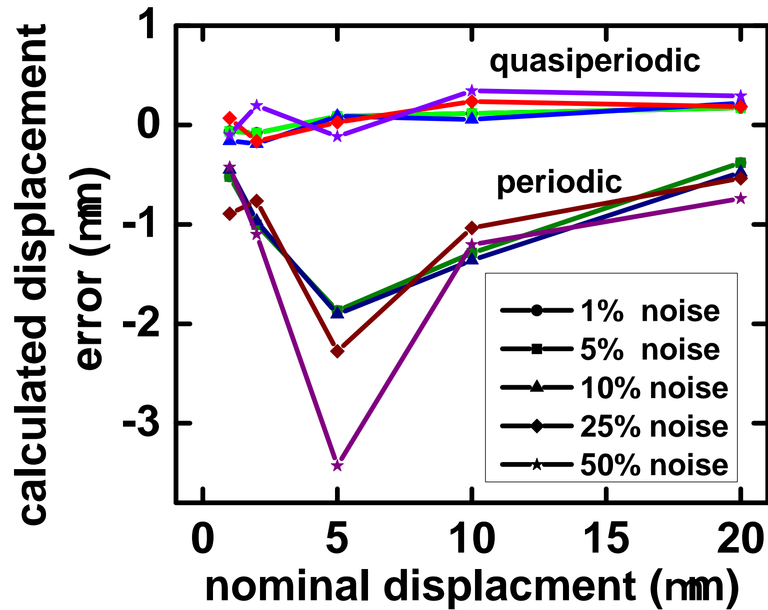


Figure 5.11: NORIS precision compared between a periodic and quasiperiodic aperture array. The periodic array performs up to 31 times worse than the quasiperiodic array. The sampled portion of the periodic array is particularly high in features so that the periodic array precision will in general be much worse than shown here.

A set of Penrose tilings with five-fold rotational symmetry is shown in 5.12, both as a tiling and with the vertices used to generate the aperture array. The rotational symmetry occurs only in one place in the infinite plane, indicated in the graphs. This particular Penrose tiling can be constructed as the intersection of five families of lines or equivalently as the projection into two dimensions of a six dimensional cubic lattice [13].

Figure 5.13 shows a micrograph of the Penrose tiling fabricated in a photomask, then the resulting diffraction pattern with a 632.8 nm laser.

In order to maximize the precision of the quasiperiodic patterns, it is important to have a process flow with very good fabrication tolerances while being optically

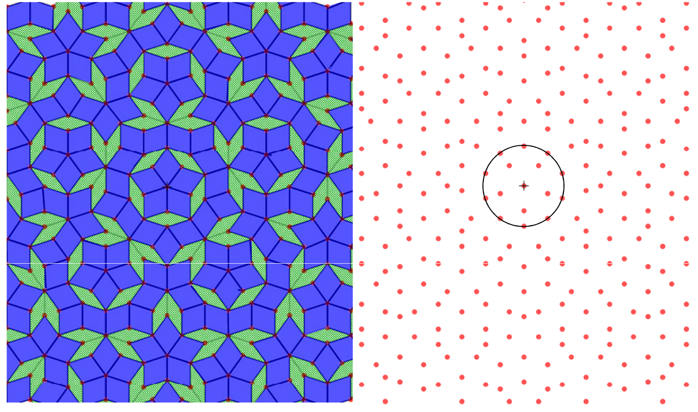


Figure 5.12: Left half shows Penrose tiling using the thin and thick rhombuses (P3). The thin rhombus has angles 36 and 144 degrees, whereas the thick rhombuses have angles of 72 and 108 degrees. Cutting these rhombuses in half to generate two congruent triangles, these triangles have edges in ratios of the Golden ratio. Right side shows the resulting vertices used to design the aperture array. Circle shows center point: this is the only place in the pattern with five-fold rotational symmetry.

transparent up to the diffracting thin film metal. Several process flows have been used. See Figure 5.14. The left half of the figure shows a nirtide suspended version. The substrate is a double side polished wafer. Between one and two μm of low pressure chemical vapor deposition silicon nitride (Si_3N_4 is deposited. Then a thin film of Al (300 - 400 nm) is e-beam evaporated onto the top side of the wafer. The aluminum is patterned by photoresist with the desired quasiperiodic array, then the aluminum is etched by a chlorine reactive ion etch. Then the backside nitride is etched to form an etch window. A seven to eight hour 30% KOH etch at 70° makes a through-wafer etch up to the nitride, and forms a suspended optically accessible thin film aperture array. The right half of the Figure shows a process flow using a silicon-on-insulator wafer. First, 200 nm LPCVD nitride is deposited

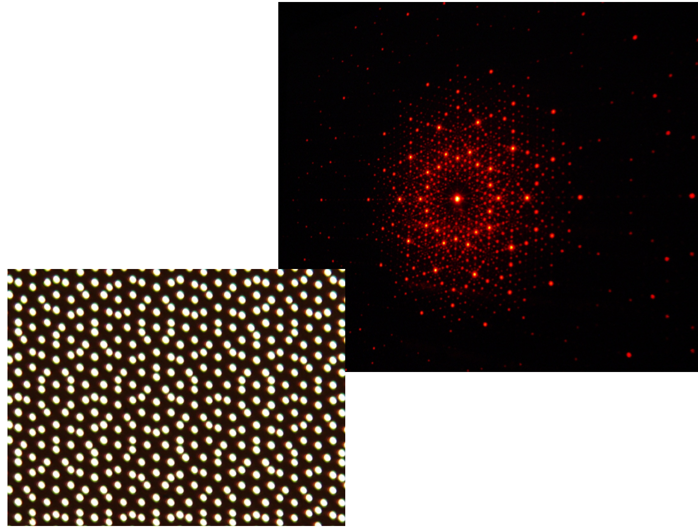


Figure 5.13: Prototypes of optical diffraction rulers can be made by fabricating a photomask with the desired aperture pattern. The micrograph on the left shows a back-illuminated photomask of a Penrose tiling aperture pattern. Image on right shows resulting diffraction pattern.

onto both sides of the wafer, but the top side nitride is removed in a fluorine reactive ion etch. Then a 100 nm film of thermal oxide is grown, which does not appear on the bottom side nitride. The desired aperture pattern is formed using ebeam lithography, then the silicon oxide is patterned by a short fluorine reactive ion etch. A high aspect ratio chlorine silicon etch etches the silicon device layer down to the buried oxide. Then a window is opened into the backside nitride by a reactive ion etch, then a through wafer etch is performed using KOH. The structure is released in buffered oxide etch, then a 200 nm ebeam evaporation of gold on the top side makes a reflective thin film. This method, though more laborious, tends to make more precise patterns because of the ebeam lithography and higher selectivity of the silicon device etch to the oxide mask. A simple metal-on-fused silica with a lift-off resist is also possible, the advantage being that the

pattern can be generated as points of metal rather than apertures, increasing the transmission of the diffraction pattern.

5.3 Image registration

Image registration is used to process the optical ruler images to determine the position of the camera with respect to the optical ruler. Image registration takes a reference and sample image, then by comparing the image intensities intuits the best position offset between the two images. The most basic method is a pixel-by-pixel comparison. Consider two images I_R(x,y) and I_S(x,y), the reference and sample images, respectively. The images extend from x=[1,x_N] and y=[1,y_N]. The cross correlation is,

$$CC(x_0, y_0) = \sum_{x,y}\sum I_R(x,y) \times I_S(x - x_0, y - y_0), \tag{5.6}$$

which yields a x_N by y_N dataset of the cross correlation. The x_0 and y_0 denote the trial offsets. The most likely image offset between the reference and sample image, x_{ML} and y_{ML}, is the value of x_0 and y_0 when CC is maximal. However, the cross correlation must be normalized. For example, a pixel with an extraordinarily high intensity value will result in a term in the sum of Equation 5.6 that is also extraordinarily high. That biases the calculation of the cross correlation, and an incorrect image registration will occur. The normalized cross correlation is then,

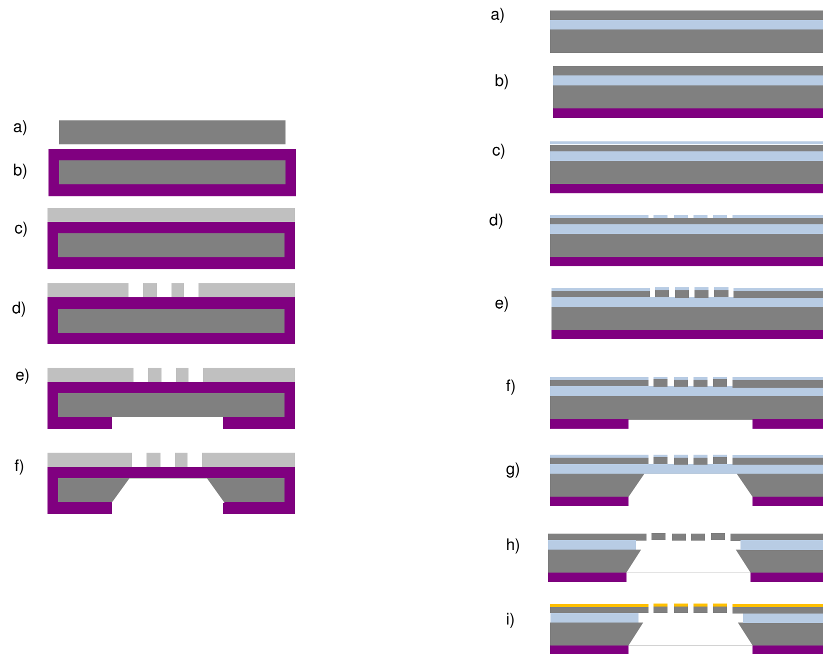


Figure 5.14: Two process fl6ws for making a wafer with an optically accessible diffraction aperture array. Left shows single side polished wafer with LPCVD nitride, thin film of Al etched by RIE for precise patterning, and backside patterning and KOH through-wafer etch to expose the thin film pattern to optical access. Right side shows a silicon-on-insulator wafer. LPCVD nitride is grown, then etched from device side in RIE; subsequent thermal oxide grows only on top device side due to nitride on back side. Due to thin thickness of oxide, precise electron beam lithography can pattern the oxide. Pattern is transferred into device layer Silicon by chlorine chemistry reactive ion etch. Backside nitride is pattern, then through-wafer KOH etched to reach the buried oxide. Then oxide is released by wet buffered oxide etch, then a thin film of Cr/Au is evaporated from the top side to act as the diffractive thin film metal: where holes were etched in the device layer Si, no metal film is evaporated. For large area aperture arrays, a simple process flow of gold thin-film metal evaporated onto a fused silica wafer can be used.

$$CCN(x_0, y_0) = \sum_{x,y}\sum \frac{\{I_R(x,y) - \bar{I_R}\} \times \{I_S(x+x_0, y+y_0) - \bar{I_S}\}}{\sigma_{I_R}\sigma_{I_S}}, \tag{5.7}$$

where the images are normalized by subtracting the mean and dividing by the standard deviation.

Calculating the cross correlation by Equation 5.7 generates a value for x_{ML}, y_{ML} that is a pixel value (e.g., position difference of (5,2) pixels). Achieving nanoscale precision in NORIS requires sub-pixel image registration techniques, much less than the micron sizes of most imager pixels.

Phase cross correlation is based on the Fourier shift theorem,

$$\mathcal{F}[I(x+x_0, y+y_0)] = e^{i(k_x x_0 + k_y y_0)}\mathcal{F}[I(x,y)]. \tag{5.8}$$

That is, an image that is translationally shifted (and assuming periodic boundary conditions) yields a phase difference in the two images; the phase difference is a plane in the x,y plane. Sub-pixel phase cross correlation finds the phase difference $k_x x_0 + k_y y_0$ to deduce the translated shifted x_0 and y_0. This method was tried using the quasiperiodic NORIS optical ruler images, but the results have been mixed. Rather, an upsampling method is used.

Images can be upsampled by padding the image with zeroes in-between the pixel values that exist. After doing so, the pixel-by-pixel cross correlation method seen in 5.7 can be used to determine the most likely offset, x_{ML} and y_{ML}. The single-step DFT algorithm algorithm is used [24], where the cross correlation first finds a small region of maximum probability, then does an upsampled cross correlation of

that region. The cross correlation is easily implemented in software, but hardware implementation using an FPGA, for example, is expected to greatly increase the image processing speed.

In addition to the algorithm used for signal processing, the hardware will also make a difference in the NORIS precision. Imaging cameras can be used with higher pixel count and smaller pixel count to improve the precision, but it is at a cost of increased signal processing times.

Simulations were performed to carry out comparisons of precision as a function of imager characteristics. The quasiperiodic optical ruler image calculated in Figure 5.10 was used. First, Figures 5.15 and 5.16 consider the effect of the pixel size. An imager with N_{pix}=80 and L_{pix} varies from 1μm to 24μm at (Δx,Δy)=(1μm,1μm), (2μm,2μm), (5μm,5μm), and (10μm,10μm). As expected, precision is improved at smaller pixel size. However, the precision undergoes a sort of phase transition where the precision is fairly flat up to 11 μm or 12 μm sized pixels. Then, the precision error jumps high.

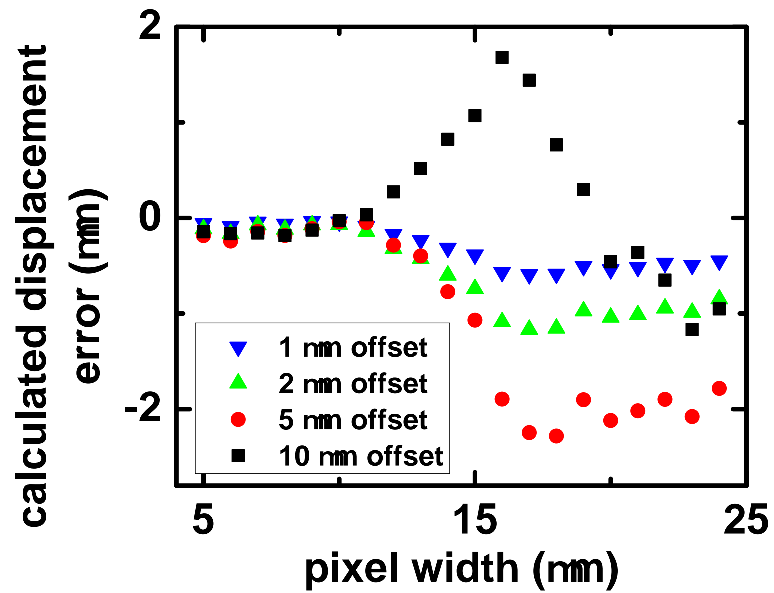


Figure 5.15: NORIS precision as function of pixel width L_{pix}, in x direction. As in 5.17, the precision is flat until a critical value of L_{pix}=11μm where the precision diverges up to 60% of the pixel size.

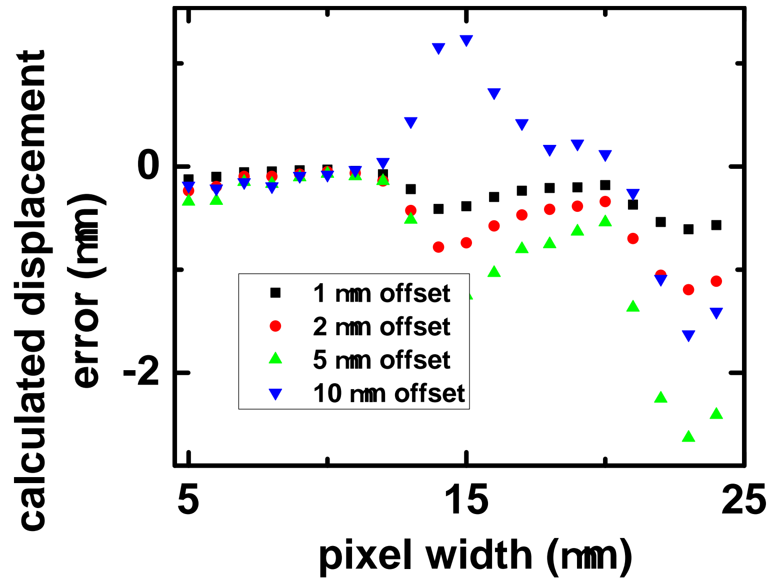


Figure 5.16: NORIS precision as function of pixel width L_{pix}, in y direction. As in 5.17, the precision is flat until a critical value of L_{pix}=11μm where the precision diverges up to 60% of the pixel size.

In the next set of data, the effect of the number of pixels on the NORIS error is seen. By incorporating more pixels, an imager will sample a higher area of the optical ruler gaining an increased sampling of spatial information. Every Fourier component, and its phase, helps the cross correlation deduce the positioning of a sampled image. Again, as expected, the error is reduced at higher numbers of pixels. However, there is again a sort of phase transition where the error does not reduce below a certain number of pixels.

Next, the errors are considered as a function of pixels per side, but by including 5% pixel noise. In the noise-free cases in the Figures 5.17 and 5.18, the transition to good precision occurred around 80 pixels, but adding the 5% of pixel noise to

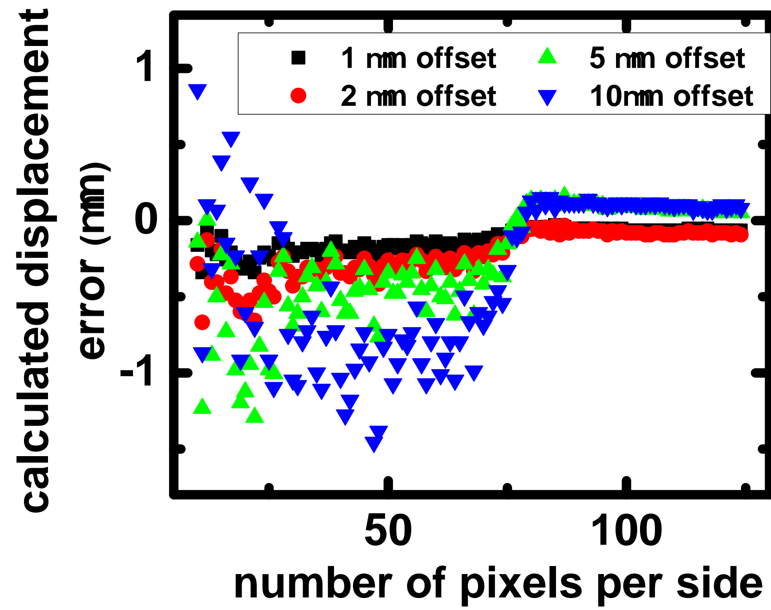


Figure 5.17: NORIS precision as a function of N_{pix}, the number of pixels per side, in x direction. The error can reach as high as one third of the pixel size at low N_{pix}. At a critical N_{pix}, the error precipitously drops and decreases less than 50 nm for $N_{pix} = 80$ to 124. The critical N_{pix} increases as noise is added.

about 86 pixels. In addition, the precision before the transition is much less stable and has higher errors as well.

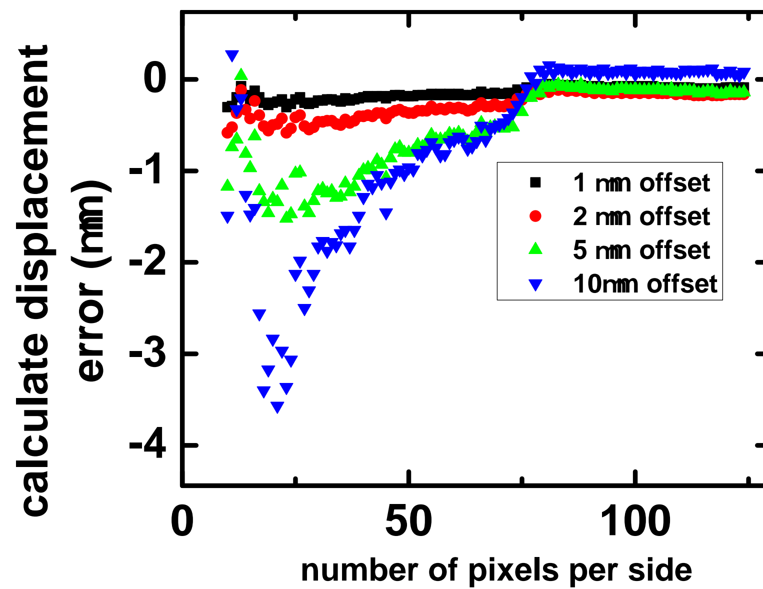


Figure 5.18: NORIS precision as a function of N_{pix}, the number of pixels per side, in y direction. The error can reach as high as one third of the pixel size at low N_{pix}. At a critical N_{pix}, the error precipitously drops and decreases less than 50 nm for $N_{pix} = 80$ to 124. The critical N_{pix} increases as noise is added.

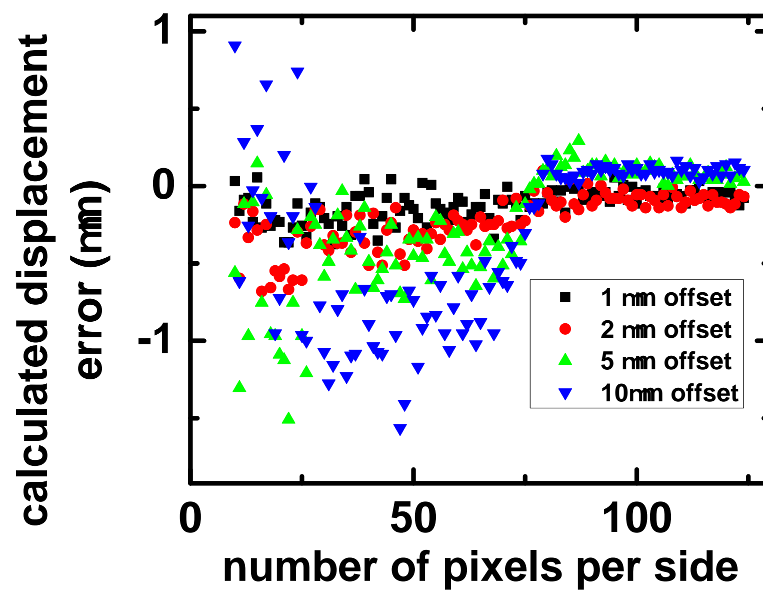


Figure 5.19: NORIS precision as a function of the number of pixels, but with 5% noise added, in x direction. In the noise-free cases in the Figures 5.17 and 5.18, the transition to good precision occurred around 80 pixels, but adding the 5% of pixel noise to about 86 pixels. In addition, the precision before the transition is much less stable and has higher errors as well.

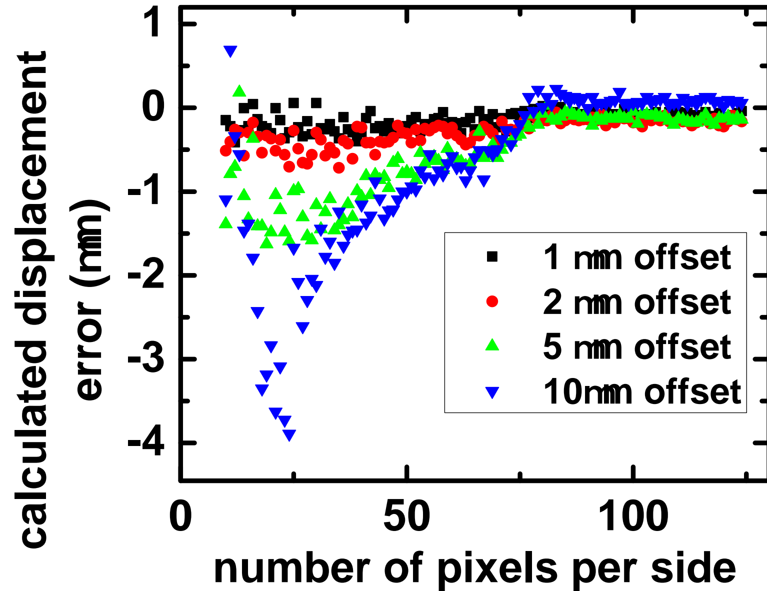


Figure 5.20: NORIS precision as a function of the number of pixels, but with 5% noise added, in x direction. In the noise-free cases in the Figures 5.17 and 5.18, the transition to good precision occurred around 80 pixels, but adding the 5% of pixel noise to about 86 pixels. In addition, the precision before the transition is much less stable and has higher errors as well.

In Figures 5.21 and 5.22, an imager with N_{pix}=100 and L_{pix}=12μm is used and moved the imager by (Δx,Δy)=(1μm,1μm), (2μm,2μm), (5μm,5μm), and (10μm,10μm). The noise N is varied from 0% to 100% of the mean image intensity, and the error in the calculated position in one direction is calculated. There is a $\pm$500nm position error increase (1/24th of pixel size) per ten percentage point increase in pixel noise; with no noise, the system measures the position to be within 100nm (1/120th of pixel size). The mean position error is independent of the position.

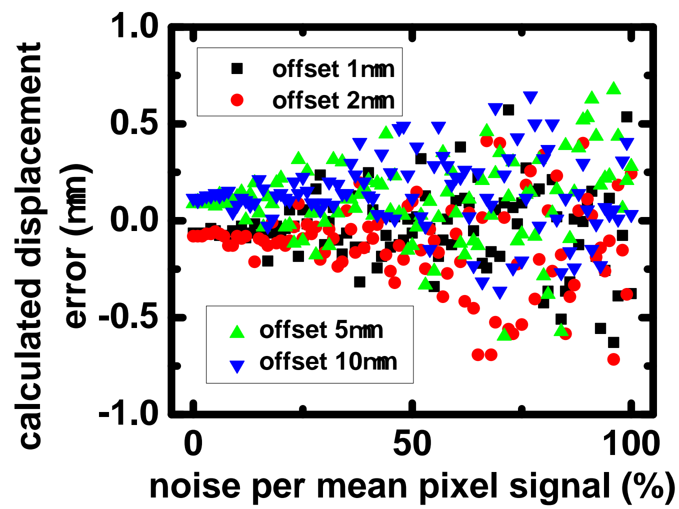


Figure 5.21: NORIS precision as a function of N, the percentage of noise per mean pixel intensity. The error is independent of offset, which is a result of the high density of features in the diffraction pattern of the quasiperiodic aperture array throughout the optical ruler.

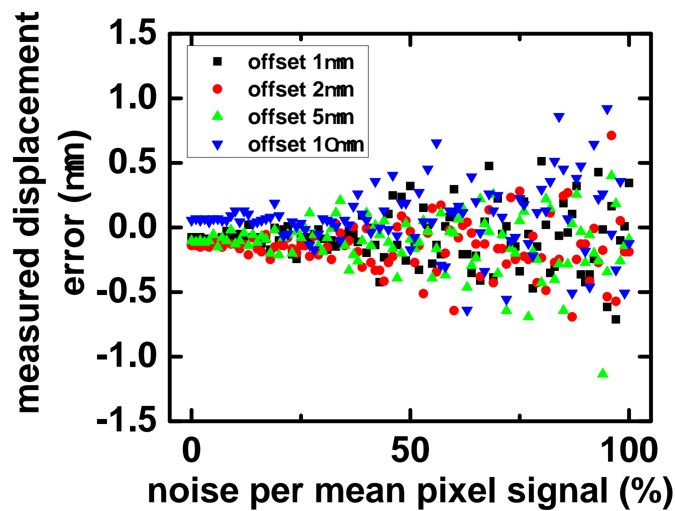


Figure 5.22: NORIS precision as a function of N, the percentage of noise per mean pixel intensity. The error is independent of offset, which is a result of the high density of features in the diffraction pattern of the quasiperiodic aperture array throughout the optical ruler.

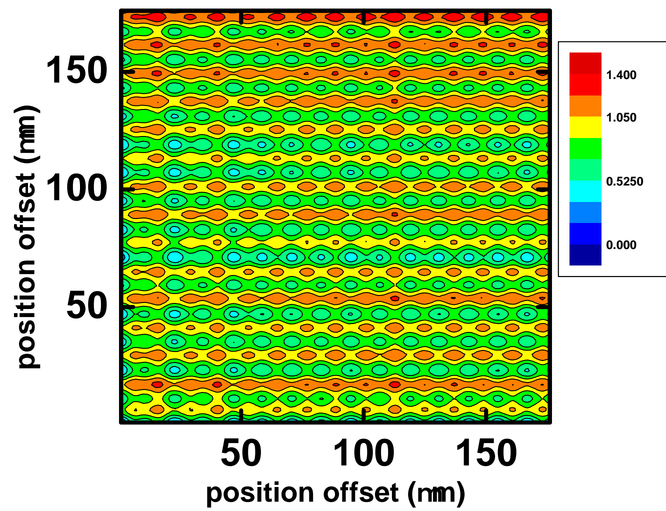


Figure 5.23: NORIS precision as a function of the displacement. The Fourier cross-correlation method assumes periodic images. However, when the imager sees a displaced image, the part of the image that disppears to one side does not appear on the other side. This introduces edge effects, and the result is seen here as a sinusoidal error in positioning.

Finally, in Figure 5.23, the distance error (as opposed to error in x or y) are plotted as a function of distance between the reference and sample images. An imager with 12μm pixels, of a 100 by 100 pixel array, was used. Note the periodic structure of the calculated position.

5.4 Empirical results

An external cavity laser in the Littrow configuration is used as the laser source. A ^{85}Rb heater cell is heated to 45°C, and as in 5.1, is positioned to provide saturation spectroscopy. The laser is locked to the F=2 to 1 transition with a frequency sweep of ¡10 MHz as measured by a Fabry Perot cavity. The laser beam travels some 152 cm until reflecting from three mirrors and to position the laser beam vertically above a wafer manufactured using the SOI process shown in Figure 5.14. The pattern was made to cover a 1 mm diameter area in the center of the wafer. The laser beam is first angled to be perpendicular to the stage surfaces. Then the wafer is placed in line with laser beam and the stage below. As per Babinet's principle, the reflection will generate the quasiperiodic pattern but reflected back towards the laser source. Especially as the reflection is very high power, the feedback to the laser can disrupt the lasing. As a result, it is often prudent to slightly misalign the diffraction wafer. The diffraction pattern generates a very visible center point with a plus-oriented cross (up-down and left-right) along the vertical and horizontal directions of the Penrose vertices, it can be placed diagonally away from axis center and be blocked by an aperture.

Certainly, the intensity of the optical ruler cannot be too late or else the imager

will not be able to detect the signal with high signal-to-noise ratios. However, the gain sweep method described above cannot be used properly if the gain is too high: the pixel intensities will saturated even at slightly larger gains, and a large sweep range will not be possible. Some care must be taken to find a suitable laser intensity to satisfy both conditions. In addition, the optical table holding the NORIS is surrounded by polyvinyl chloride curtains and, above, by fiberglass. These reduce environmental acoustic noise.

As show in Figure 5.1, the diffraction wafer is illuminated by the laser light. The optical ruler is imaged at a distance of 25.4 mm, as simulated above, then positions were calculated using the methods described. The OV7670 CMOS imager was used, with a VGA (640 by 480 pixel array) and an 8-bit ADC converter, with an 8-bit (256 level) gain setting. The chip was mounted on a PC board that contained all the passives required to run the CMOS imager, then mounted on an nPoint XYZ100A which is a flexural piezo stage with capacitive sensors for precise positioning over the stage range of 100 μm in the x-y directions. The piezo stage was then mounted on a Newport ILS150PP stepper stage, with 500 nm resolution.

The first dataset demonstrate the mechanical stability of the NORIS system. Figure 5.24 shows the position of the imager, idle, as calculated by NORIS over an hour. Neither direction drifts more than 38 nm, with standard deviations of 27.1 and 23.8 nm.

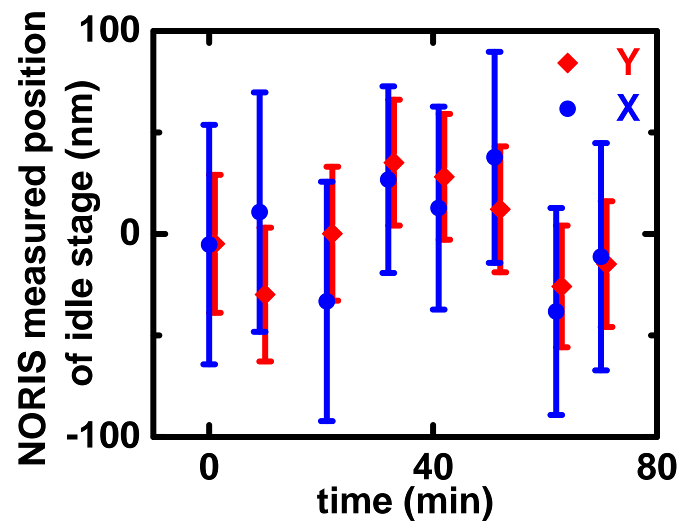


Figure 5.24: Long-term stability measurement of NORIS in x and y directions of an idle stage. The drift is less than 38 nm, with standard deviations of 27.1 nm and 23.8 nm.

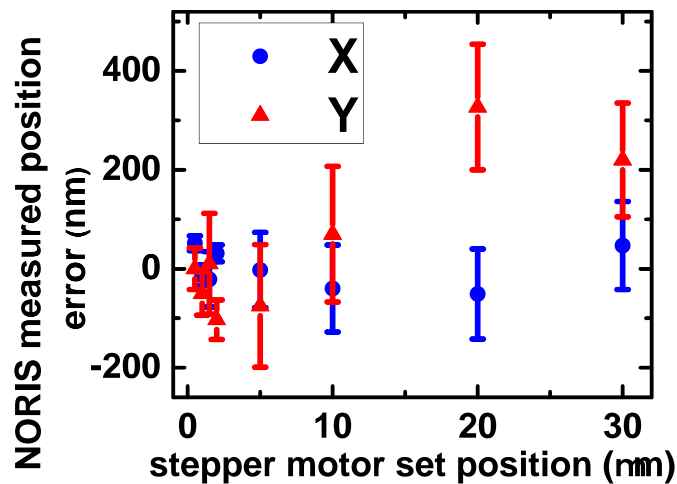


Figure 5.25: Precision of a long travel stepper motor, corrected for angular misalignment. This data also demonstrates lack of half-wavelength ambiguity in position measurement and the sub-pixel positioning which allows for a continuous resolution over long distances.

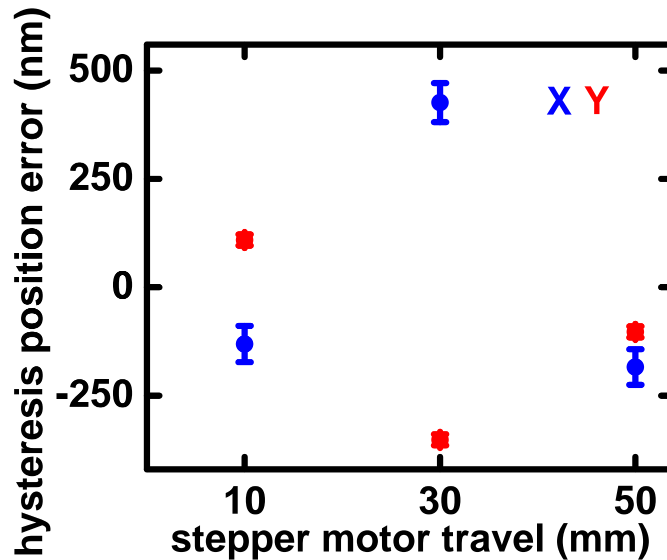


Figure 5.26: Hysteresis of a long travel stepper motor. The position is measured after the stage traverse 10, 30, and 50 mm then returns to its origin.

The dataset in Fig. 5.25 uses NORIS to examine the errors in positioning by a long travel stepper motor. The data show the error as the stepper motor moves 30 μm, correcting for a small angular offset. The motor steps being 500 nm, the stepper motor begins to pick up errors up to 1.6% of its travel distance. The data also demonstrates the lack of half-wavelength ambiguity in NORIS and the path- and actuator-independence of NORIS. Fig. 5.26 shows the hysteresis accumulated by the stepper motor after it traversed tens of millimeters.

Fig. 5.27 shows the nanometer precision of NORIS. The piezoelectric stage was moved in steps of 100 nm, measured by its precise capacitive sensors, and its position calculated by NORIS. The system is accurate to within 50 nm, with a mean error of 29 nm (1/124th of pixel size). The residuals exhibit an oscillatory

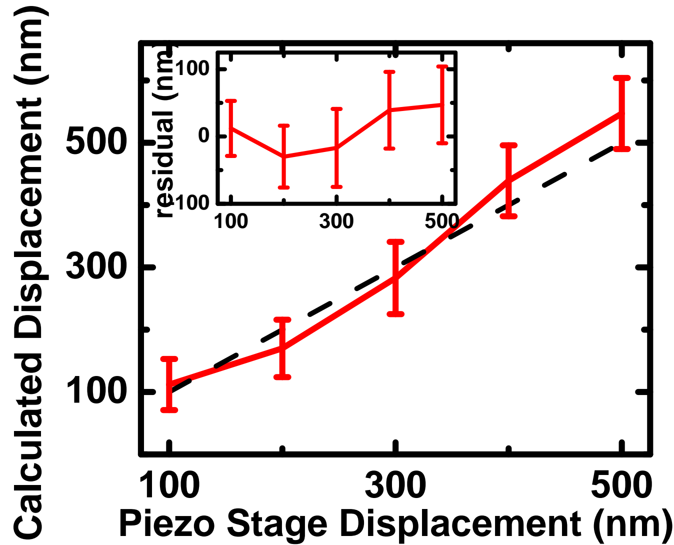


Figure 5.27: Comparison of NORIS measured position to high precision capacitive sensor stage position, showing sub-50nm precision (mean offset 29nm = 1/124th of pixel size), a precision of 5×10^{-7} over four inches. The slight oscillatory behaviour is due to the signal processing.

behaviour, which is a result of the signal processing. Note the oscillatory behaviour of the positioning error; see Figure 5.23 above.

Figure 5.28 shows one method of integrating z-positioning into NORIS. The cross correlation method does not suffice for z metrology because the shift in pixel that is measured for x-y is not high enough to detect changes in z as well. Here, a beam splitter was simply positioned on the PC board holding the imager, such that the incoming frequency stabilized laser formed a standing wave between a thin film of gold on the PC board, and the diffractive wafer. A resonance is set up in the resulting Fabry Perot cavity, the Figure showing the detected intensity as a function of the change in the distance between the imaging camera and the diffraction wafer. As in the interferometric systems described above, this method suffers from translation symmetry and thus a half wavelength ambiguity. In many

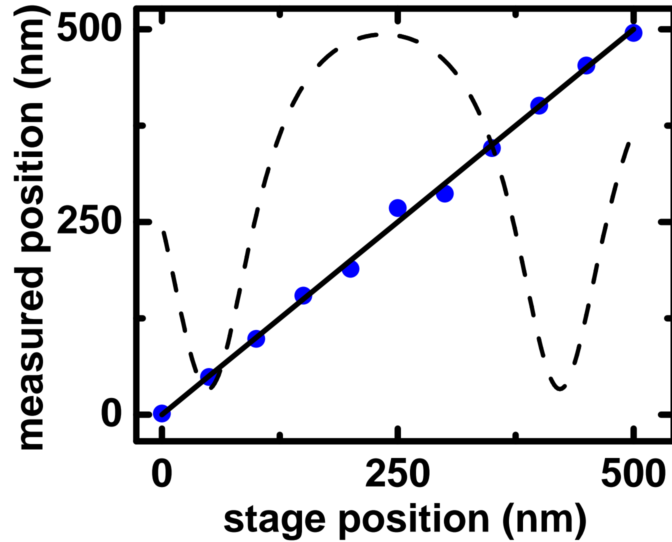


Figure 5.28: Comparison of NORIS measured position to high precision capacitive sensor stage position, showing sub-50nm precision (mean offset 29nm = 1/124th of pixel size), a precision of 5×10^{-7} over four inches. The slight oscillatory behaviour is due to the signal processing.

applications, in particular for scanning probes for which NORIS is best applied, the z positioning is not a variable. For example, a sample will be fixed to a known position and multiple z positions are not necessary for operation. In cases where precise z positioning is required between the imaging surface and the diffracting wafer, it is proposed that a Pound-Drever-Hall technique would suffice, whereby the central lobe of the diffraction pattern would form a standing wave between the diffracting wafer and a manufacturing wafer / imaging surface.

CHAPTER 6

PRECISION PATTERNING OF GRAPHITE

Parts of this section have been submitted to Nanotechnology.

Graphene is receiving much attention for its extraordinary electrical and mechanical properties, c.f. [21]. Graphene discovered as recently as 2004, when it was isolated by mechanical exfoliation [46]. The carbon atoms in graphene are sp^2 hybrid orbitals bonded to three neighboring carbon atoms, in addition to a weakly interacting π bonds. The weakly interacting π bonds result in some of its interesting characteristics, most importantly that single sheets of graphene can be exfoliated from thick layers of graphite. One of the anticipated uses for graphene is as a field-effect transistor. Graphene, however, lacks a band-gap until it is confined by constraining its width. The resulting graphene nanoribbons have been predicted to have single-atomic dependent properties with a periodicity of three atoms: that is, for graphene nanoribbons fabricated with random widths two of every three would have a band gap, while one in three would be metallic [61]. There is a need for graphene nanofabrication methods that are beyond e-beam or photo-lithographic methods, which rely in resist molecules that are much larger than a single carbon atom in the graphene lattice.

As a first example, Figure 6.1 shows etching of few layer graphene by a custom made STM designed in combination with NORIS. A 3V bias was applied between the tip and the sample, resulting in a 6000 Åetch; a subsequent drift in the image

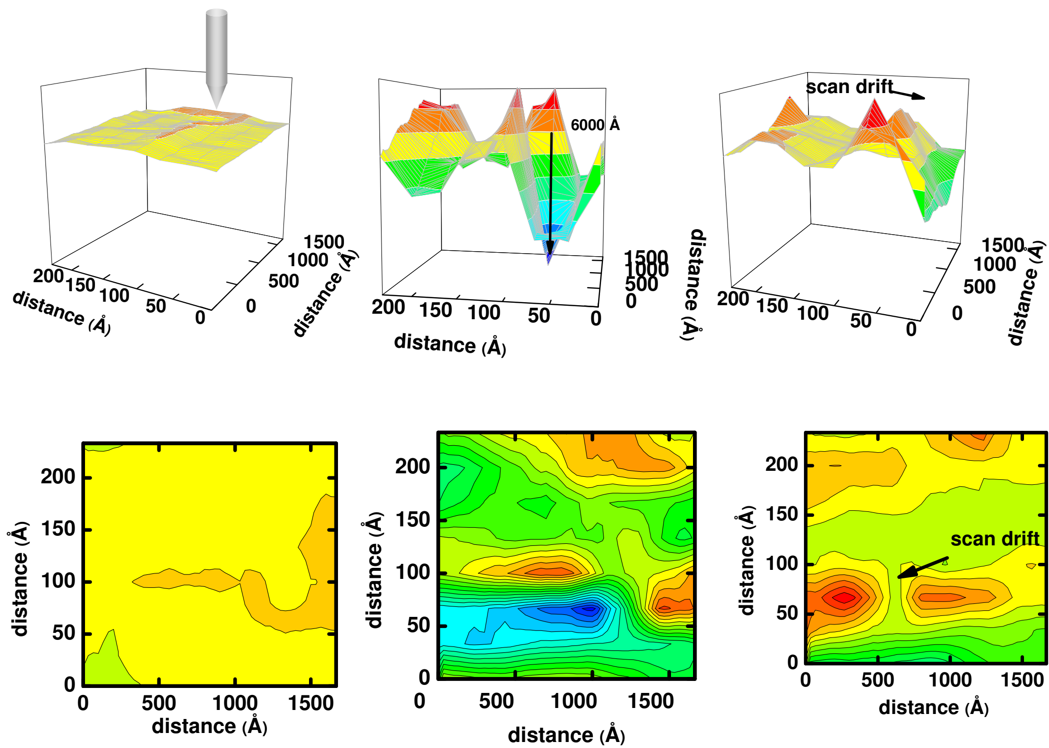


Figure 6.1: In custom scanning tunneling microscope fabricated with NORIS optics, a few layer graphene sheet was etched with a small hole by applying a voltage to the STM tip. After etching, the STM was scanned to view the hole; a subsequent image shows drift in the image.

scanning is also apparent.

Though single domains of graphene can be mechanically robust, it can be chemically etched by an electrochemical etch.

6.1 Graphite etching process

The process of graphite etching is due to an aqueous electrochemical oxidation and removal of the surface carbon atoms. The water is supplied by the meniscus that forms between the carbon surface and the scanning probe tip from the ambient moisture. As a result, HOPG etching does not occur under vacuum. This is best described by the Pourbaix diagram for carbon, which shows a narrow region of stability at low potentials [50]. The main graphite etching reaction is the generation of carbon dioxide [45, 44, 42],

$$C + 2H_2O \rightleftharpoons 4H^+ + 4e^- + CO_2, \tag{6.1}$$

which is essentially an irreversible reaction whose rate is increased at positive carbon, or cathode, voltages. At high bias voltages, the carbon is etched rapidly. At low bias voltages the carbon can be oxidized rather than etched: while the Pourbaix diagram indicates the possibility of etching at low voltages, the chemical kinetics favor the oxidation [35]. Graphene has been etched at voltages as small as +2 V sample bias. At the reverse polarities, the Pourbaix diagram shows the possibility of the generation of methane by combining carbon with four hydrogen ions and four free electrons, but this is a highly unlikely reaction. This results in the HOPG cutting only at a positive sample bias, as seen in Figure 6.2.

It is found that generally, the hydrogen ions and free electrons in Reaction 6.1 combine quickly during the etching process and disperse as hydrogen gas. First, note that the electrochemical etching at the scanning probe involves etching carbon with a small volume of solvent as the meniscus. Given the amount of carbon that is typically etched and the resultant amount of ionic hydrogen, it would be possible for the molal ionic strength to range up to tens of thousands. Even for a much

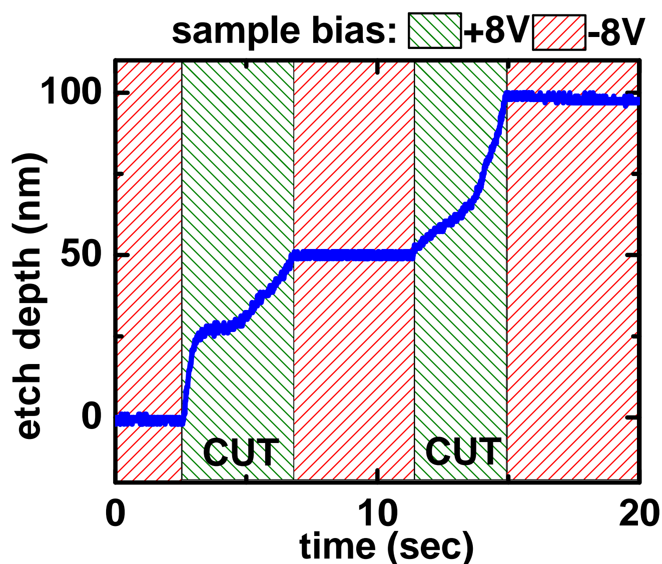


Figure 6.2: When the sample bias voltage is positive with respect to the scanning probe tip, the essentially irreversible reaction in Equation 6.1 etches the graphite. The opposite polarity results in no etching of the graphite as the generation of methane requires the carbon to react with four hydrogen ions and four free electrons. The polarity-dependent etching of HOPG allows for electrochemical cleaning of the probe and possibly graphite annealing.

smaller range, the activity coefficient for Reaction 6.1 would change by an order of magnitude (cf. [48]). However, the etch profiles in Figure 6.3 shows that the etch rate is linear, indicating that the reaction rate and therefore the activity coefficient is constant. Second, it is possible to etch graphite without current, as seen in Figure 6.3. Free charge is only seen at higher cutting voltages, where some of the hydrogen ions and the free charge can be pulled to the tip and substrate, respectively.

Operating the atomic force microscope (AFM) in the tapping mode, Figure 6.4 shows fine structure of HOPG etching that has previously been unobserved. While contact mode images tend to show very smooth pits, the tapping mode

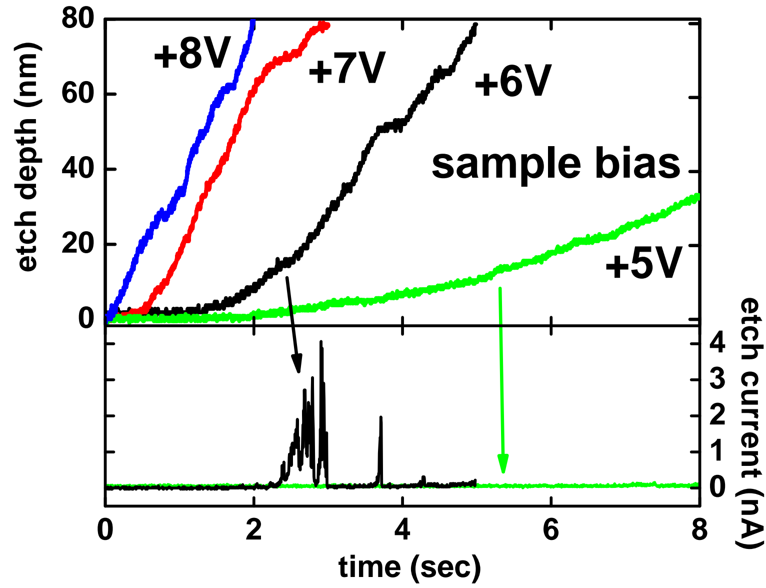


Figure 6.3: Etch profiles at several sample bias voltages. Note the constant etch rates, especially at +7 V and +8 V sample bias, which imply a constant activity coefficient. The current is also measured as the tip etches the carbon. At a bias of +5 V, the hydrogen ion and free electron currents are not measured by the scanning probe tip. However at a bias of +6 V, the tip and substrate capture some of the current.

images typically show crevice structure inside the pit. Figure 6.4 (b) also shows a tapping mode image of a smooth line cut by moving the tip along the surface. In addition, by comparing tapping mode images to contact mode images it is found that water is present after features are cut into the HOPG. This is confirmed as the water can be removed by running the tip across the water under contact or by pumping down the chamber to vacuum then reimaging the HOPG; see Figure 6.4 (a). The presence of the water is surprising, as HOPG is hydrophobic. This may be a combined effect from surface modifications to the carbon and a change in pH of the water due to the etching process. The etching process in addition to removing carbon can also oxidize the carbon, which will decrease its hydrophibicity [16]. Graphite's wettability can also be raised by increasing the acidity [23] caused

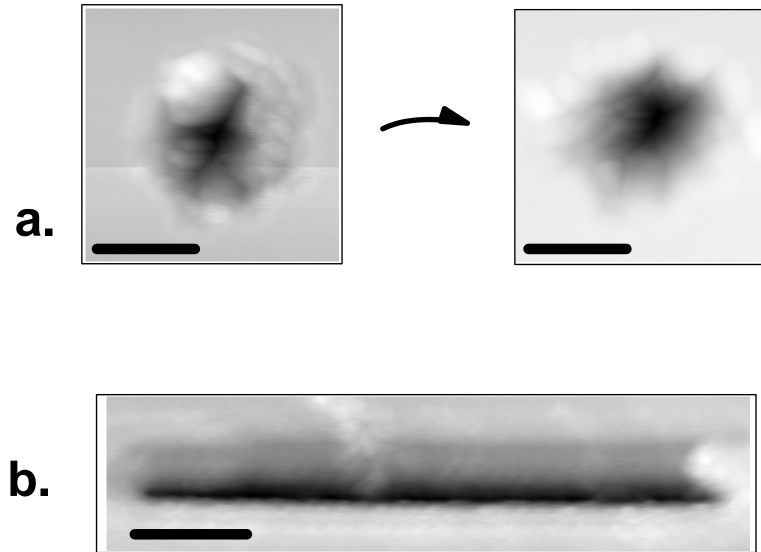
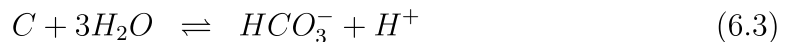


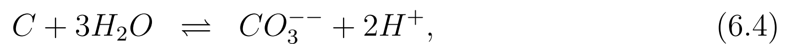
Figure 6.4: (a) Tapping mode image of large hole (scale bars show 50 nm) showing some crevice structure. After imaging, three contact mode scans were made of the hole. Subsequent tapping mode image shows removal of water. (b) Tapping mode image of a smooth line (scale bar shows 100 nm).

by the formation of carbonic acid and carbonates,

$$C + 3H_2O \rightleftharpoons H_2CO_3 \tag{6.2}$$



$$C + 3H_2O \rightleftharpoons HCO_3^- + H^+ \tag{6.3}$$



$$C + 3H_2O \rightleftharpoons CO_3^{--} + 2H^+, \tag{6.4}$$

as evident in the Pourbaix diagram for carbon.

While it is expected that the reaction in Equation 6.1 to cleanly etch the carbon without residuals, energy dispersive x-ray spectroscopy with an electron beam microprobe shows the presence of carbon on the AFM tip surface. After several hundred hole etches, the deposited carbon is clearly visible in an SEM image of the tip. The increased tip diameter results in poor imaging and lithographic resolution. It is found that this can be simply remedied by applying a large negative sample

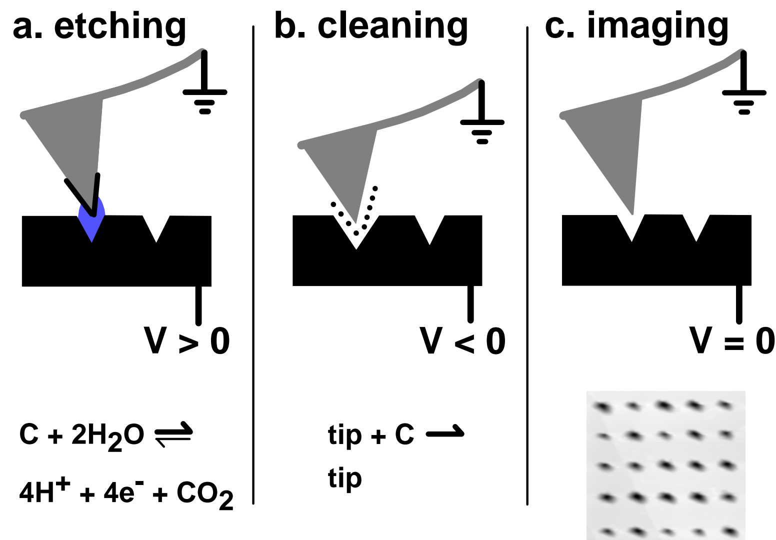


Figure 6.5: Schematic summary of modes of operation for AFM graphite nanolithography: (a) Etching occurs at positive sample bias voltages with a meniscus formed between the tip and the sample via an electrochemical etch chemistry. While etching, carbon deposits form on the AFM tip. (b) Tip cleaning of carbon deposits occurs at negative sample bias voltages. (c) Image scans are at zero bias voltage.

bias, opposite in polarity from the voltages used in HOPG etching. As demonstrated above, large reverse biases can be applied without etching the HOPG. This process was used by Spinney, et al. [63] as a method of carbon deposition onto a gold surface. It is found, however, that AFM images of the HOPG surface before and after cleaning the tip are identical. Intermittently, the AFM probe is cleaned in this manner to maintain a clean, sharp tip over several hundreds of writes. Additionally the negative sample bias allows non-destructive current injection from the tip into the graphite, which will anneal the graphite edges by Joule heating [31]. The etching, cleaning, and scanning modes of operation are summarized in Figure 6.5.

The lithographic feature size is determined by the size of the meniscus that

forms between the scanning probe tip and the HOPG surface. The meniscus size is mainly determined by the environment, probe parameters, and tip height [30]. The temperature dependence can be exploited to control the feature size. As the temperature of the HOPG surface increases, the evaporation rate at the edges of the meniscus increases, resulting in a narrower meniscus and smaller feature size. The data is shown in Figure 6.6, where a 35 ^{o}C increase in HOPG temperature results in a 24 % decrease in the full width half maximum of 24 nm deep holes. The Arrhenius equation predicts an increase in the etching rate as the surface temperature is elevated. Instead, the etching rate slows at higher temperature due to increased instability of the meniscus. The thermally-reduced meniscus yields smaller feature sizes but increased etch times. It is estimated that the average etch rate decreases from a maximum 2×10^{6} atoms/second to a minimum 2×10^{5} atoms/second.

6.2 Precision writing of a line array

Scanning probe etching of HOPG is driven by the tip-substrate meniscus, and as a result is sensitive to the tip shape and surface, tip-sample separation, environment, applied bias, etch time, and temperature. Slight variations among tips might require the applied bias to vary from six volts up to fifteen volts or more before a given tip will etch HOPG. Even the same tip will etch inconsistently over time due to changes in a variety of these parameters. In order to nanofabricate many features consistently, it is necessary to incorporate a feedback control system. A digital control system was implemented that monitored the AFM piezo height position and controlled the applied bias voltage. Conductive AFM tips were used

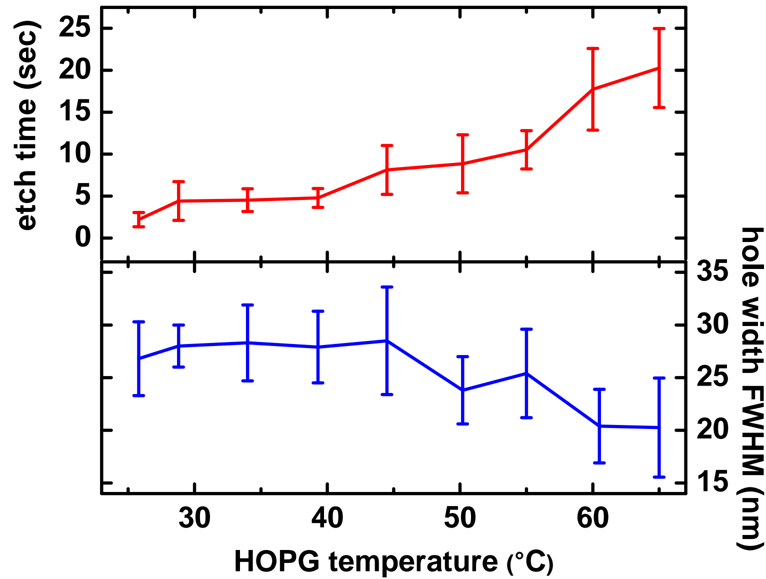


Figure 6.6: Temperature dependence of HOPG etching with V_H=10.5 V, V_L=9 V. Top shows decrease in cutting rate as the temperature of the HOPG is raised. A decrease in cutting rate can be offset by increasing the sample voltage. Bottom shows improved lithography resolution as temperature of HOPG is raised. Increased temperature leads to increased volatility due to evaporation at the edges of the meniscus, resulting in a smaller meniscus resulting in a smaller etch feature size but also much slower etch rates.

with a platinum thin film: platinum has been found to catalyze the electrochemical corrosion of carbon [74, 52]. The control is shown in the inset of Figure 6.7, where a voltage bias is applied to the substrate based on the cut depth. The cutting begins at a larger voltage V_H, typically more than +10 V. Higher biases consistently initiates the carbon etching, and etch at a more rapid rate. However, fast etches lasting less than a second result in inconsistent dimensions and more tip damage, possibly due to more mechanical damage and deposited carbon. Therefore, the bias voltage is decreased to a smaller voltage V_L, typically one or two volts smaller than V_H, when the tip has reached 80 to 90 percent of the total desired etch depth. Etching at V_L results in slower, smoother cuts with well controlled etch

dimensions. Etching at too small of a voltage impractically increases fabrication times, and the increased time allows longer lateral drifts in the tip position resulting in degraded feature quality. In addition, the graphite may undergo oxidation rather than etching, as described above.

Using this feedback control loop, several examples of etches are shown in Figure 6.7. Etches (a), (b), and (c) all start with a three second tip cleaning by applying a -10 V sample bias voltage. Etch (a) is an optimal etching profile. The etch starts immediately after the tip cleaning, and etches quickly before slowing down just before the desired etch height is reached. Etch (b) shows some variation from an optimal etch profile. The etching begins very slowly (7 Å/sec) for 2.8 seconds until the etch rate suddenly increases to a substantial rate. The etch also shows two pauses, indicated by arrows. An etch will sometimes slow or even stop for several seconds, then suddenly resume etching into the carbon surface. This slow decrease and abrupt jump in etching rate may be a result of the loss and reformation of the tip-sample meniscus. Etch (c) shows a large deviation from the etch profiles of (a) and (b). The initial delay is slightly less than that of (b), but the overall etch rate is lower. Etching ceases several times, once as long as five seconds, and in the end its total etch time is twice as long as that for (a).

An array of trenches was fabricated to demonstrate the scalability of our precision HOPG-patterning technique. These lines were cut by measuring the HOPG height at the endpoints, and linearly interpolating a desired cut depth along the line. The lines are three or four pixel widths, where a pixel is about 40 nm wide. The line length was chosen at random by the computer immediately before each one was fabricated. The values for V_H and V_L were +11.5 V and +11 V, but

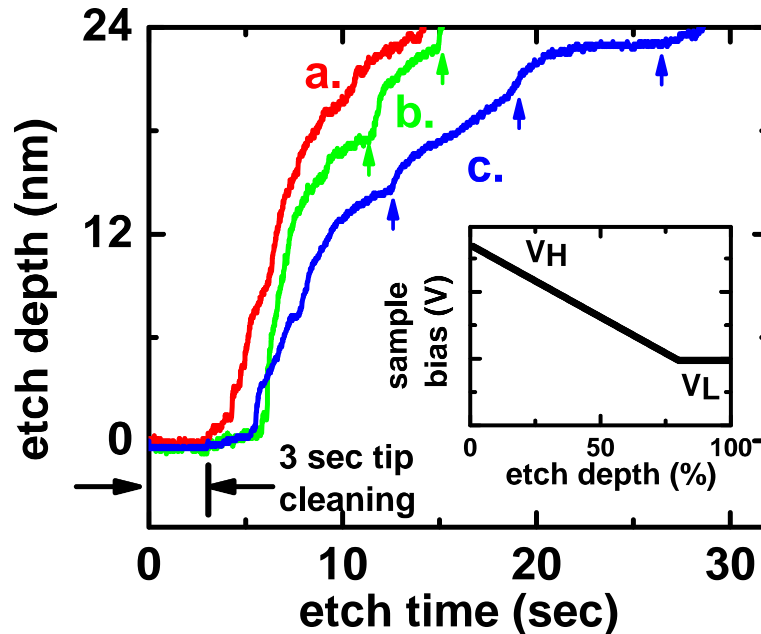


Figure 6.7: Etch profiles using etch control depicted in inset with $V_H = +11.5$ V and $V_L = +11$ V; a three second tip cleaning at -10 V sample bias precedes all the etches. (a) An optimal etching profile, where the etching depth and deceleration are smooth. (b) The etch begins very slowly for 2.8 seconds. Arrows indicate pauses during etching, possibly due to disappearance and reappearance of the meniscus. (c) Very slow etch with several etching pauses. Even with the same parameters as (a) and (b), the etch time is twice as the nominal etch in (a). Inset shows control feedback for precision etching of graphite. The applied bias starts at a large V_H, rapidly initiating the cut and etching most of the desired depth. The applied bias decreases linearly with etching depth, until at some percentage of the desired etch depth the applied bias is set to a low V_L, typically a volt below V_H. The V_L yields a well-controlled etch depth and width by slowing down the etch rate.

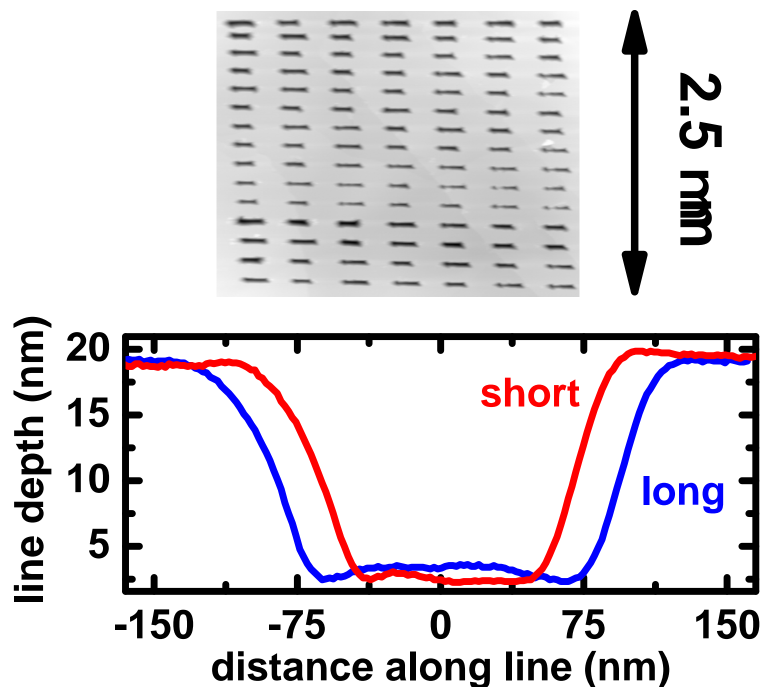


Figure 6.8: AFM image shows large array of 18.3±3 nm deep etched lines in HOPG. Lines are randomly selected to be 136 nm and 183 nm, yielding 6 nm (4.4 %) and 5 nm (2.7 %) precision, respectively. Figure below shows sample cross sections of the short and long lines.

offsets were added to these voltages during fabrication to keep the etch time for each line at around 40 seconds. A -10 V sample bias was applied for three seconds to clean the tip before etching each line.

In order to assess the precision of the fabricated lines, the depth and full-width at half-maximum depths were measured using AFM scans. The result from a seven by fifteen array of lines is 136±6 nm and 183±5 nm, with etch depths that are 18.3±3 nm. Figure 6.8 shows typical profiles of a long and a short line sampled from this array. There is a consistent asymmetry in the edge slopes due to an imaging artifact.

6.3 Conclusion

Feedback-controlled nanofabrication of a 105 line array in HOPG has been demonstrated, using some 370 etching operations on an atomic force microscope. The lines were randomly selected to be of two lengths which were precise to within 4.4 % and 2.7 %, or 6 nm and 5 nm, respectively; the depths were measured to be 18.3±3 nm. The precision of nanofabrication was maintained by a bias voltage-etch depth feedback control loop of the etching rate. In addition, etching results in carbon build-up on the tip which reduces fabrication and imaging precision. A negative sample bias was applied to remove carbon residuals from the tip without sample damage. This work demonstrates that non-prototypical scanning probe nanofabrication for carbon-based devices is possible.

APPENDIX A

CONTROL LOOPS

This appendix discusses control loops. Control loops appear several times, both explicitly and implicitly, in this thesis. Atomic locking requires a control loop to fix the frequency laser to the atomic transition. A control loop is required for thermal control in NORIS. A digital control loop is required for the precision scanning probe nanofabrication of HOPG. While there has been much work on control theory, experimental implementation of complicated theory is unnecessary and probably not very effective. These systems are best controlled by sensitive observation of the system dynamics and intuitive changes to the control loop.

A.1 Control loop stability

The data on the linewidth narrowing by the feedback of the VCSEL 2.11 brings to light the need for control loops in atomically stabilized systems. In that case, the linewidth narrowing naturally arose from the setup. As long as some of the laser spectrum matched the spectral hole, the hole would filter the laser wavelength and thereby narrow its linewidth. The center wavelength was not controlled, but was a natural constant. Therefore the frequency of the laser did not have to be strictly controlled to match the spectral hole. It is easy to see that most systems will require some sort of control loop since rather than directly using atomic properties as in this example, an external system will be controlled by probing an atomic sample. Atomic clocks, for example, lock their local oscillators to the physics package by measuring the error signal from a phase sensitive detector,

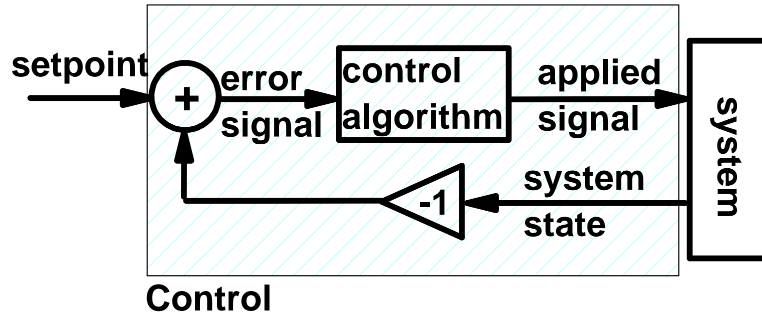


Figure A.1: Schematic of feedback controlled system. The user inputs a desired setpoint. At the summing junction, the system state is subtracted from the desired setpoint. The error signal is fed into a control algorithm, which calculates a response to the system to correct for the error. The signal is applied to the system to try to move its state to the desired setpoint.

then applying the error signal to the voltage controlled local oscillator. This way, the local oscillator stays locked with the atomically defined RF frequency. Figure A.1 is a schematic showing how a control system would run a system, just as an atomic physics package controls a local oscillator or interferometry is used for stage metrology. An error signal is generated as the difference between a desired setpoint (i.e. output frequency), then a control algorithm uses the error signal to apply a signal to the system which changes the system (i.e. frequency shift in oscillator transduces change in atomic resonance), until the system state is found be the same as the desired setpoint.

Control theory provides the basis for efficiently and accurately controlling our systems[65]. Figure A.2 shows a schematic of a control loop. The transfer function L(s) is the frequency response of the control loop which will control a system. First, the stability of the control is determined using the Nyquist criterion. This is done by studying the Nyquist plot. First, the simply connected closed contour $\Gamma \in \mathbb{C}$

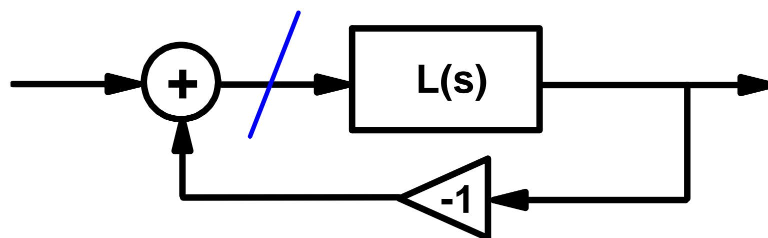


Figure A.2: Schematic of a loop. The loop transfer function is denoted L(s), and the negative feedback is summed back into the loop. The blue line indicates the point of the loop where it is theoretically broken and the stability can be calculated. As described in the text, the poles can be calculated to deduce the loop stability.

is defined to be clockwise, extending infinitely from the negative purely imaginary numbers up to the positive purely imaginary numbers, then forming a half-circle that rotates clockwise from the positive back to the negative imaginary numbers. The Nyquist plot consists of the transfer function L(s) along this closed contour Γ. Inevitably, the transfer function L(s) goes to zero gain at infinite frequencies and L(s) is studied where the values of s go from -inf to +inf. Recall that the stability of a transfer function depends on its poles. Poles with positive, real parts implies a positive, real part in its eigenvalue of the constituent differential equation. This in turn implies that the response of the system will be to sustain or grow the mode at the pole. Therefore, transfer functions with poles in the positive half of $\mathbb{C}$ will generate unstable feedback loops, as each cycle of a signal through the loop will increase its strength. The Nyquist criterion can easily find the number of poles and zeroes of a loop transfer function by examining the Nyquist loop. In Figure A.2 note that the gain of the loop G=T(s)/(1+T(s)). Instability points are found at the poles where 1+T(s)=0. The Nyquist criterion is based on Cauchy's argument

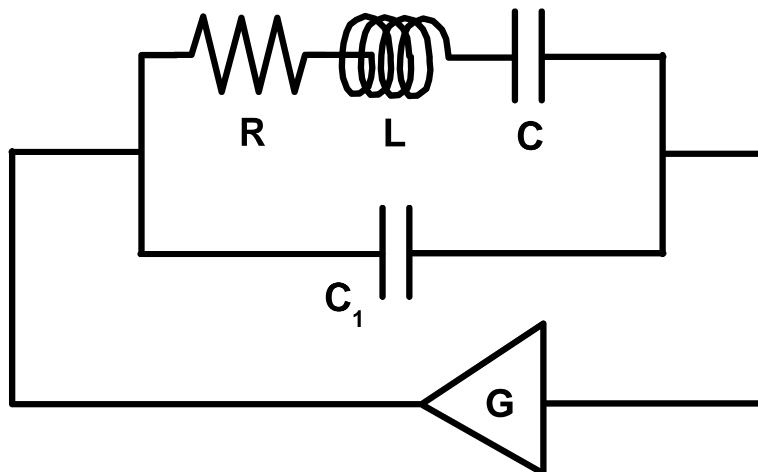


Figure A.3: A control loop for a MEMS oscillator. The MEMS oscillator is modeled by a serial RLC tank and a feedthrough capacitance.

principle, where the difference between the number of poles and zeroes with the region of a simply connected contour is reflected in the line integral along that contour (cf. Cauchy integral theorem). The result is the Nyquist criterion:

Nyquist Criterion. *The loop transfer function T(s), when s traverses the contour* Γ*, will enclose the point -1, N times where N = Z - P, where Z is the number of poles of the closed loop system in the positive real half of* $\mathbb{C}$ *(indicating closed loop stability), and P is the number of poles enclosed by the loop transfer function T(s) in the positive real half of* $\mathbb{C}$*. A loop is stable iff Z=0.*

As an example, a simple model for a MEMS resonator is considered with a negative feedback loop. The mechanical resonator is modeled as a capacitor C1 in parallel to a series RLC, and a feedback gain G is added. See Figure A.3. The open loop transfer function T(s) is,

$$T(s) = \frac{Ls^3 + Rs^2 + (C_1 + C)s}{C_1 Ls^2 + C_1 Rs + C_1 C}. \tag{A.1}$$

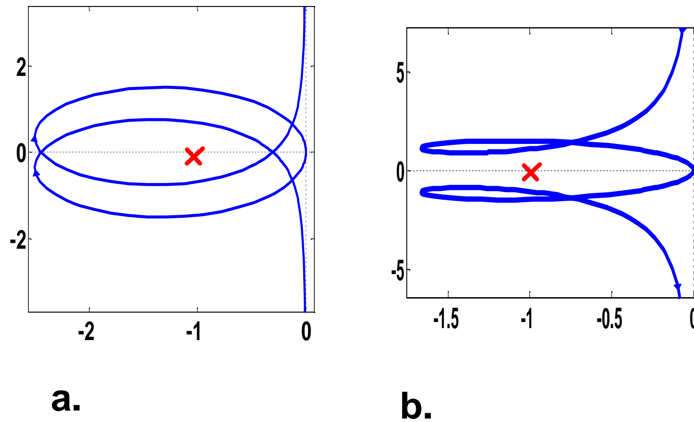


Figure A.4: Nyquist plots. a. Nyquist plot for MEMS oscillator in Fig. A.3, with G=5, R=2, L=4, C=2, C_1=10. The Nyquist loop circles -1+0i twice, i.e. N=2. Z=N+P=2, so that the loop is unstable. b. Nyquist plot for G=5, R=3, L=1, C=5, C_1=10. The Nyquist loop no longer circles -1+0i, so that N=0 and therefore has no unstable poles. The transition from instability to stability (no oscillation) is a reflection of the decrease in the $Q=1/R\sqrt{L/C}$ of the resonator.

Two sets of parameters are considered. The first set has the values G=5, R=2, L=4, C=2, C_1=10. The Nyquist loop is plotted in Fig. A.4(a). The Nyquist loop circles the point -1 twice clockwise, i.e. N=2. The open loop transfer function has no poles in the real half of the complex plane so the closed loop will have Z=N+P=2 poles and therefore be unstable. In contrast, consider a second system with values G=5, R=3, L=1, C=5, C_1=10. In Fig. A.4 (b) it is seen that the Nyquist loop no longer circle the point -1, so that N=0. The open loop transfer function still has no unstable poles, so P remains 0 so that Z=0. This loop is stable and will not uncontrollably oscillate. The transition from instability to stability is due to the increase in R and C, and decrease in L, which effectively lowered the $Q=1/R\sqrt{L/C}$ of the tank of the MEMS resonator.

A.1.1 PID control

A common control algorithm is the proportional-integral-derivative (PID) controller. In such a control system, the feedback signal is a sum of the P, I, and D of the error signal. It can be said that the three terms correct based on past, current, and future behaviour of the system. This is reflected in their spectral response, where the I and D favor the lower and higher frequencies, respectively. The PID control is written as $K_i/s+K_p+K_d s$. Control loops can also use parts of a PID. Here, PID control loops are used later in stabilizing a laser frequency to an atomic saturated absorption line, but a P controller is used to control AFM lithography and a PI controller has been used to control a homemade STM. Typically, the D is used at very low gains. The D feedback can reduce oscillations, but can quickly become destabilizing as its frequency response $K_d s$ has a theoretically infinite high frequency gain, which practically can cause great oscillations in the system response. To prevent this, extra poles can be added to the D circuit to attenuate the high frequencies: in other words, to transform it into a band pass rather than a high pass frequency response. The I is especially important to reduce settling offsets. At s=0, the I term has infinite gain, so that the closed loop gain is 1, so no offset error should exist. Setting the PID frequency responses and gains depends on the system parameters, which are often unavailable, immeasurable, and/or dynamic. Therefore, setting PID gains is often a matter of trial and error, especially as the systems here are not meant to combat large changes in the system but small, Gaussian noise type disturbances. This is often done by setting a medium P setting, then slowly increasing the I before the system begins to oscillate. D can be added as well. However, a common method for estimating the ideal parameters have been developed by Ziegler and Nichols. Their step response

method measures the step function response and aims to find the optimal settings. In the steepest part of the step response, the variable a is its parameter change and τ is its time delay, so that a/τ would be its steepest slope. Their tuning rules are set for different controllers. For a P controller, the prescribed constant is K_p=1/a; for a PI controller, the prescribed constants are K_p=0.9/a, K_i=3τ; and for a PID controller, the prescribed constants are K_p=1.2/a, K_i=2τ, and K_d=1/(0.5 τ). Ziegler and Nichols also devised constants for a frequency method, similar to what was described earlier. First the controller is set to operate under P only, and P is increased until the system begins to oscillate. At this setting, the critical value of K_p is recorded as K_c, and the period of oscillations as T_c. Then for a P controller, the prescribed constant is K_p=0.5 K_c; for a PI controller, K_p=0.4 K_c and K_i=0.8 T_c; and for a PID controller, K_p=0.6 K_c, K_i=0.5 T_c, and K_d=0.125 T_c. There have been additional tuning rules developed by Tyreus and Luyben, Marlin, and others. Practically, however, PID tuning requires trial-and-error and empirical quantification of the system stability.

APPENDIX B

NONPERIODIC APERTURE ARRAYS

This appendix shows some prototypes of nonperiodic aperture arrays fabricated in a five-inch photomask, where holes are etched into a chrome thin film. A variety of patterns were attempted: a Fibonacci spiral, a Golden spiral, a radial circular pattern, and arrays of circular spirals. Since diffraction image scales with the pattern image, so the actual length scales of the image versus the aperture array are relative. In these data sets, the patterns shown are about 750 μm in size, and the diffraction pattern image spans approximately three inches.

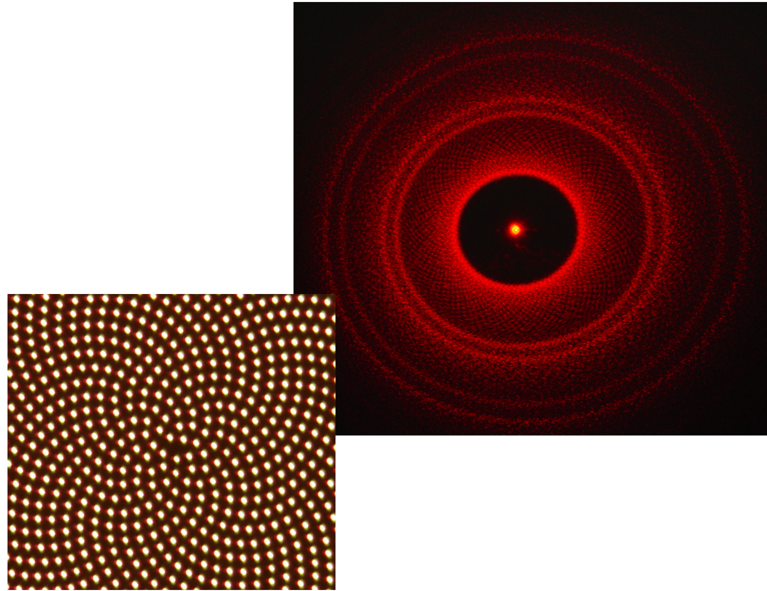


Figure B.1: Fibonacci spiral design, where $r \approx \sqrt{\theta}$.

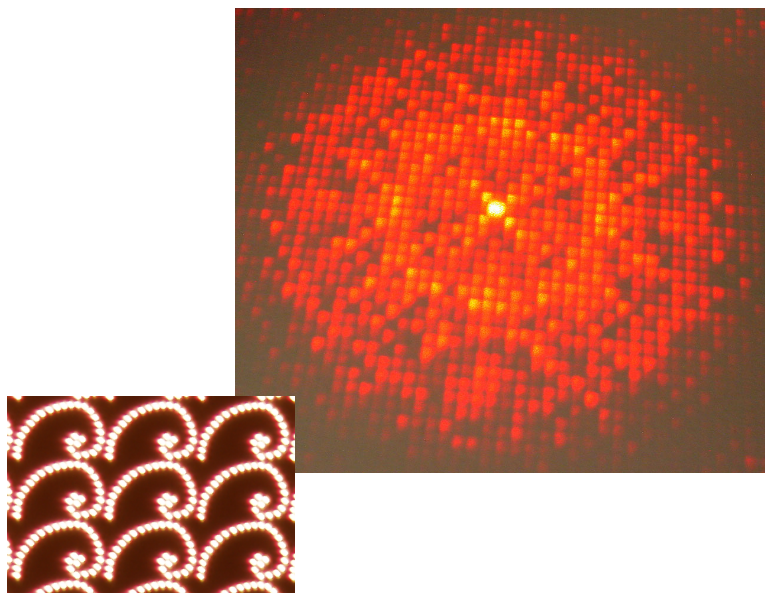


Figure B.2: Arrays of Golden spirals. Holes were placed approximately along the Golden spiral, $r = a \exp b\theta$, where $b = \ln \phi/(\pi/2)$ where ϕ is the Golden ratio.

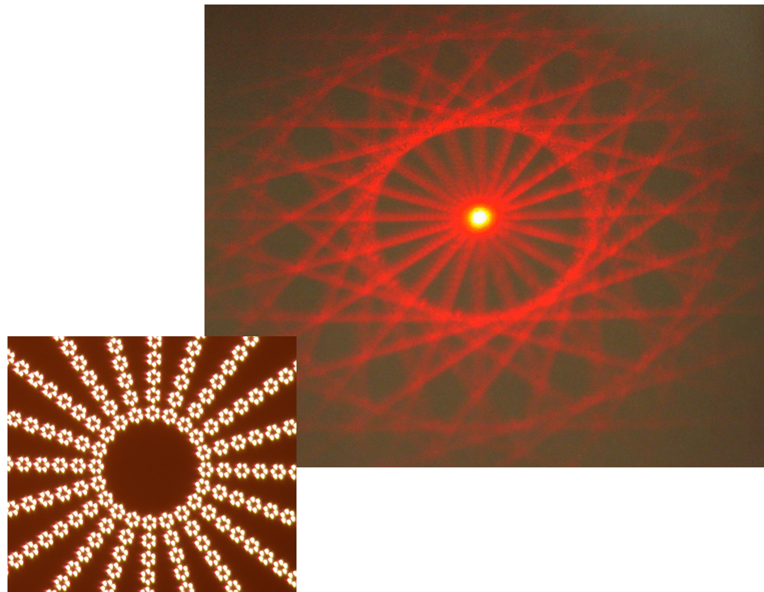


Figure B.3: Radial structure of a circular pattern. Circles of apertures are placed radially away from the center point.

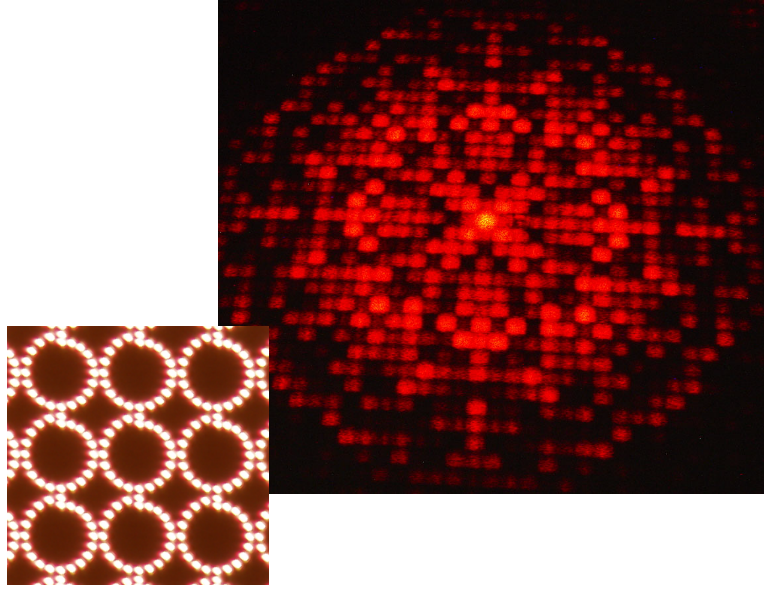


Figure B.4: Arrays of spiral of circles. Circle of apertures, whose radius decreases slowly around the circumference, are placed in an array.

Therefore, it is seen that a variety of nonperiodic structures can be used for diffraction images.

BIBLIOGRAPHY

[1] C. Affolderbach and G. Mileti. *J. Opt. Las. Eng.*, 43:291, 2005.

[2] L C Angrave, N E Booth, R J Gaitskell, G L Salmon, and M Harston. *Phys. Rev. Lett.*, 80:1610, 1998.

[3] E. Arimondo, M. Inguscio, and P. Violino. *Rev. Mod. Phys.*, 49:31, 1977.

[4] F L Arnot. *Proc. Roy. Soc. A*, 133:615, 1931.

[5] G. P. Barwood, P. Gill, and W. R. Rowley. *J. Phys. E: Sci. Instrum.*, 966:966, 1988.

[6] G. P. Barwood, P. Gill, and W. R. C. Rowley. *J. Mod. Opt.*, 37:749, 1990.

[7] J. R. Beacham and K. L. Andrew. *J. Opt. Soc. Am.*, 61:231, 1971.

[8] S Belogurov, G Bressi, G Carugno, E Conti, and D Iannuzzi. *Nucl. Instrum. Methods A*, 452:167, 2000.

[9] H. M J. M. Boeston, C. C. Tsai, J. R. Gardner, D. J. Heinzen, and B. J. Verhaar. Observation of a shape resonance in the collision of two cold ^{87}rb atoms. *Phys. Rev. A*, 55:636 – 640, 1997.

[10] E C Bullard and H S W Massey. *Proc. Camb. Phil. Soc.*, 26:556, 1930.

[11] C J Cuneo, J J Maki, and D H McIntyre. *Appl. Phys. Lett.*, 64:2625, 1994.

[12] L. Cutler and C. Searle. *Proc. IEEE*, 54:136, 1966.

[13] N. G. de Bruijn. Algebraic theory of penrose's non-periodic tilings of the plane i & ii. *Ned. Akad. Wetensch. Proc. Ser. A.*, 43:39–66, 1981.

[14] S. A. Diddams, D. J. Jones, J. Ye, S. T. Cundiff, J. L. Hall, J. K. Ranka, R. S. Windeler, R. Holzwarth, T. Udem, and T. W. Hänsch. *Phys. Rev. Lett.*, 84:5102, 2000.

[15] G. W. F. Drake, editor. *Atomic, Molecular, & Optical Physics Handbook.* American Institute of Physics, 1996.

[16] T. W. Ebbesen. *J. Phys. Chem. Solids*, 57:951–955, 1995.

[17] C J Eiting, V Krishnamoorthy, S Rodgers, T George, J D Robertson, and J Brockman. *Appl. Phys. Lett.*, 88:064101, 2006.

[18] U Fano, L Curtiss, I Feister, and I Stegun. *Tech. Rep. 13.* National Bureau of Standards, 1952.

[19] G. W. Fox, O. S. Duffdendack, and E. F. Barker. The spectrum of co_2. *PNAS*, 13:302–307, 1927.

[20] R J Gaitskell, L C Angrave, N E Booth, A D Hahn, G L Salmon, and A M Swift. *Phys. Lett. B*, 370:163, 1996.

[21] A. K. Geim and K. S. Novoselov. The rise of graphene. *Nat. Mater.*, 6:183 – 191, 2007.

[22] J. W. Goodman. *Introduction to Fourier Optics.* Roberts & Company Publishers, 2005.

[23] E. V. Gribanova, A. N. Zhukov, I. E. Antonyuk, C. Benndorf, and E. N. Baskova. *Diamond and related materials*, 9:1–6, 2000.

[24] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup. Efficient subpixel image registration algorithms. *Opt. Lett.*, 33:156–158, 2008.

[25] H Guo and A Lal. Self-powered photon source. In *Solid state sensors, actuators and microsystems*, pages 1474 – 1477, Boston, MA USA, 2003.

[26] Ralf K Heilmann, Carl G Chen, Paul T Konkola, and Mark L Schattenburg. Dimensinal metrology for nanometre-scale science and engineering: Towards sub-nanometre accurate encodders. *Nanotechnology*, 15:S504–S511, 2004.

[27] C. H. Henry. *IEEE J. Quantum Electron.*, QE-18:259, 1982.

[28] C. H. Henry. *IEEE J. Quantum Electron.*, QE-19:1391, 1983.

[29] Y Isozumi and S Shimizu. *Phys. Rev. C*, 4:522, 1971.

[30] J. Jang, G. C. Schatz, and M. A. Ratner. *Phys. Rev. Lett.*, 92:085504, 2002.

[31] X. et. al Jia. *Science*, 323:1701–1705, 2009.

[32] T Kitahara, Y Isozumi, and S Shimizu. *Phys. Rev. C*, 5:1810, 1972.

[33] Joen Koenig. *A Century of Excellence in Measurements, Standards, and Technology: A Chronicle of Selected NBS/NIST Publications 1901-2000.* 2001.

[34] E J Konopinski. *Rev. Mod. Phys.*, 15:209, 1943.

[35] C. Kozlowski and P. M. A. Sherwood. *J. Chem. Soc., Faraday Trans. 1*, 80:2099–2107, 1984.

[36] A Lal and J Blanchard. *IEEE Spectrum*, 41:36, 2004.

[37] S K Lamoreaux and J R Torgerson. Neutron moderation in the oklo natural reactor and the time variation of α. *Phys Rev D*, 69:121701, 2004.

[38] E. L. Lewis. Collisional relaxation of atomic relaxed states, line broadening and interatomic interactions. *Phys. Rep.*, 58:1 – 71, 1980.

[39] H Lewis. *Phys. Rev.*, 78:526, 1950.

[40] H Li and A Lal. Self-reciprocating radioisotope-powered cantilever. *J. Appl. Phys.*, 92:1122 – 1127, 2002.

[41] P. J. Lu and P. J. Steinhardt. Decagonal adn quasi-crystalline tilings in medieval islamic architecture. *Science*, 315:1106–1110, 2007.

[42] M. Matsumoto, T. Manako, and H. Imai. *J. Electrochem. Soc.*, 156:B1208–B1211, 2009.

[43] A. N. Matveev, N. N. Kolachevsky, J. Alnis, and T. W. Hänsch. *Quantum Electron.*, 38:895, 2008.

[44] R. L. McCarley, S. A. Hendricks, and A. J. Bard. *J. Phys. Chem.*, 96:10089–10092, 1992.

[45] W. Mizutani, J. Inukai, and M. Ono. *Jpn. J. Appl. Phys.*, 29:L815–L817, 1990.

[46] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov. Electric field effect in atomically thin carbon films. *Science*, 306:666–669, 2004.

[47] M. Ohtsu, H. Fukada, T. Tako, and H. Tsuchida. *Jpn. J. Appl. Phys.*, 22:1157, 1983.

[48] K. S. Pitzer. *Activity coefficients in electrolyte solutions.* CRC Press, 1991.

[49] D. W. Pohl. Some design criteria in scanning tunneling microscopy. *IBM J. Res. Develop.*, 30:417 – 427, 1986.

[50] M. Pourbaix. *Atlas of electrochemical equilibria in aqueous solutions.* National Association of Corrosion Engineers, 2nd edition.

[51] D. Robinson and P. Milanfar. Fundamental performance limits in image registration. *IEEE Trans. Image Process.*, 13(9):1185–1199, 2004.

[52] L. M. Roen, C. H. Paik, and T. D. Jarvi. *Electrochem. Solid-State Lett.*, 17:A19–A22, 2004.

[53] E. A. Rotberg, B. Barreet, S. Beattie, S. Chudasama, M. Weel, I. Chan, and A. Kumarakrishnan. Precision lifetime measurements on alkali atoms and on helium by beam-gas-laser spectroscopy. *Physica Scripta*, T65:48–56, 1996.

[54] J. Rutman. *Proc. IEEE*, 66:1048, 1978.

[55] E B Saloman. Energy levels and observed spectra lines of xenon, xe i through xe liv. *J. Phys. Chem. Ref. Data*, 33:765, 2004.

[56] J. E. Sansonetti. Wavelengths, transition probabilities, and energy levels for the spectra of rubidium (rbi through rbxxxvii). *J. Phys. Chem. Ref. Data*, 35(1):302–421, 2006.

[57] H. Schnatz, B. Lipphardt, J. Helmcke, F. Riehle, and G. Zinner. *Phys. Rev. Lett.*, 76:18, 1996.

[58] A. E. Siegman. *Lasers.* University Science Books, 1986.

[59] J. E. Simsarian, L. A. Orozco, G. D. Sprouse, and W. Z. Zhao. Lifetime measurements of the 7p levels of atomic francium. *Phys. Rev. A*, 57:2448 – 2458, 1998.

[60] H. D. Smyth. The emission spectrum of carbon dioxide. *Phys. Rev.*, 38:2000–2015, 1931.

[61] Y.-W. Son, M. L. Cohen, and S. G. Louie. Energy gaps in graphene nanoribbons. *Phys. Rev. Lett.*, 97:216803, 2006.

[62] L V Spencer. *Phys. Rev.*, 98:1597, 1955.

[63] P. S. Spinney, S. D. Collins, and R. L. Smith. 7:1512–1515, 2007.

[64] D. A. Steck. Rubidium 85 d line data. *http://steck.us/alkalidata*, 2008.

[65] Åstr˙ *Feedback Systems*.

[66] John Townsend. *A Modern Approach to Quantum Mechanics*. 2000.

[67] S. Tsurubuchi and T. Iwai. Simultaneous ionization and excitation of co_2 by electron-impact. *J. Phys. Soc. Japan*, 37(4):1077–1081, 1974.

[68] T. H. Udem, J. Reichert, R. Holzwarth, and T. W. Hänsch. *Phys. Rev. Lett.*, 82:3568, 1999.

[69] K. Vahala and A. Yariv. *IEEE J. Quantum Electron.*, QE-19:1102, 1983.

[70] K. Vahala and A. Yariv. *IEEE J. Quantum Electron.*, QE-19:1096, 1983.

[71] S. Viciani, M. Gabrysch, F. Marin, F. Monti di Sopra, M. Moser, and K. H. Gulden. *Opt. Coommun.*, 206:89, 2002.

[72] M Walhout, A Witte, and S L Rolston. *Phys. Rev. Lett.*, 72:2843, 1994.

[73] J A Wilkerson, J McLaren, and U Raut. *Nucl. Instrum. Methods A*, 500:345, 2003.

[74] J. Willsau and J. Heitbaum. *J. Electroanal. Chem.*, 161:93–101, 1984.

[75] Y. Yamamoto. *IEEE J. Quantum Electron.*, QE-19:34, 1983.

[76] J. Ye, H. Schnatz, and L. W. Holberg. *IEEE J. Select. Topics. Quantum Electron.*, 9:1041, 2003.

[77] N. Yoshimizu, A. Lal, and C. R. Pollock. *Appl. Phys. Lett.*, 89:021909, 2006.

[78] A. Zarka, A. Abou-Zeid, D. Chagniot, J.-M. Chartier, O. Číp, J.-F. Cliche, C. S. Edwards, F. Imkenberg, P. Jedlićka, B. Kabel, A. Lassila, J. Lazar, M. Merimaa, Y. Millerioux, H. Simonsen, M. Têtu, and J.-P. Wallerand. *Metrologia*, 37:329, 2000.