

# A Novel Application of HEP Technology to the Problem of Audio Preservation\*

Vitaliy Fadeyev and Carl Haber  
Physics Division

Lawrence Berkeley National Lab  
Berkeley, California USA

\*thanks to LBNL Tech Transfer for support



# Outline

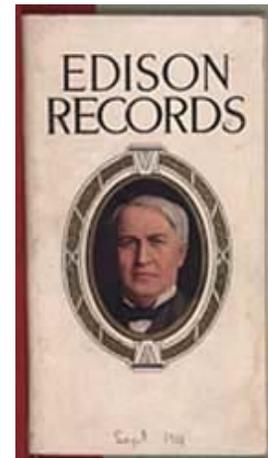
- Introduction
- Background
  - History
  - Issues
- The Imaging Method
  - Basic Idea
  - Relationship to HEP methods
  - Advantages
  - Detailed Discussion of Elements
- Proof of Concept Test
- Conclusions
- Future Directions

# Introduction

- We have investigated the problem of optically recovering mechanical sound recordings **without contact** to the medium
- Our approach evolved naturally out of tracking and testing methods in use with position sensitive detectors for HEP.
- This work may address some concerns of the preservation, archival, and research communities:
  - The reconstruction of damaged media
  - The playback of delicate media
  - Mass digitization and storage
- Message to take away from today's presentation:
  - Optical techniques can produce acceptable reproductions and some improvements.
  - Measurement, data storage, and computing technologies may be approaching performance levels required for this application.
  - This is an interesting and useful application of the HEP approach.

# History (1)

- Recorded sound was introduced by Edison in 1877 who embossed audio data onto metal foil
- A variety of media and methods used since then
  - Wax cylinders with vertical modulation
  - Shellac disks with vertical or lateral modulation
  - Vinyl disks with lateral or 45/45 (stereo) modulation
  - Acetate instantaneous recordings, lateral modulation
  - Metal reversed stampers (disks) and galvanos (cylinders)
  - Magnetic tape and wire
  - Compact digital disks (CD)
- Essentially all pre-1948 recordings were mechanical



# History (2)

- What does the archive contain?
  - Primary recordings of key artists, Caruso....
  - Field recordings of sources which underlie much of modern music, American and European folk traditions...
  - Recordings of cultures lost due to modernization, anthropological and linguistic data...
  - Speeches & spoken words of historical figures, Churchill, Roosevelt...
  - Early radio broadcasts (acetates)...
  - Live performances, events,....



# Major Sound Archives

- The US Library of Congress 2.5M recordings
- The British National Library: 2.5M recordings
- The New York Public Library: 500K recordings
- The Berlin Phonogramm-Archiv: 145K total, 30K cylinders
- The Vienna Phonogramm-Archiv: 50K
- The Stanford Archive of Recorded Sound: 250K recordings
- The Belfer Archive at Syracuse University: 340K recordings
- The Yale Historical Sound Collection: 160K recordings
- The Edison National Historical Site (NJ, USA): 37K cylinders

# Issues and Concerns

- Can recordings be mass digitized in an efficient way to enable preservation and access for future users?
  - Diverse formats
  - Damaged samples which require intervention or are impossible to play at all
  - Further damage to delicate samples
- Can samples which are of particular value to someone be recovered or improved in a useful way?

# Traditional Contact Playback



Bulky stylus riding in a narrow groove => Issues with

- tracking
- condition of the groove
- wear of the groove

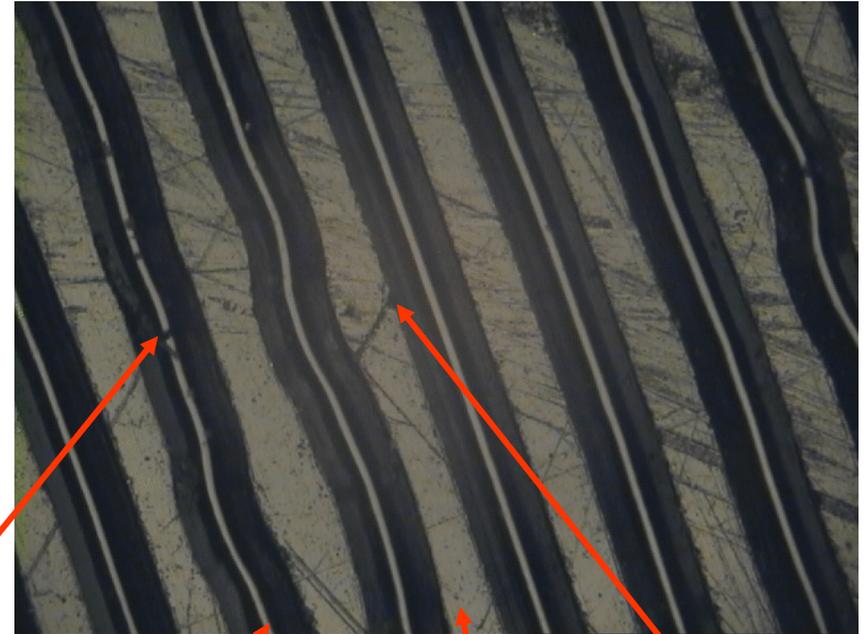
Presence of skilled manpower or supervision is de-facto required.

# A Non-contact Method

- Using optical techniques, the pattern of grooves or undulations in a recording surface can be **imaged**.
- To cover a surface (thousands of) sequential views can be acquired.
- Views can be stitched together.
- The images can be **processed** to remove defects and **analyzed** to **model** the stylus motion.
- The stylus motion model can be sampled at a standard frequency and **converted** to digital sound format.
- Real time playback is **not required** de-facto, method is aimed at reconstruction and digitization.

# Example of Groove Image

- Two dimensional view (2D)
- Field is 1.39 x 1.07 mm
- Groove width is 160  $\mu\text{m}$
- Lighting is perpendicular to surface.
- Bright line is groove bottom.
- Acquired with a digital camera and magnification.



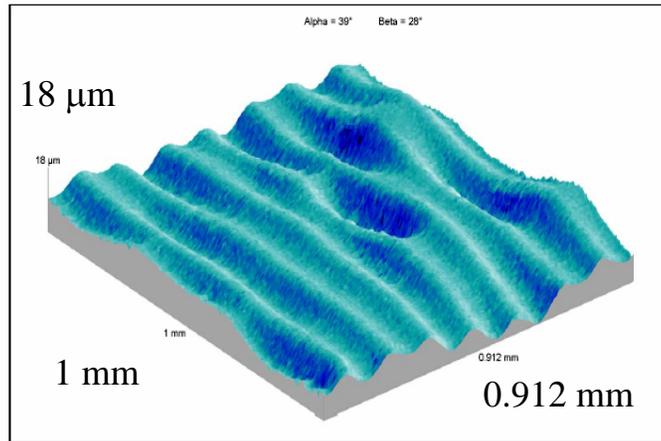
Dust particle

Groove bottom

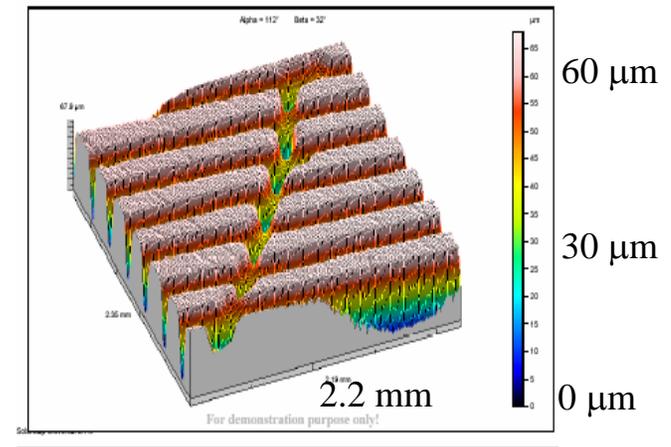
Record surface

Small scratch

# Example of 3D Images: Surface Profiles

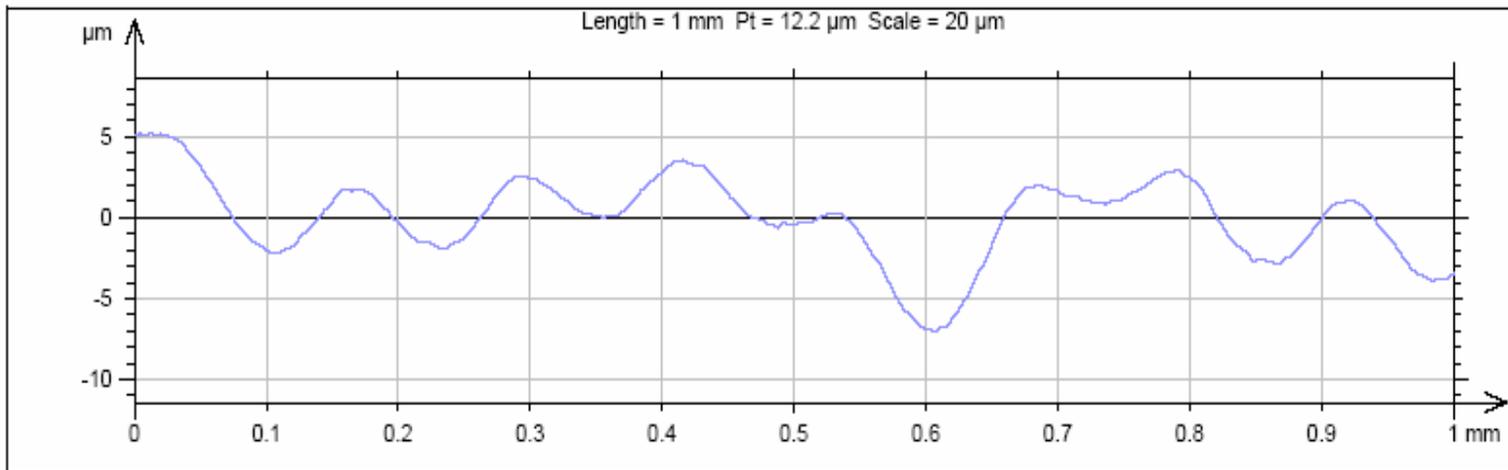
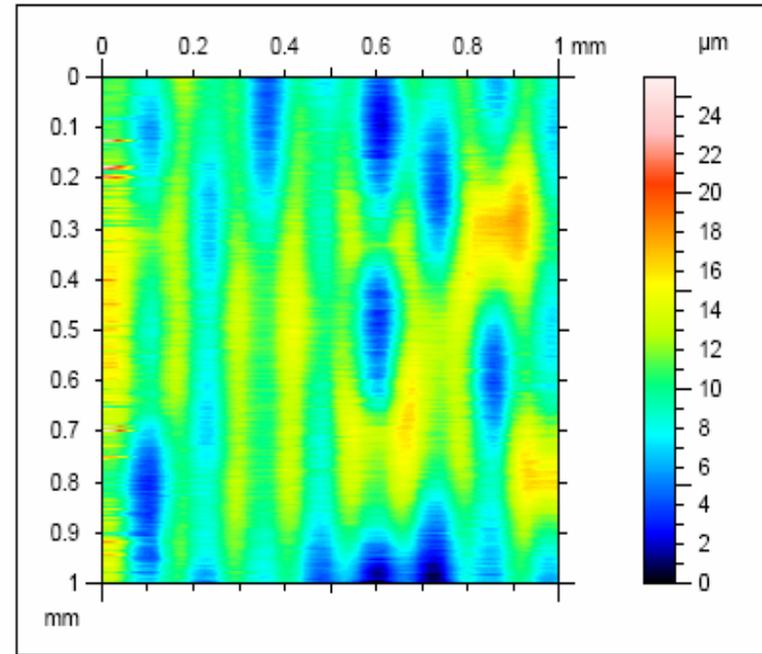
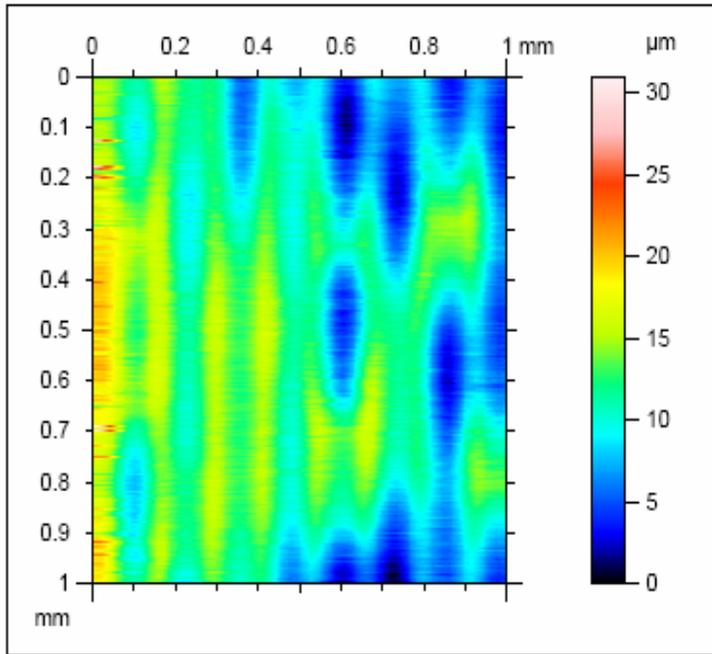


Edison "Blue Amberol" cylinder



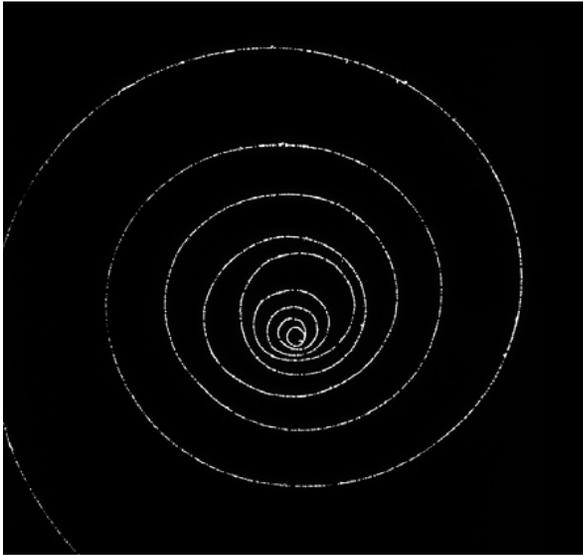
78 rpm disk with large scratch

- Fields are few  $\text{mm}^2$
- Acquired with confocal laser scanning probe
- Suitable for detailed groove reconstruction and vertical modulated recordings



# Relationship to HEP

- The problems of pattern recognition, track finding, and noise suppression are standard in HEP analysis.
- In HEP we commonly deal with large data sets.
- To build and test tracking systems, such as the silicon detectors for ATLAS and CDF, we use precision optical metrology methods.
- Following the grooves in a record is not that different than following a track in detector. Similarity also to the automated bubble chamber film scanning methods.
- Familiar concepts of pattern recognition, tracking, and noise reduction can be applied to **reconstruct** and **repair** recordings.



Photograph of bubbles formed along the trajectory of an electron as it loses energy in a bubble chamber.

Measurement of tracks in a particle detector is similar to following the groove in a mechanical recording – pattern recognition, noise reduction issues are familiar.

Measurement methods and precision required to build detectors are similar to that required for audio reconstruction



# Bubble Chamber “spiral reader” installed in 50B-6238 circa 1970



Sept 16, 2003

LBNL Physics Division RPM

15 Vitaliy Fadeyev  
LBNL

# Advantages of Imaging Method

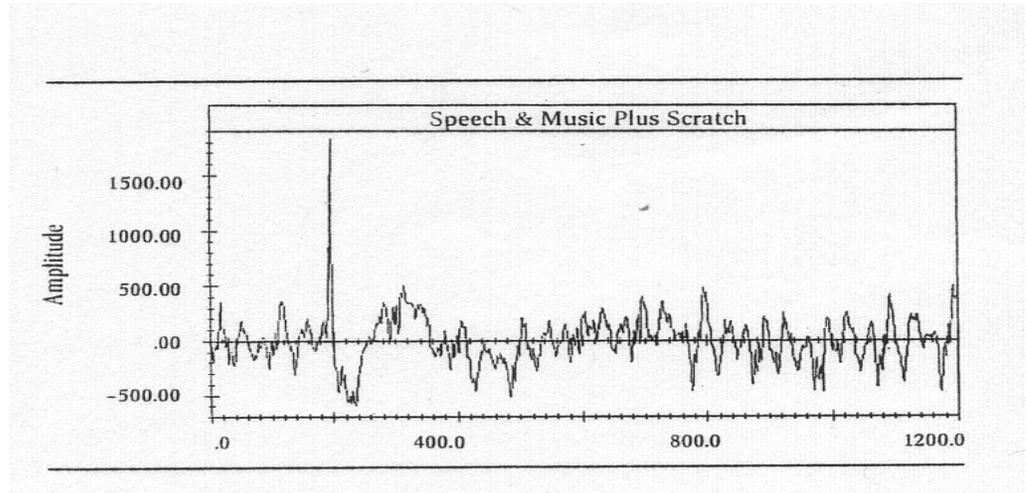
- Delicate samples can be played without further damage.
- Broken samples can be re-assembled virtually.
- Independent of record material and format – wax, metal, shellac, acetates...
- Effects of damage and debris (noise sources) can be reduced through image processing. Scratched regions can be interpolated.
- Discrete noise sources are resolved in the “spatial domain” where they originate rather than as an effect in the audio playback.
- Dynamic effects of damage (skips, ringing) are absent.
- Classic distortions and systematics (wow, flutter, tracing and tracking errors, pinch effects etc) are absent or removed as geometrical corrections
- No mechanical method needed to follow the groove.
- Suggests a method for mass digitization, full 3D maps.

# Material and Format Independence

- Image of metal “stamper” used to mold plastic record.
- Molding technology is obsolete.
- Can be played with special “cowboy” stylus which rides ridges
- But easily imaged in 2D



# Dynamic Effects



- Continuum of imperfections up to a full skip.
- Mechanical stylus responds dynamically to these imperfections
- Result is a “ringing” which may persist as an artifact if only clicks are removed in a standard digital remastering
- Plot is from Rayner, Vaseghi, and Stickells, FIAF Joint Technical Symp, “Archiving the Audio-Visual Heritage”, May 20-22, 1987

# Relationship to Other Work

- **Laser turntables (ELP)**: these devices work off a reflected laser spot only and are susceptible to damage and debris and sensitive to surface reflectivity.
- **Stanke and Paul, “3D Measurement and modeling in cultural applications”, Inform. Serv. & Use 15 (1995) 289-301**: use of image capture to read cylinder “galvanos”, depth was sensed from greyscale in 2D image – lacks resolution required for good reconstruction.

# Relationship to Other Work

- S.Cavaglieri, Johnson, and Bapst, Proc of AES 20<sup>th</sup> International Conference, Budapest, Oct 5-7, 2001: use of photographic contact prints and scanner to archive groove pattern in 2D – insufficient resolution, no 3D analog.
- O.Springer (<http://www.cs.huji.ac.il/~springer/>): use of desk top scanner on vinyl record – lacks resolution for useful reconstruction, no notion of magnification nor image processing to improve data.
- Penn et al at Belfer Lab: no information available, some sort of real time interferometric playback system (??), no notion of image processing

# Elements

- **Imaging**
  - Requires sufficient resolution to measure the minimal undulations of the surface. The scale is  $\sim 0.2\text{-}1\ \mu\text{m}$  (lateral) for the pre-vinyl era.
  - Lateral recordings can be imaged in 2D or 3D.
  - Cylinders, vertical disks, and complex damage structures require 3D
  - Method must be fast enough for efficient application.
- **Processing**
  - Effective algorithms to capture information content of grooves and reject noise.
- **Analysis**
  - Transform spatial pattern of groove position into audio response – physical model.
- **Conversion to standard audio formats**

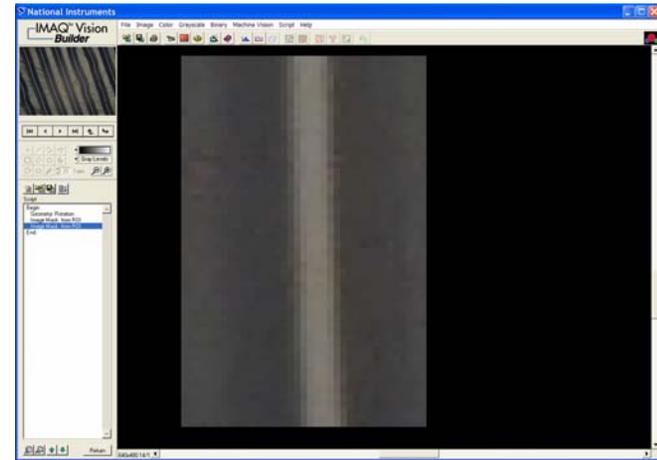
<b>Parameter</b>	<b>78 rpm, 10 inch</b>	<b>Cylinder</b>
Revolutions per minute	78.26	80-160
Max/Min radii inches (mm)	4.75/1.875 (120.65/47.63)	2– 5 fixed (50.8-127)
Area containing audio data (mm <sup>2</sup> )	38600	16200 (2")
Total length of groove	152 meters	64-128 meters

Groove width at top inches ( $\mu\text{m}$ )	0.006 (160)	variable
Grooves/inch (mm) $G_d$	96-136 (5.35-3.78)	100-200 (3.94-7.87)
Groove spacing microns	175 – 250	125 – 250
Ref level peak velocity@1KHz	7 cm/sec	NA
Max groove amplitude (microns)	100 - 125	< 10
Noise level below reference, S/N	17-37 dB	
Dynamic range	30-50 dB	
Groove max amplitude at noise level	1.6 - 0.16 $\mu\text{m}$	

# Imaging Methods

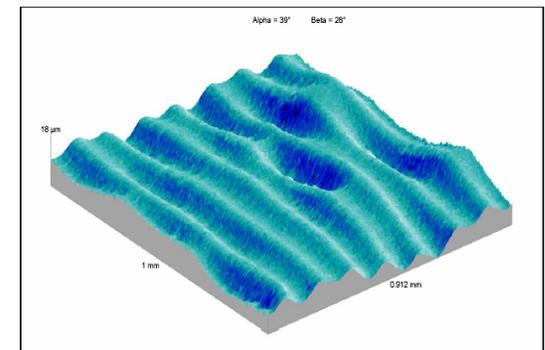
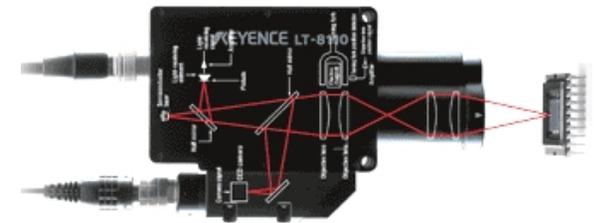
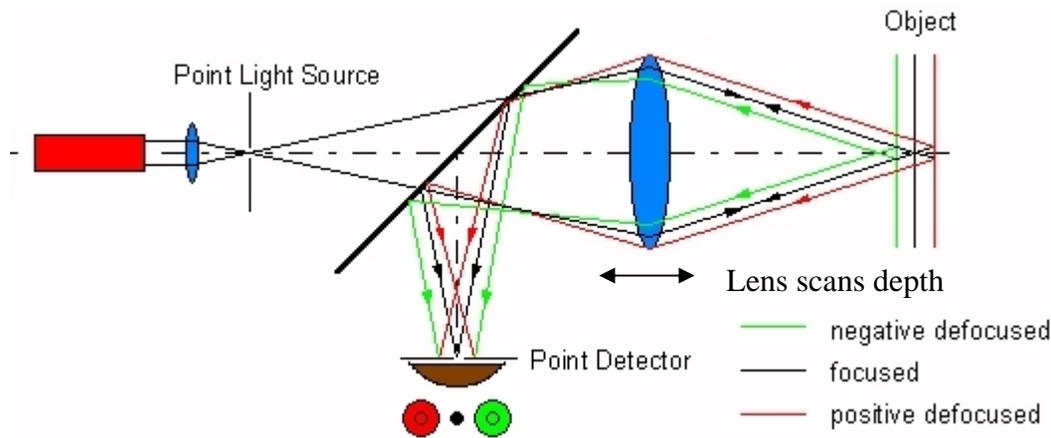
- **Electronic Cameras** – 2D or horizontal only view, frame based
- **Confocal Scanning** – 3D or vertical+horizontal view, point based
- **Chromatic sensors** – 3D, point based
- **White Light Interferometry** – 3D, frame based

# Electronic Camera



- CCD or CMOS image sensor: frame or line format
- Typical frame field of view is 0.7 x 0.54 mm
- Cameras contains 1 x 8000, or 768 x 494 pixels, or up to few Mega-Pixels
- 1 pixel = 0.91 x 1.09 microns on the record surface
- Magnification and pixel size yield sufficient resolution for audio data measurement due to pixel interpolation
- Entire frames acquired at 30-1000 fps

# Laser Confocal Scanning Microscope



- Acquires image point by point
- Vertical resolution is  $\sim 0.1$  micron
- Commercially available
- Point light source is reflected from measurement surface and detected at point detector. Only in-focus rays give signal in detector.
- Complete depth scan occurs 700 times/sec for each point, averaging?
- Horizontal resolution set by point size 1-2 microns

# Chromatic Confocal Sensor

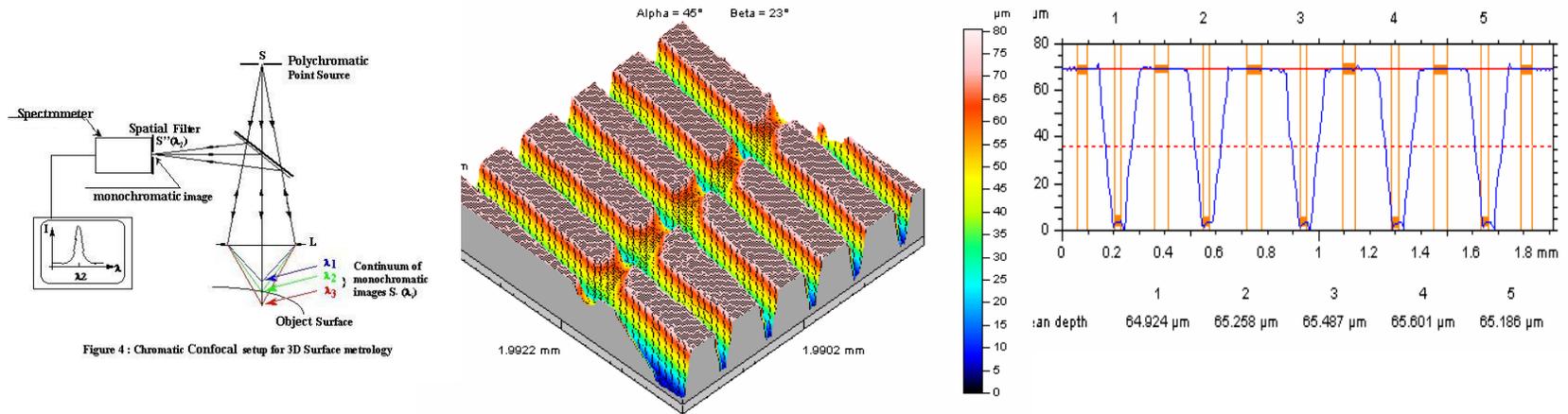
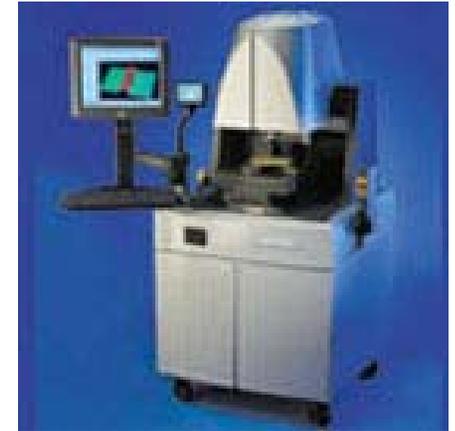
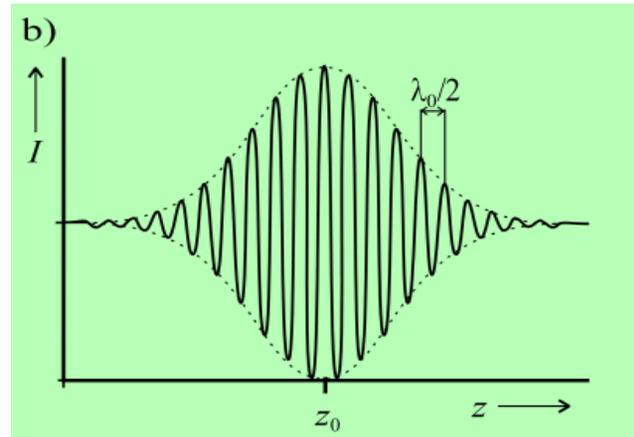
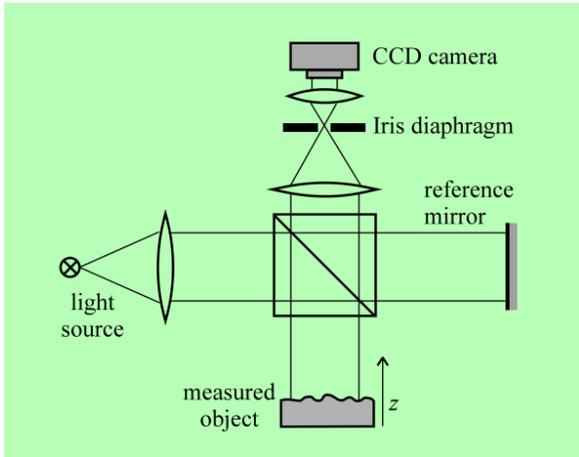


Figure 4: Chromatic Confocal setup for 3D Surface metrology

- Different colors image to different depths on the sample simultaneously – potential for faster scan, up to 1000 points/second may be possible, depending on surface
- Signal detected by spectrometer (color sensitive).
- Issue of reflection off sloped surfaces – data loss

# White Light Interferometry



- Interference principle – waves combine constructively for equal distances traveled.
- Sample is scanned in depth and imaged by frames, at each depth a different interference pattern is found. Frame size is typically 0.6 x 0.4 mm.
- Horizontal resolution is like 2D electronic camera but takes many vertical slices.
- Current systems run 60 fps and require 1-20 seconds per view for vertical scan.
- Scan time depends upon surface angle and reflectivity. Cylinders better than discs
- Potential for faster systems with high fps cameras?

# White Light Interferometry

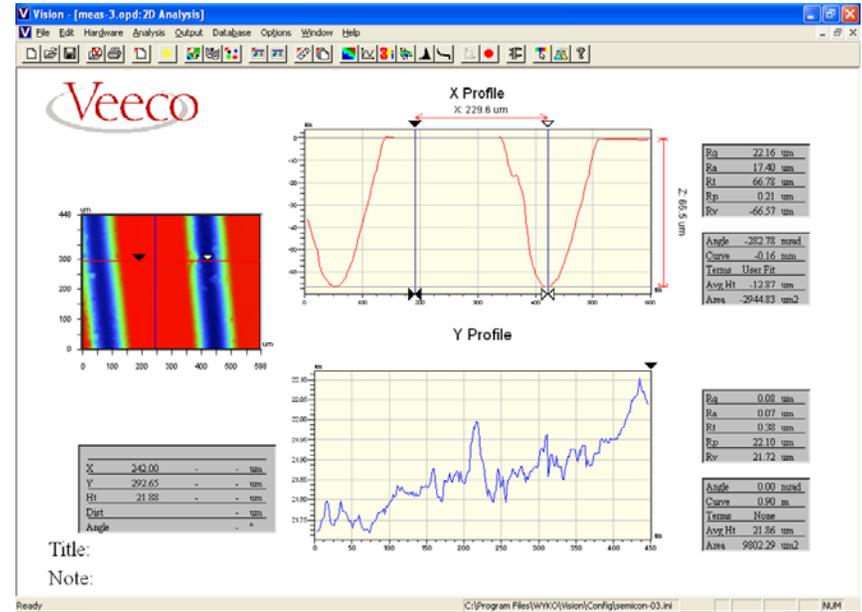
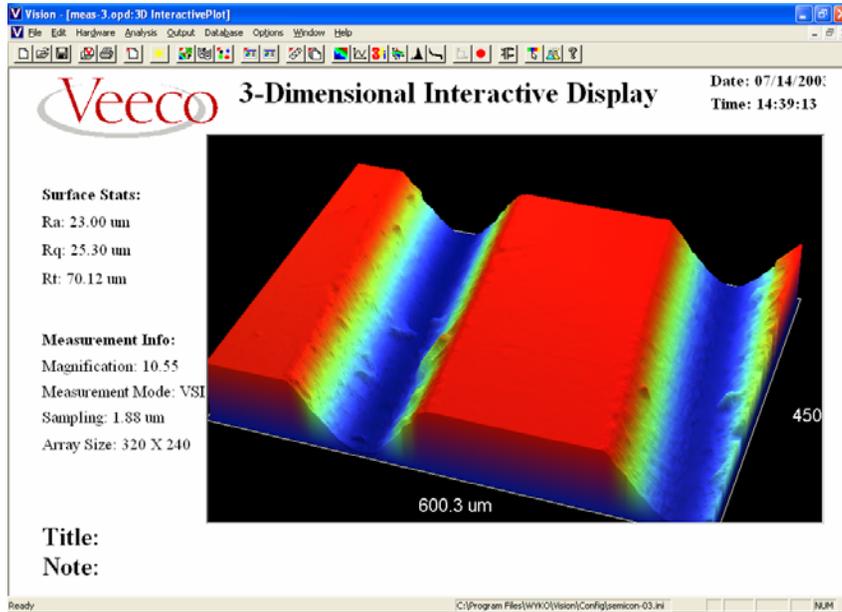
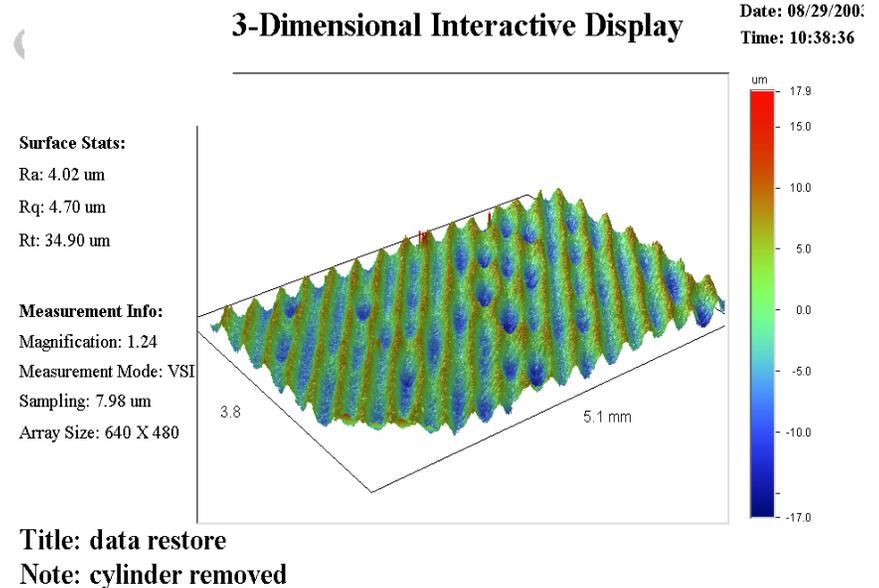
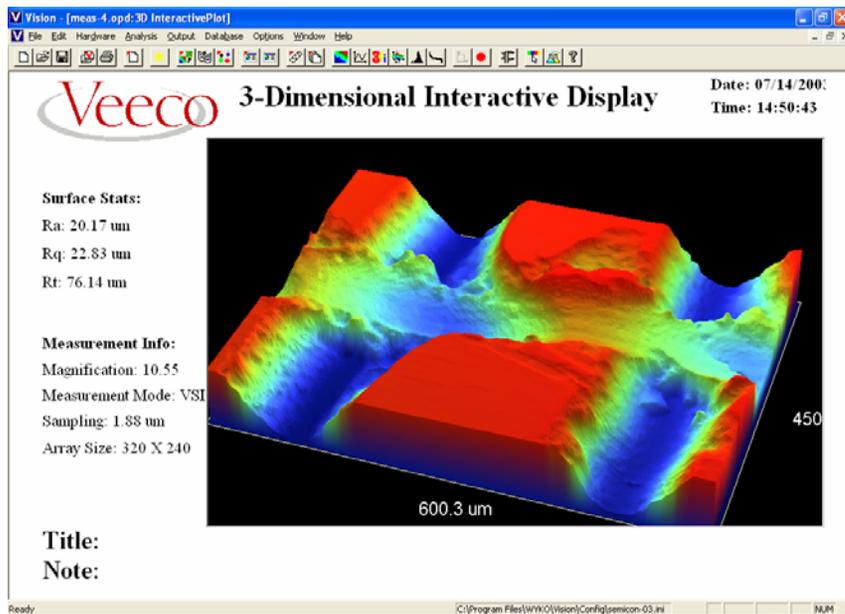


Image from 78 rpm record surface

Issue of data loss from groove sides – angle effect, time required to measure

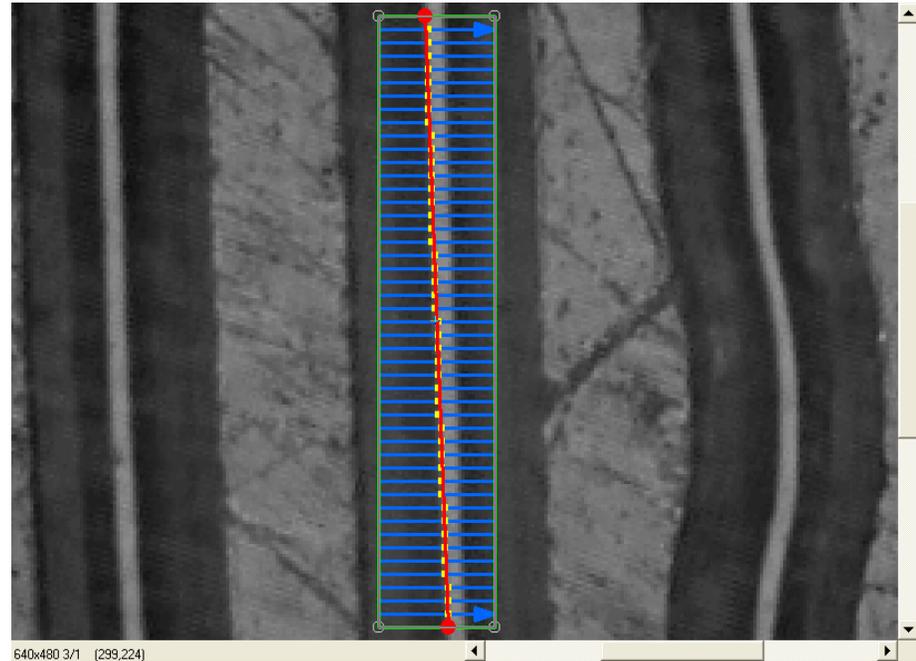
# White Light Interferometry: Scratch and Cylinder Surface



# Image Processing

- Pixel is an intensity or height measurement at a horizontal position  $(x,y)$
- Image processing is performed on the pixels to extract information from the image, including:
  - Measure profiles and distances
  - Find transitions (edge detection)
  - Shape measurement (morphology) and transformation
  - Alter values based upon neighboring pixels (filtering)

# Image Processing



- Brightness profile in grayscale image across a feature
- Edge detection along a series of lines

# Image Processing



Effect of iterated “dilation” operator on 1x3 pixel clusters  
Dust particle is removed from the image

# Signal Analysis

- Once the groove pattern is properly imaged and acquired an analysis is performed to extract the audio data.
- Based upon the physics of the recording process.
- Groove data is already in digital form so analysis methods are numerical.

# Physics of Mechanical Sound Recording

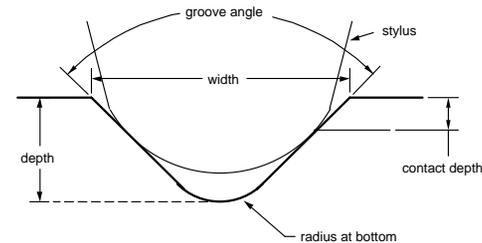
- Playback stylus rides in groove
- For magnetic recording and playback styli, signal is proportional to stylus velocity

$$A_p = \frac{v_p}{2\pi f}$$

“constant velocity recording”

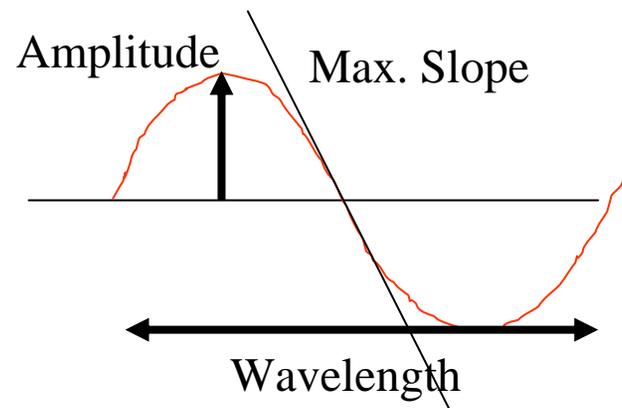
- Mediated by equalization scheme to attenuate low frequencies and boost high frequencies
- Levels are compared by amplitude

$$dB = 20 \log \left( \frac{v}{v_{ref}} \right)$$



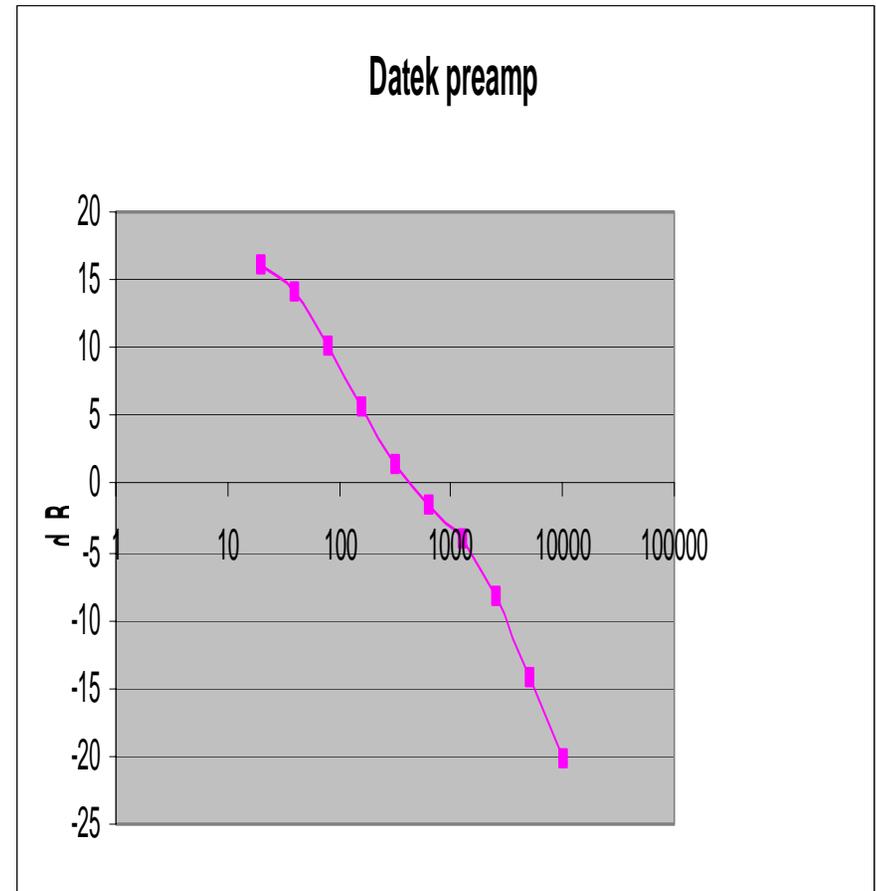
parameter	78 rpm coarse	33 1/3 ultrafine
width	0.006 - 0.008	0.001
depth	-0.0029	-0.0006
contact depth	0.0008	0.0004
radius	0.0015-0.0023	0.00015
angle	82 - 98	87 - 92

units are inches or degrees



# Equalization and Reconstruction

- The constant velocity recording characteristic is modified as follows:
  - Low frequencies are attenuated to avoid excessive groove excursion
  - High frequencies are boosted above surface noise floor
- Playback is equalized to compensate for this.
- Optical reading of **groove displacement is differentiated numerically** to determine stylus velocity and then equalized



# Speed of Method (1)

- Frame based methods
  - 78 rpm disk:  $38600 \text{ mm}^2 = 123,000$  frames
    - 0.54 x 0.7 mm frame
  - @30 fps: 1.2 hours/scan (including 20% overlap)
    - Faster frame rates possible but require high speed photographic methods (sample in motion)
  - Cylinder:  $16200 \text{ mm}^2$  , use WLI method, **but vertical scan requires 1.5-3 seconds per frame:**  
1.2-5 hours per scan (with 60 fps camera)

# Speed of Method (2)

- Line scan methods
  - Apply to 2D scan of 78 rpm disc
  - Media in continuous motion
  - 1 x 8000 pixels => 0.008 mm<sup>2</sup> area per frame
  - 38600 mm<sup>2</sup> = 5,800,000 frames
  - @ 18,000 fps: ~6 minutes/scan

# Speed of Method (3)

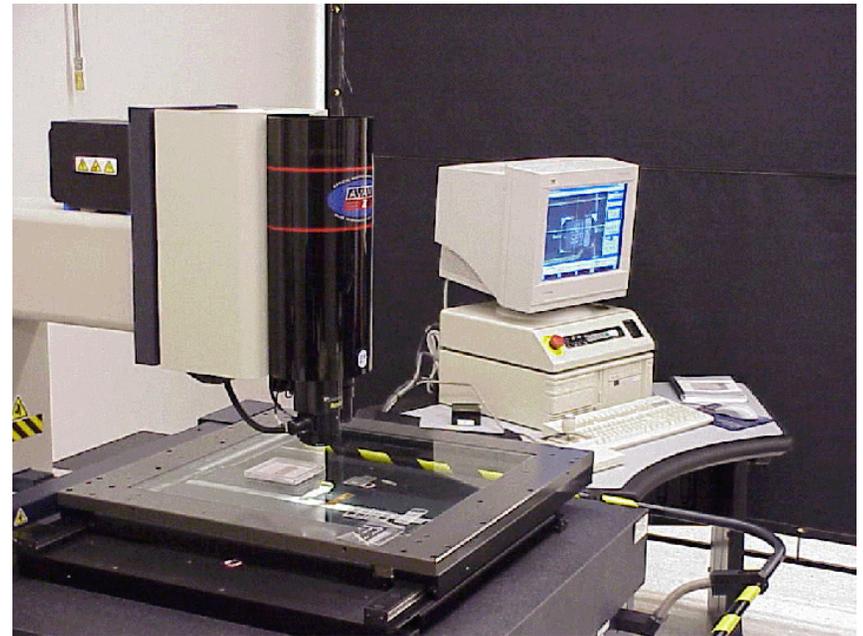
- Point scan based methods (confocal microscopes @1000 points/second)
  - Very sensitive to number of points required for reconstruction
  - High density example: 4x4  $\mu\text{m}$  grid:
    - 78 rpm disk: 152 meters x 160  $\mu\text{m}$ : 1.5 billion points: 400 hours per scan
    - Cylinder: 16200  $\text{mm}^2$  : 1 billion points: 280 hours per scan
  - Low density example: sample groove with 3 points across, at 8  $\mu\text{m}$  intervals along length. Identify defects and return with high density selectively.
    - 78 rpm disk: 3 x 152 meters / 8  $\mu\text{m}$  = 57 million points: 16 hours per scan
    - Cylinder ~12 hours per scan

# Speed of Method - Comments

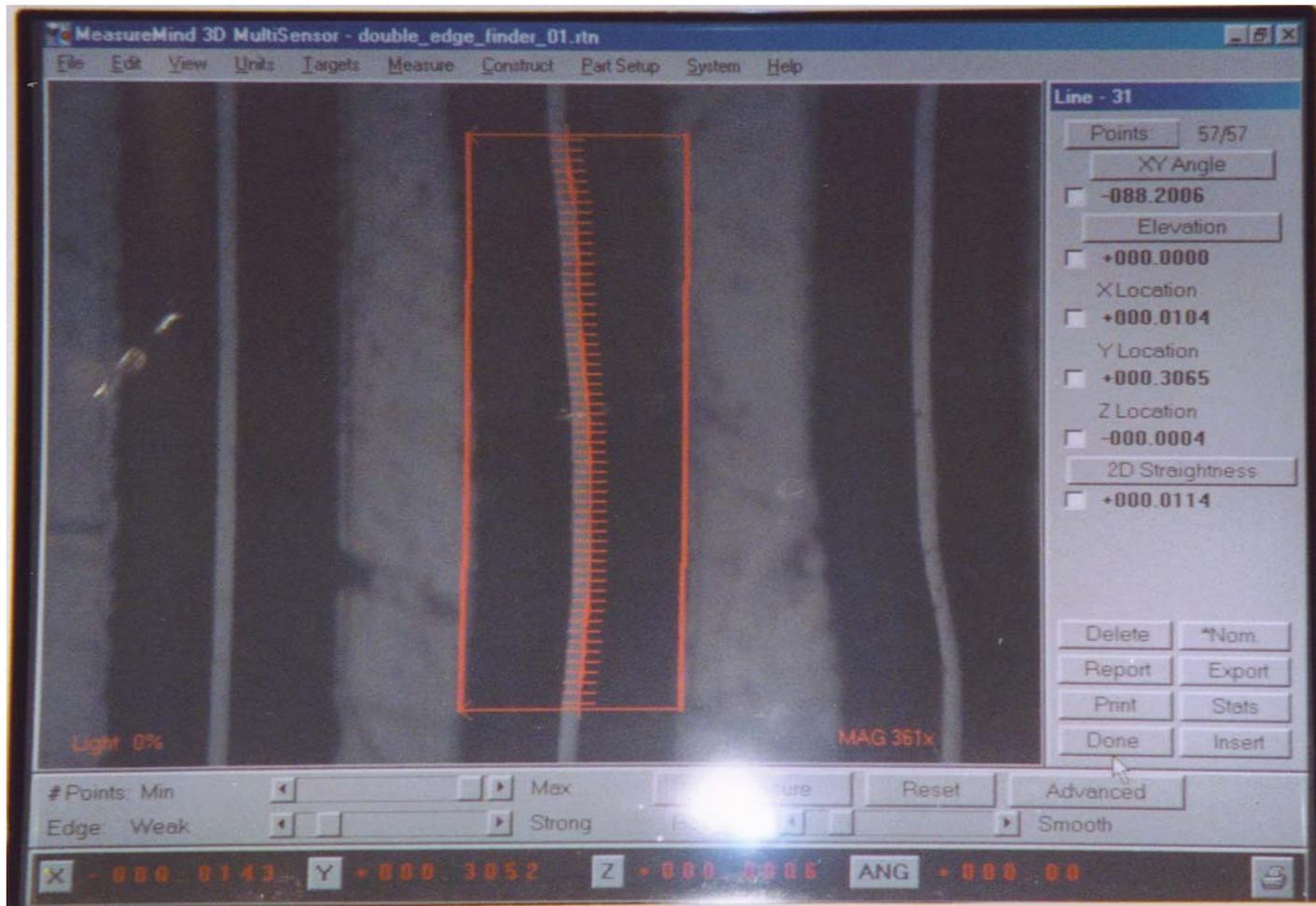
- Discussion only covered image acquisition, but data transmission, real time processing, storage requirements, and mechanics are also significant.
  - Raw 2D images of 78 rpm disk: 190 (grayscale) Mbytes per second of audio data (88 Kbytes/sec in WAV file)
  - Immediate pre-processing (DSPs) could provide reduction.
- Only 2D camera is reasonably efficient for mass digitization. Slow scans OK for special reconstructions. But 3D required for cylinders.
- Speed of 3D surface profilers could increase with new technologies (faster cameras, higher frequency drivers...) recall – none of this was possible 10 years ago.

# Test of Concept using 2D Imaging

- Precision optical metrology system already in use for Particle Physics detector construction at LBL. Non-contact.
- “SmartScope” manufactured by Optical Gauging Products.
- System features zoom microscope with electronic camera and precision stage motion in x-y-z.
- Includes image acquisition with pattern recognition and analysis & reporting software
- Wrote program to scan grooves, report, and process data (offline).

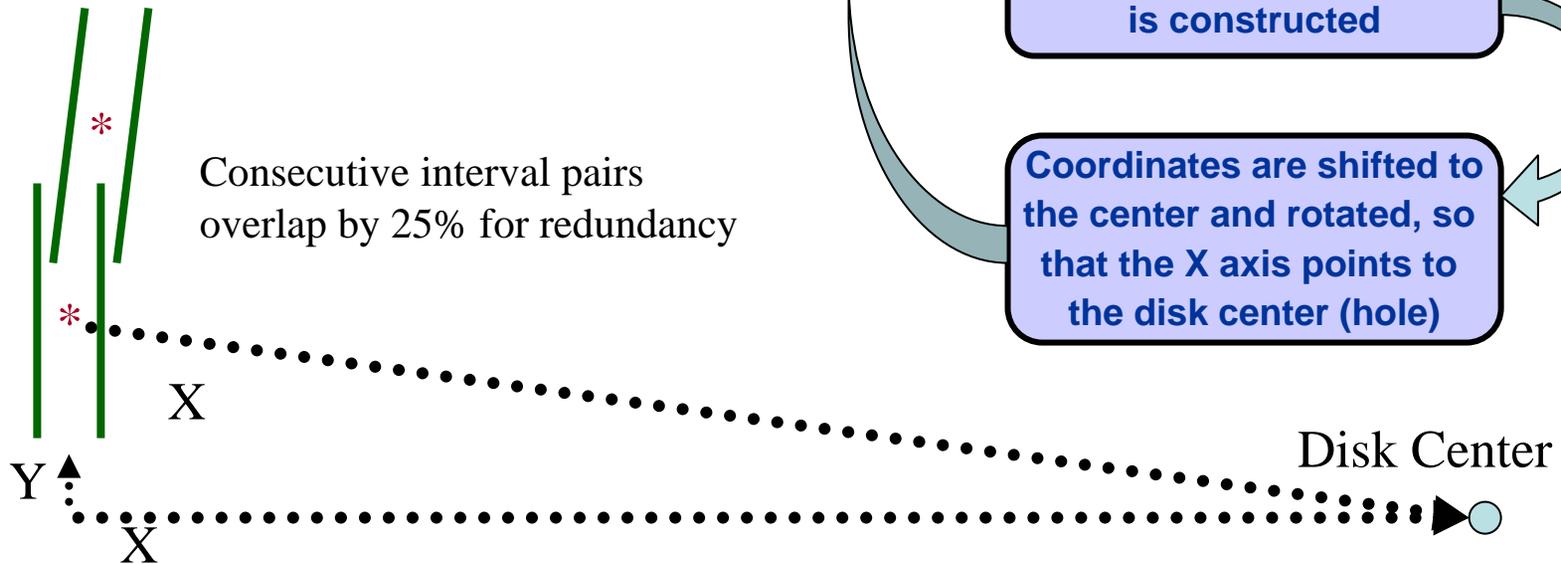


# Example of Smart Scope Edge Finding on Groove Bottom



# The SmartScope Program

An algorithm was devised to follow the groove approximately spiraling to the disk center:



# Video of the SmartScope Run



Sept 16, 2003

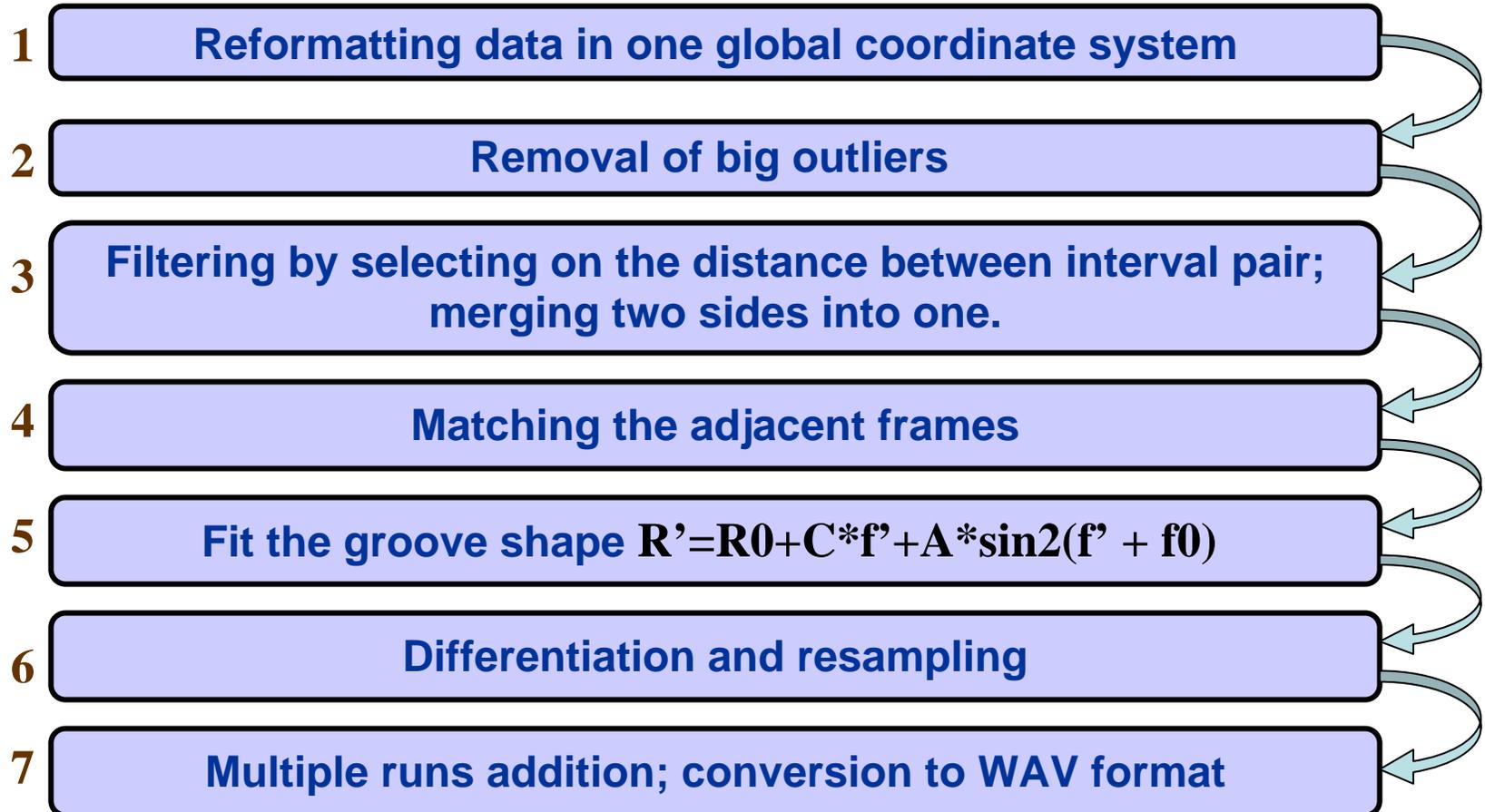
LBLN Physics Division RPM

43 Vitaliy Fadeyev  
LBLN

# Technical Issues

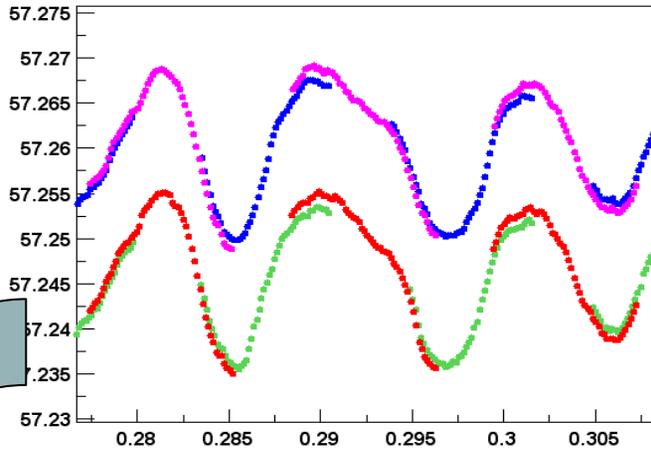
- System used was general purpose, not optimized for this application. It required ~40 minutes to scan 1 second of audio.
- No image recording (for offline use), system performed processing up front with built in software. Output data were found points along features.
- Automatic noise reduction due to rejection of dust or scratches in edge finding process. Also used bottom width measurement as a noise rejection tool offline.
- Number of points along groove, could be increased.
  - 8  $\mu\text{m}$  steps = 66 KHz sampling
- Merging of adjacent frames
- Groove pattern must be differentiated numerically- algorithm selection

# Offline Data Processing



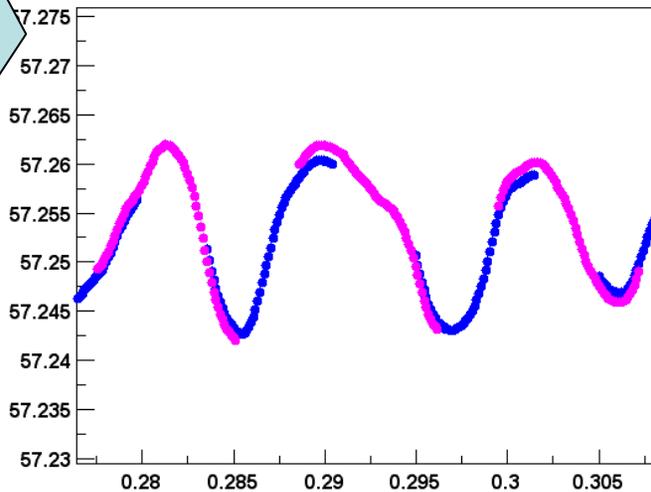
*Raw data after step (2)*

Groove trace data, R[mm]-vs-phi[rad]

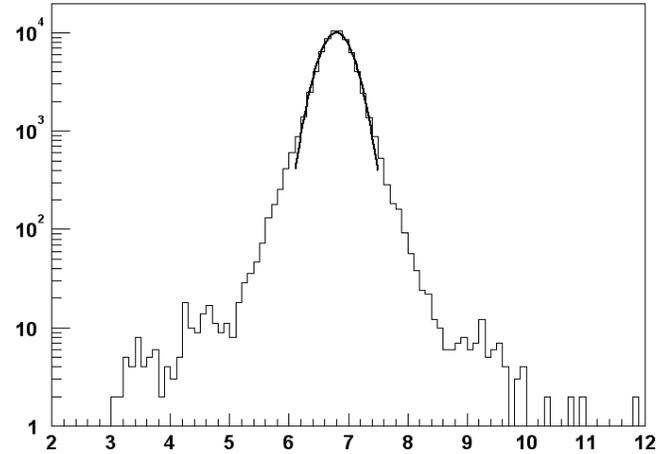


*Merged data after step (3)*

Groove trace data, R[mm]-vs-phi[rad]



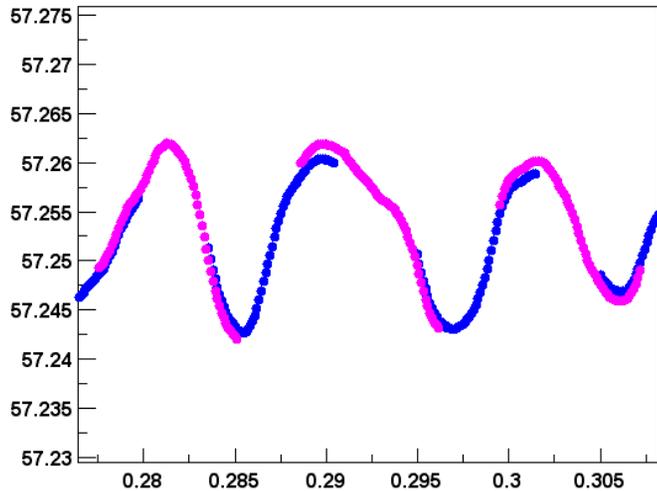
gap



Gap between two edges [ $\mu\text{m}$ ]  
We impose 2.5 sigma cut to  
reduce noise.

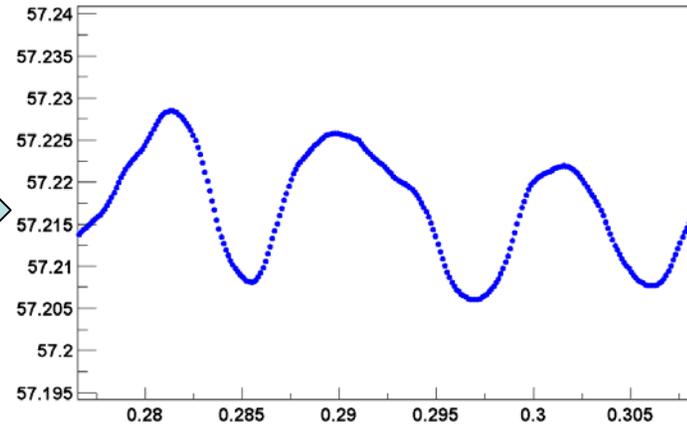
*Merged data after step (3)*

Groove trace data, R[mm]-vs-phi[rad]



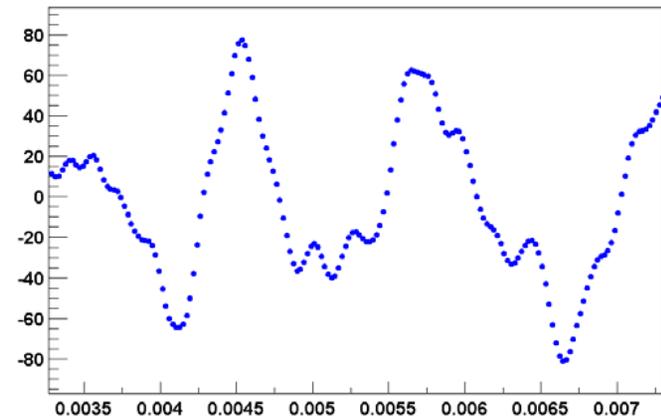
*Matched intervals after step (4)*

Groove trace data, R[mm]-vs-phi[rad]



*Differentiated data after step (6)*

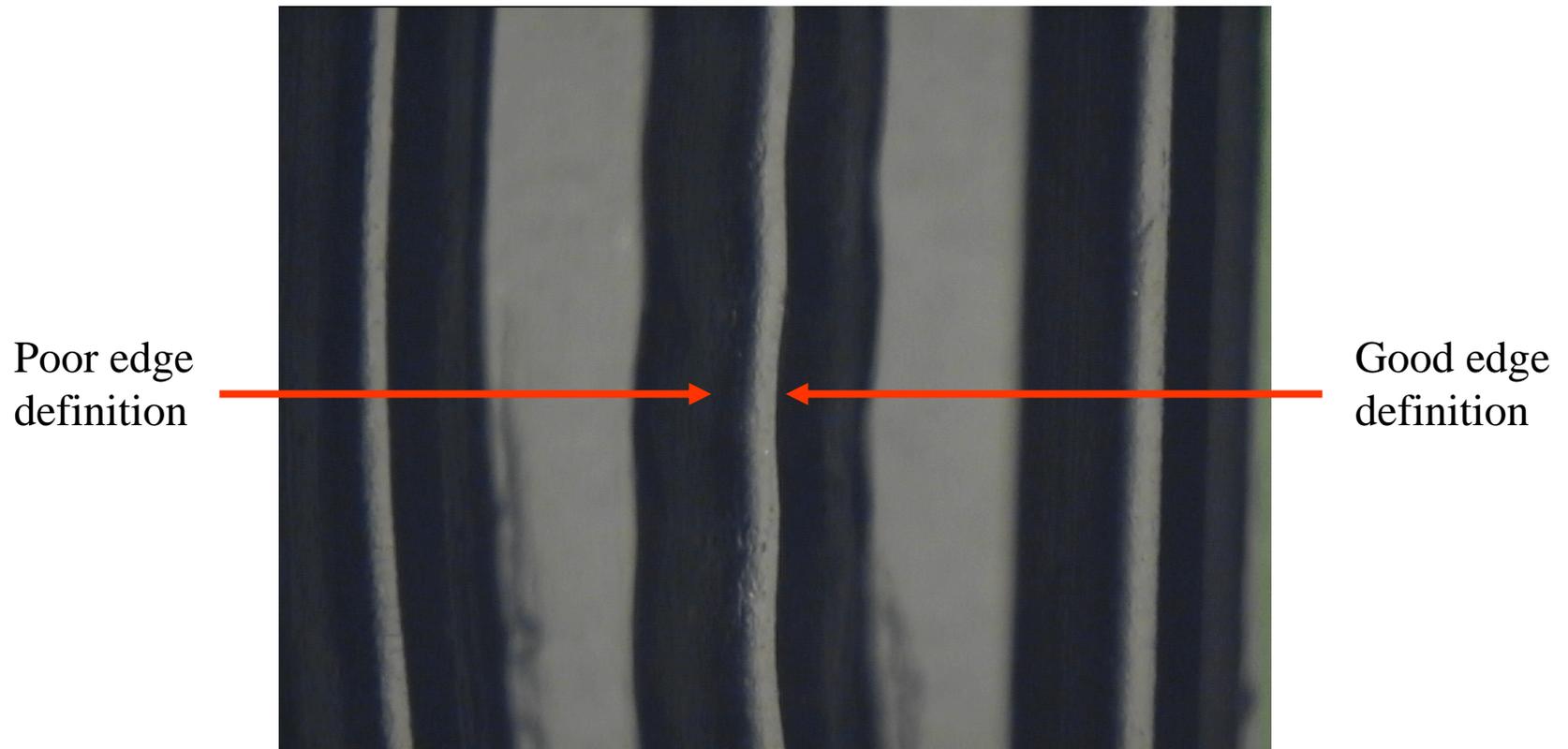
Groove trace data, dR/dT[mm]-vs-time[sec]



**(5) Fit the groove shape  $R' = R_0 + C*\phi' + A*\sin 2(\phi' + \phi_0)$   
(just an offset for the local data)**

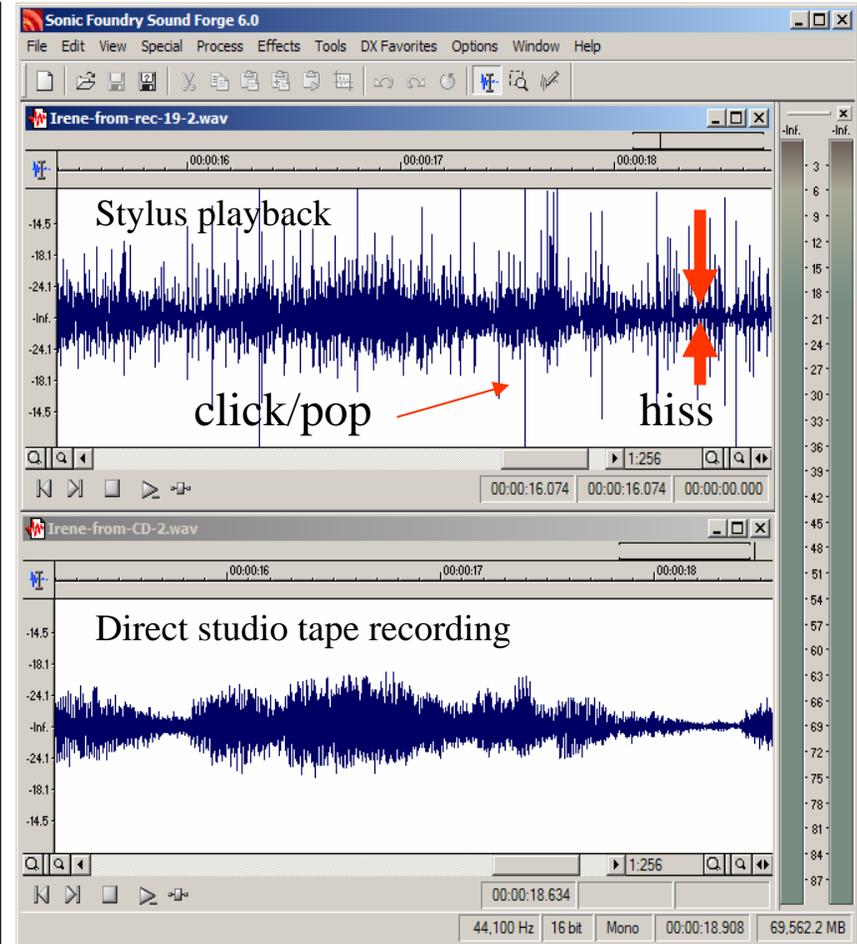
**(6) Numerical differentiation through the 4<sup>th</sup> degree  
polynomial fit to the neighboring 15 points.**

# Issue of Edge Quality in 2D Image



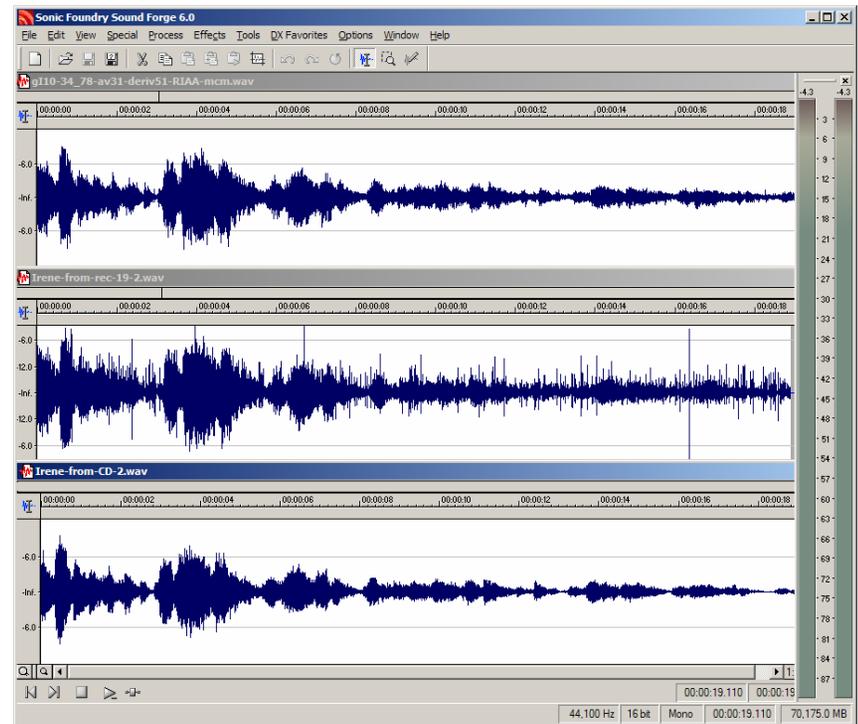
# Noise Sources

- Surface noise or hiss
  - High frequency due to continuous imperfections in groove surface
- Transient impulse noise or “clicks and pops”
  - Due to discrete imperfections such as scratches, random and isolated
- Wow and flutter
  - Not really noise, systematic distortions such as motor speed, off axis rotation



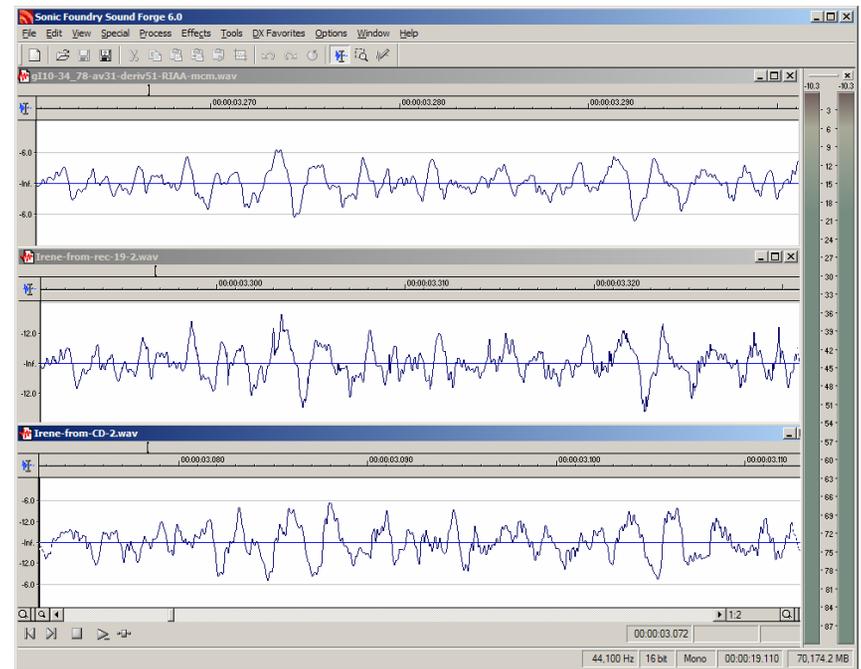
# Results: 1<sup>st</sup> Sample

- Sample is 19.1 seconds
- From ~1950 78 rpm disk
- Top: Imaging method
- Middle: Played by stylus
- Bottom: Professionally re-mastered CD version
- Record was not cleaned.
- No “standard” digital noise reduction software used.



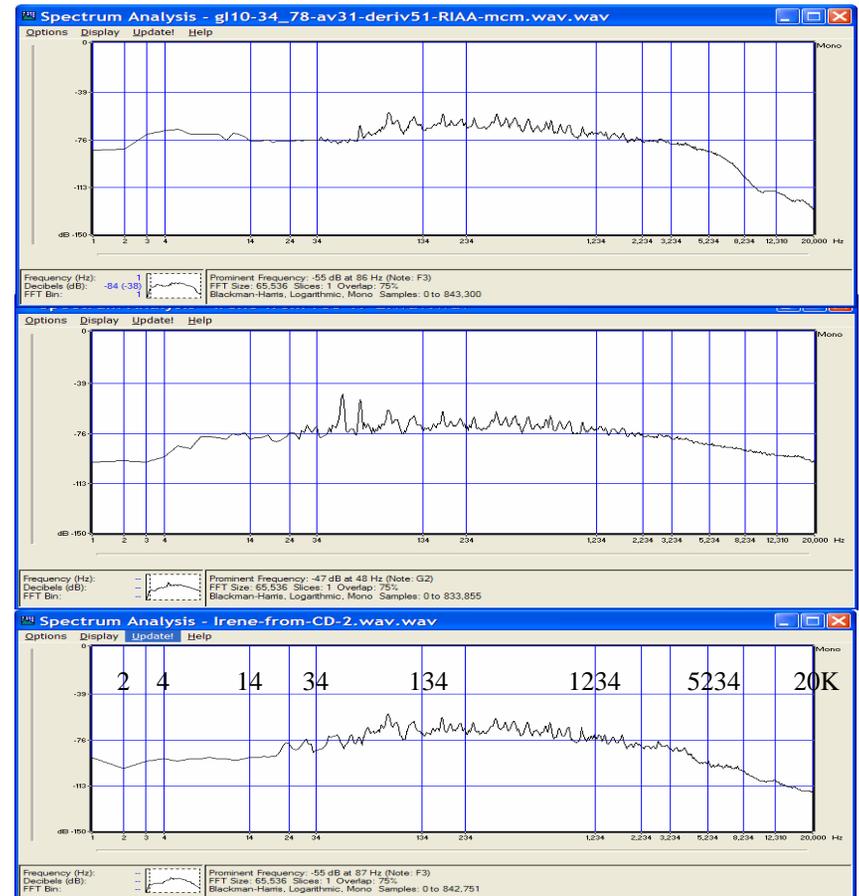
# Zoom In for Detailed Comparison

- 40 ms portion shown
- Striking similarity between optical and stylus reconstruction
- Optical lacks clips and pops, certain noise features, high frequency structures ( $\sim 10$  kHz)
- Qualitative match to CD/tape version



# Frequency Spectra

- FFT spectra of optical (top), stylus (middle), and CD/tape (bottom)
- Audio content in range 100 - 4000 Hz very similar
- More high frequency content in stylus version
- Effects of equalization and differentiation?
- Low frequency structure in optical sample (audible).



# Sound Comparison

“Goodnight Irene” by H. Ledbetter (Leadbelly) and J.Lomax, performed by The Weavers with Gordon Jenkins and His Orchestra ~1950

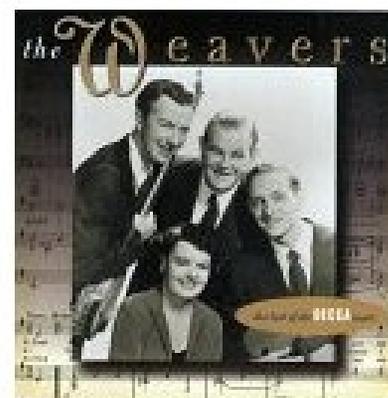
-  Sound from the CD of *re-mastered tape*.
-  Sound from the *mechanical (stylus)* readout.
-  Sound from the *optical* readout.



Sept 16, 2003



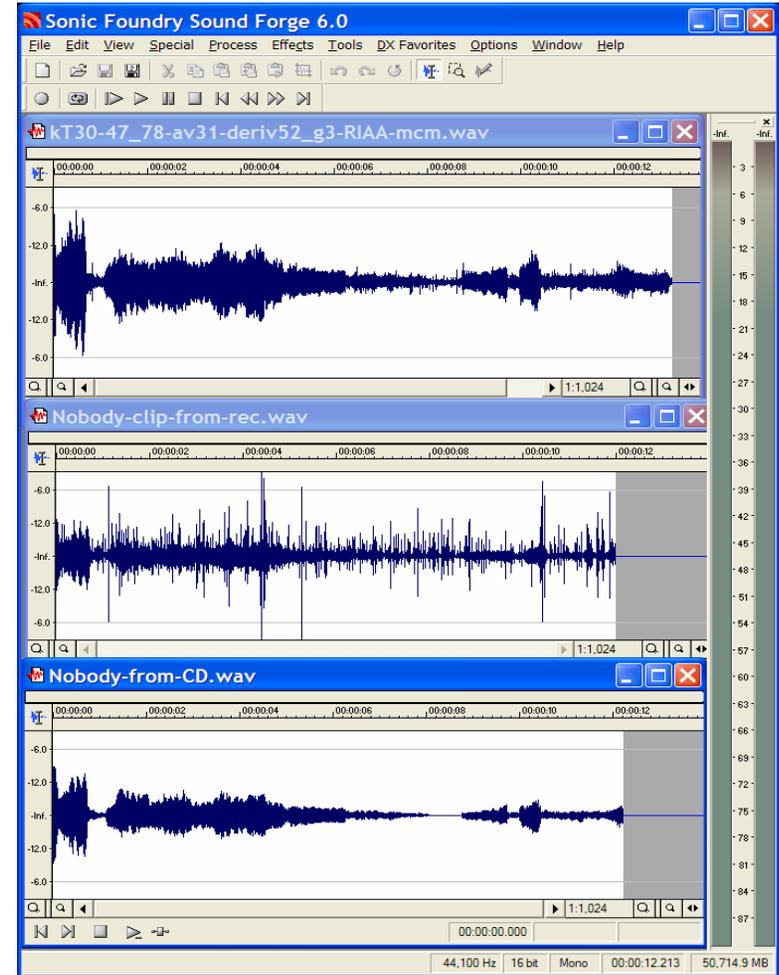
LBNL Physics Division RPM



53 Vitaliy Fadeyev  
LBNL

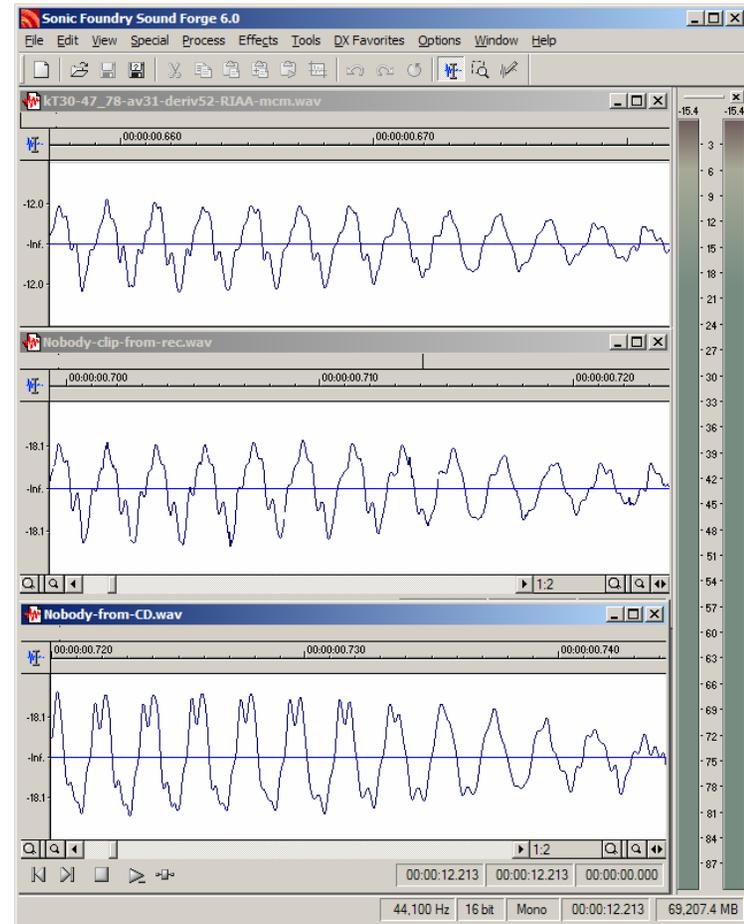
# Results: 2<sup>nd</sup> Sample

- Sample is 13.2 seconds
- From 1947 78 rpm disk
- Top: Imaging method
- Middle: Played by stylus
- Bottom: Professionally re-mastered CD version
- Sample contains a pause at ~8 second point which can be used for noise studies



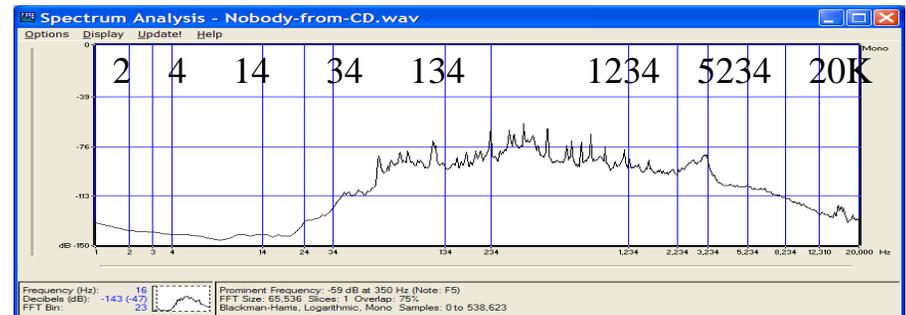
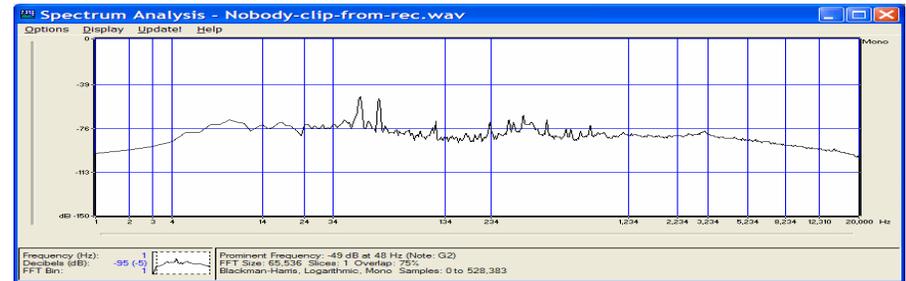
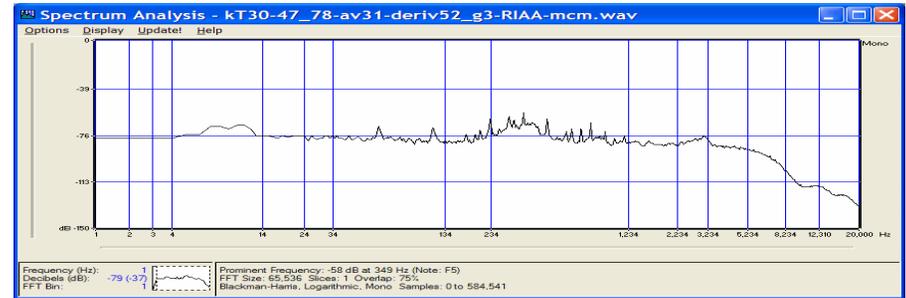
# Zoom In for Detailed Comparison

- 25 ms portion shown
- Striking similarity between optical and stylus reconstruction
- Qualitative match to CD/tape version
- Strong structure around 500 Hz is typical of this sample – single voice and piano



# Frequency Spectra

- Top (optical), middle (stylus), bottom (CD) versions
- Similar mid-range
- Low frequency difference



# Sound Comparison

“Nobody Knows the Trouble I See ” , traditional, performed by Marian Anderson, Matrix D7-RB-0814-2A, 1947



Sound from the CD of *re-mastered tape*.



Sound from the *mechanical (stylus)* readout.



Sound from the *optical* readout.



Sept 16, 2003



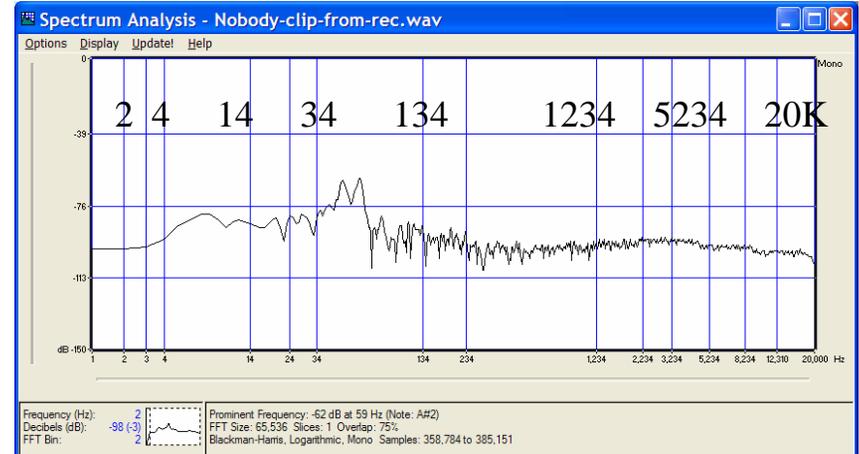
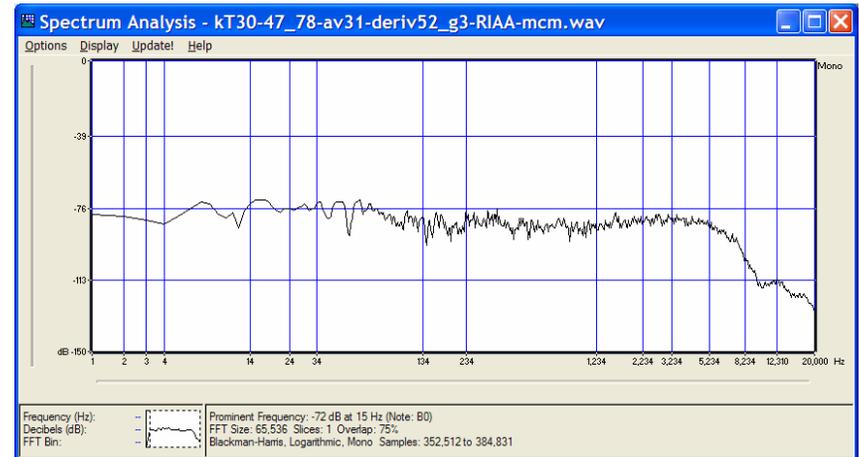
LBNL Physics Division RPM



57 Vitaliy Fadeyev  
LBNL

# Noise Study

- Spectra of noise only segment at ~8 seconds
- Top (optical) bottom (stylus) versions
- Mid-range level is higher in optical sample



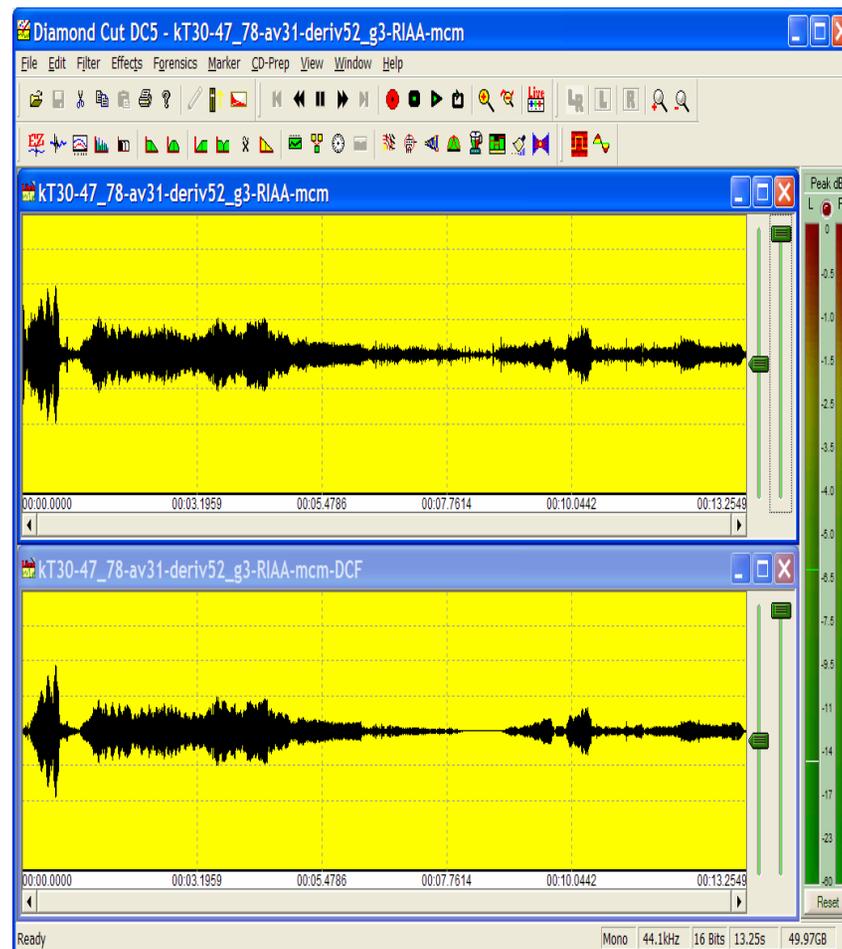
# Effect of Classic Noise Reduction

- Option to use commercial continuous noise filtering software on optical sample.
- Derive noise spectrum from quiet portion.
- Apply to full sound clip.
- Result

– Before

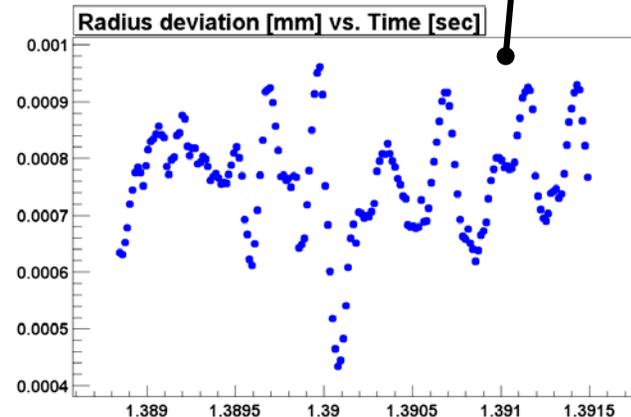
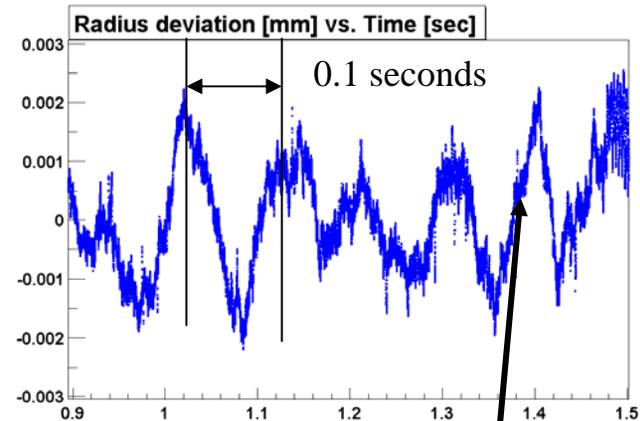


– After

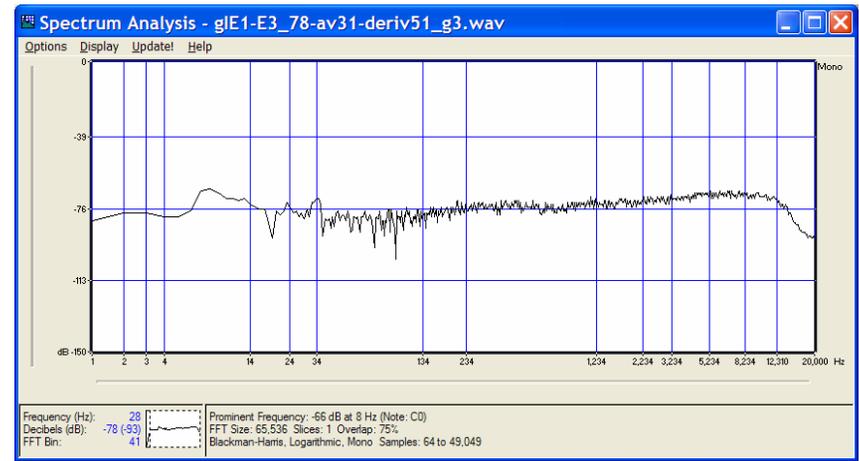
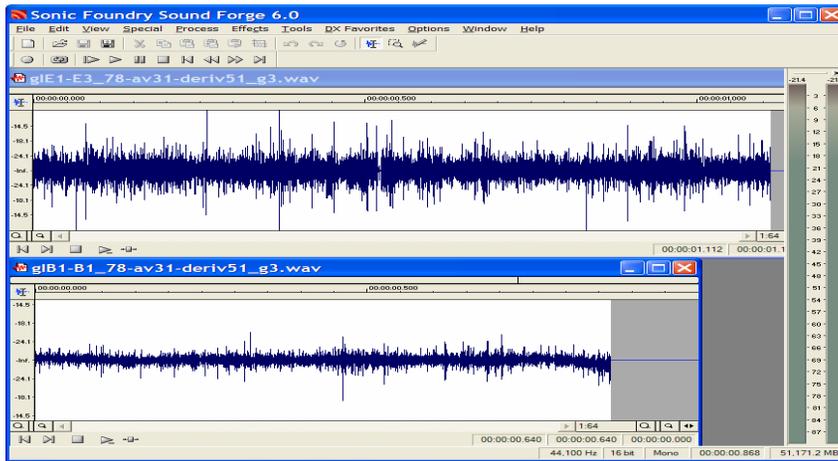


# Physical Origin of Noise in Optical Reconstruction

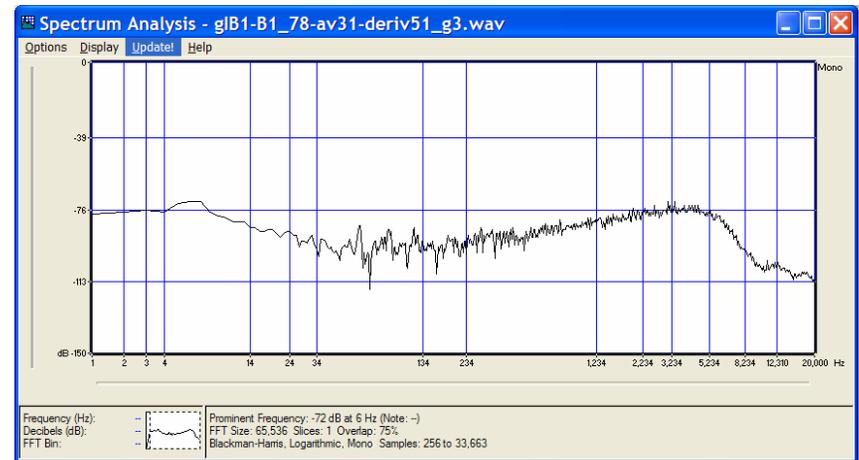
- View of raw groove shape data from region of pause, before differentiation into velocities.
- Upper plot is 0.6 second portion.
- Lower plot shows deviations about 10 Hz waveform.
- Each point is an independent edge detection across the groove bottom.
- Clear structures, spanning multiple points are resolved of typical scale:  
100 microns (0.2 ms) x 0.2 microns !!!



# Measurement of Noise at $R_{\min}$ & $R_{\max}$



Optical readings  
Upper sample is at outer radius  
Lower sample is at inner radius  
From “Goodnight Irene” disk  
If noise is dominated by surface structures  
of constant size distribution the outer radius  
amplitude and frequency should be greater  
due to greater linear speed there



# Conclusions

- Image based methods have sufficient resolution to reconstruct audio data from mechanical media.
- Method readily reduces impulse (click & pop) noise.
- Wide band “hiss” is present in 2D optical reconstruction.
  - Origin is not known definitively.
  - Insights into physical basis of noise. Observed structure may suggest ways to reduce further. More study required.

# Conclusions

- Basic process is data intensive compared to simple stylus playback.
  - 2D approach may be suitable for mass digitization.
  - How general is the 2D image quality?
  - At present 3D methods may be suitable for reconstruction of particular samples since they require ~hours per scan. Point scan is more flexible than frame based approach.
- Future potential in surface profiling field for full 3D reconstruction. This would be a good area for further research and collaboration. Optimizations.
- Report available [LBNL-51983](#), submitted to JAES
- Info at URL [www-cdf.lbl.gov/~av](http://www-cdf.lbl.gov/~av)

# Future Directions

- Report has been circulated in Audio Preservation community and well received.
- Presentations at Library of Congress, NY Public Library, and Stanford archives.
- Serious interest expressed in the further development of this work for both mass digitization and recovery of delicate media by the Library of Congress.

<b>Parameter</b>	<b>78 r.p.m., 10 inch</b>	<b>33 1/3 r.p.m., 12 inch</b>
Groove width at top	150-200 $\mu\text{m}$	25-75 $\mu\text{m}$
Grooves/inch (mm) $G_d$	96-136 (3.78-5.35)	200-300 (7.87-11.81)
Groove spacing	175-250 $\mu\text{m}$	84-125 $\mu\text{m}$
Reference level peak velocity@1KHz	7 cm/sec	7 cm/sec (0.0011 cm)
Maximum groove amplitude	100-125 $\mu\text{m}$	38-50 $\mu\text{m}$
Noise level below reference, S/N	17-37 dB	50 dB
Dynamic range	30-50 dB	56 dB
Groove max amplitude at noise level	1.6 - 0.16 $\mu\text{m}$	0.035 $\mu\text{m}$
Maximum/Minimum radii	120.65/47.63 mm	146.05/60.33 mm
Area containing audio data	38600 mm <sup>2</sup>	55650 mm <sup>2</sup>
Total length of groove	152 meters	437 meters