PROC SPIE 1531-2

Comparison of optical and mechanical measurements of surface finish

E. L. Church

USA ARDEC, Picatinny NJ 07801-5000

J. C. Dainty and D. M. Gale

Imperial College, London SW7 2BZ

P. Z. Takacs

Brookhaven National Laboratory, Upton NY 11973-5000

ABSTRACT

The topography of a smooth machined silicon flat has been measured with a phase-shifting Linnik microscope and a Talystep mechanical profilometer and the results compared in the frequency domain. Excellent agreement is found for the strong low-frequency components that determine the gross properties of the surface. Differences observed for some small high-frequency components may have implications for the understanding of microscope imaging properties.

1. INTRODUCTION

In 1988 we reported the comparison of optical and mechanical measurements of a smooth test surface using a Linnik microscope and a Talystep profilometer [1]. The present paper confirms and extends that earlier work. Although these studies are stated in terms of the comparison of the finish parameters of a particular surface measured by particular instruments, they invoke fundamental issues that underlie the general field of surface metrology.

2. LINEAR MEASURING SYSTEMS

Any practical measuring system must be linear over its operating range. An operation L is linear when

$$L(aZ_1 + bZ_2) = aL(Z_1) + bL(Z_2) \tag{1}$$

where a and b are constants. From this and shift invariance it follows that

$$L(x) = P(x) * Z(x) \tag{2}$$

where P(x) is the impulse response of L and the star denotes convolution. In the frequency domain Eq 2 is

$$L(f) = P(f) \cdot Z(f) \tag{3}$$

Where P(f) is the transfer function of L and the dot denotes an ordinary product.

P(f) depends only on the measurement system and Z(f) depends only on the object being measured. An elegant way of stating this result is to say that the eigenfunctions of a linear system are sinusoids, and P(f) is their eigenspectrum. By convention, P(f) is normalized so that P(0) = 1 through the choice of the instrumental 'calibration constant'.

If measurement systems were not linear we could not use Eq 1 to separate signal from noise by subtraction, and could not use Eq 3 to separate the signal from the effects of the measurement process by division. We consider only linear systems hereafter.

A 'perfect' measurement system has P(f) = 1 for all f . An 'ideal' system has P(f) = 1 over its operating range, while 'real' systems have 0 < P(f) < 1. A real system, however, can be converted into an ideal system using a-priori knowledge of its transfer function.

This is done by first transforming the measured data into the frequency domain, dividing both sides of Eq 3 by the known form of P(f) to get Z(f), and transforming back to configuration space to get the restored form of Z(x). This restoration process has been used extensively in the literature to compensate measured data for the attenuations inherent in optical measurements [1,2,3].

In practice, P(f) comes from different sources. Data analysis, for example, sets a lower-wavelength limit equal to the Nyquist wavelength, 2D, where D is the sampling interval, and an upper limit of ND , where N is the number of data points in the profile. (Note that ND is $\approx$ the trace length, (N-1)D , for large N.) Additional and possibly more restrictive constraints come from anti-aliasing mechanisms and the measurement process itself.

Since a linear measurement is completely characterized by its transfer function, the details of the test surface used in the comparison of two techniques is irrelevant as long as it generates a signal with sufficiently high signal-to noise ratios in each. In this paper, however, we state the comparison in terms of finish parameters of our particular test surface, since they may have interest in their own right, and finish parameters are the lingua franca of metrology.

3. TEST SURFACE

The test surface used in the present study was the same fly-cut silicon flat used in our earlier work [1,2]. Its profiles are dominated by well-defined tool marks with a period of about 3.4 micrometers (μm) and a root-mean-square (rms) roughness of about 150 Angstroms (1 Angstrom = 0.1 nanometer).

Figures 1 and 2

..... show 20-μm segments of the mechanical and optical profiles of this surface, with its curious double-peaked tool marks [4]. The optical trace is smoother than the mechanical one, which reflects the greater high-frequency attenuation of the optical system, and there is some variation between individual tool marks. On the whole, however, the surface is remarkably uniform, as shown in the microphotograph in Fig 2 in Ref 2 (not reproduced here).

We have chosen this surface because of its durability, the fact that it is an 'optical-quality' surface, that its features fall within the operating range of the instruments being tested, and the fact that its finish is 'one dimensional'. The last permits the use of conventional chisel-shaped styli, and allows Eq 3 to be used in its simple one-dimensional form.

A we shall see, the spectrum of the topography of this surface is dominated by the fundamental and first half-dozen harmonics of the tool marks. The nominal range of surface wavelengths of interest then lies between, say, one-tenth to ten times its fundamental period; that is, between 0.34 and 34 μm.

4. MECHANICAL MEASUREMENTS

The first set of mechanical measurements were made by Dr. T. V. Vorburger using the digitized Talystep at the National Bureau of Standards (now the National Institute of Standards and Technology (NIST)) in Gaithersburg MD [1]. This consisted of a series of 80-μm long traces each containing 4000 points with a sampling interval of D = 0.02 μm, taken at different locations on the surface. Individual traces were divided into five partially overlapping data sets of 1024 points each. A total of thirty such 1024-point data sets were used in the analysis.

The second set of measurements were made by Dr. J. M. Bennett using the digitized Talystep at the Royal Institute of Technology (RIT) in Stockholm, Sweden. This consisted of a series of traces between 15 and 175 μm in length with a sampling interval of D = 0.03408 μm. A total of ten 1024-point data sets were taken from these and used in the analysis described below.

The Nyquist and trace-length limits on P(f) are 0.04 and 80 μm for the NIST data and 0.07 μm and 35 μm for the RIT data. These span the full range of wavelengths important in the present test surface.

Additional constraints on the short-wavelength limit of P(f) come from the mechanical filtering action of the stylus assembly, possible electronic anti-aliasing filters, and the filtering action of the non-vanishing size of the stylus tip [1,5].

The short-wavelength cut-off of the mechanical and electronic filtering action of the NIST instrument was measured directly and found to be $\sim$ 2/40 = 0.05 μm, which is comparable with the Nyquist wavelength mentioned above [6]. The nominal radii of the stylus tips used in the NIST and RIT measurements was $\sim$ 0.1 - 0.2 μm, which is negligible relative to the rms radius of curvature of the present test surface: $\sim$ 2.8 μm [1].

Based on these considerations we have taken the transfer function of the Talystep measurements to be

$$P(f) = 1 \qquad (4)$$

over the range of surface wavelengths of importance in the present test surface.

5. OPTICAL MEASUREMENTS

The optical measurements were made using the research Linnik microscope at Imperial College, London, which was designed for the measurement of sub-micrometer integrated-circuit structures [7].

Figures 3 and 4

..... show the layout and a schematic of the current version of this instrument.

It uses a matched pair of 100 X microscope objectives with NA = 0.94 (0.90) and operates at the wavelength λ = 0.6328 (0.555) μm. Brightfield (Kŏhler) illumination is used with a coherence parameter of S = 0.29 (0.30), with polarizers before and after the sample oriented so that the E vector is parallel to the tool marks. (Values in parentheses refer to the version of the instrument used in the 1988 studies [1].)

Two types of detection were used. The 1988 data and batch A of the current data were recorded using a 384 (x 512) pixel CCD camera with the test surface stationary. Each measurement, averaged over 25 frames to reduce electronic noise, gave a 24.27-μm trace of the surface containing 364 points D = 1/15 = 0.0667 μm apart. A standard piezo phase-shifting scheme was used with a four (three) bucket analysis. Measured phase fluctuations were converted to height by multiplying by $\lambda/4{\cdot}pi$.

Batch B was recorded using a photomultiplier with a pinhole aperture as the test surface was scanned by a mechanical stage. Each scan, averaged over 16 frames, gave a 9.80-μm trace containing 500 points D = 1/51 = 0.0196 μm apart. The diameter of the detector pinhole projected on the sample was 2w = 0.23 μm.

Figure 5

..... shows a typical profile measured in this way, on the same scales as Figs 1 and 2.

The Nyquist and trace-length limits on P(f) are 0.133 and 24.27 μm using the CCD detector, and 0.0392 and 9.80 μm using the photomultiplier. These also span the range of wavelengths of importance in the present test surface.

Additional limits on P(f) come from the optical system of the measuring microscope. Pether [8] has analyzed the microscope configuration used here and has concluded that for a sufficiently smooth surface P(f) is given by the triple product

$$P(f) = P_1(f) \cdot [P_2(f) * P_3(f)] \qquad (5)$$

where

$$P_1(f) = \mathrm{Sin}(pi{\cdot}w{\cdot}f)/(pi{\cdot}w{\cdot}f) \qquad (6)$$

is the transfer function of the (rectangular) detector aperture,

$$P_2(f) = \mathrm{Cyl}(\lambda f/NA) \tag{7}$$

is that of the microscope objective, and

$$P_3(f) = \mathrm{Cyl}(\lambda f/S \cdot NA) \tag{8}$$

is that of the illumination. Here Cyl(x) = 1 for x < 1 and = 0 otherwise.

Over the range of surface frequencies of importance here, Eq 6 is approximately unity and P(f) is the convolution of the two cylinders given by Eqs 7 and 8. The analytic expression for this is given in the Appendix.

Figure 6

..... is a sketch of this function, which involves three regions: I, II and III. In region I,

$$0 < f < (1 - S) \cdot NA / \lambda \ , \tag{9}$$

P(f) = 1, and frequency components would be measured without attenuation. In region II,

$$(1 - S) \cdot NA / \lambda < f < (1 + S) \cdot NA / \lambda \ , \tag{10}$$

the frequency components are attenuated by factors between 1 on the low end and 0 on the high end. In region III,

$$(1 + S) \cdot NA / \lambda < f < \infty \ , \tag{11}$$

no signal should be measured since it lies above the optical cut-off of the microscope in the present approximation.

The limits in Eq 10 correspond to spatial wavelengths of 0.522 (0.474) and 0.948 (0.881) μm for the current (earlier) version of the Linnik microscope used here. As we shall see, these fall in the middle of the range of wavelengths of importance for the present test surface, which permits us to investigate its imaging properties in all three regions.

6. QUALITATIVE COMPARISONS

Armed with the above information it is possible to make a direct comparison between mechanical and restored optical profiles of our test surface. However, because of difficulties in aligning the surface to sub-micron accuracy it is more meaningful to compare statistics of a set of profiles taken over the surface. The principal statistic of interest for optical surfaces is the power spectral density of their surface-height fluctuations, since this determines its scattering properties.

We estimate the (one-sided) power spectral density, S(f), of the surface from the measured profiles, Z(x), using the well-known periodogram estimator:

$$\hat{S}(f_m) = A_v \left\{ \frac{2D}{N} \left| \sum_{n=0}^{N-1} e^{i 2\pi m n / N} Z(nD)\, W(n) \right|^2 K(n) \right\} \qquad (12)$$

where D is the sample spacing, N is the total number of sampled data points in the profile, Z(nD) is the detrended profile, W(n) is a data window and K(n) is a bookkeeping factor which equals 1 except for K(0) = K(N/2) = 1/2, and 'Av' stands for the arithmetic average over different profiles [9,10]. The surface frequency is given by

$$f_m = \frac{m}{ND} \quad , \qquad m = 1\ (1)\ N/2 \qquad (13)$$

Individual profiles were detrended by subtracting a quadratic polynomial determined by ordinary-least-squares fitting to the raw data. A Blackman (Hamming) data window was then applied, and, for the batch-B data, the sets were symmetrically padded with zeros to give N = 1024 points before substitution into Eq 12. Log-10 of the resulting 512 values of S(f) were then plotted versus f on a linear scale, with data points connected by straight lines.

Figures 7 and 8

..... show the power spectra of the Talystep measurements of the test surface. The first is the average of 30 1024-point NIST profiles, and the second is that of 10 1024-point RIT profiles.

They are very similar. Nine lines are visible in each and the spectra decay to almost identical backgrounds at high frequencies. The RIT lines, however, appear to be sharper and taller than the NIST lines, and lines 5, 6, 7 and 9 are distinctly more intense. The cause of these differences is unknown. One possibility is that the high-frequency detail of the profiles is not uniform over the surface and the two measurements sample that non-uniformity differently.

Figures 9 and 10

..... show the raw (unrestored) spectra obtained from the Linnik measurements using the CCD detector. The first is from the 1988 paper and the second is based on the 1991 measurements, which involve different system parameters. The vertical lines divide the spectrum into the three regions defined by Eq 10.

In region I, where P(f) is unity according to Eq 9, the spectra are indistinguishable. In region II, where P(f) falls between 1 and 0, the new measurements fall below the earlier ones because of the shift in system parameters. In region III, above the optical cut-off, both spectra decay to approximately the same noise level ($\sim$ twice phase quantization noise).

The unexpected feature is the apparent presence of spectral lines above the cut-off: lines 8 and 9 in the earlier data and lines 7 and possibly 9 in the current data.

Figure 11

..... shows the average of 48 raw spectra obtained from recent Linnik measurements using the photomultiplier detector. As expected, the spectral lines are broader than those in Figs 9 and 10 because of the shorter trace length in this case (9.80 vs 24.27 μm).

Not only do these data show evidence for line 7 above the optical cut-off, but lines 3 and 5 appear to be attenuated relative to their neighbors.

7. QUANTITATIVE COMPARISONS

Quantitative values of various finish parameters can be obtained from the power spectra shown in the earlier figures. Mean-square roughness values are obtained by integrating the spectra over appropriate bandwidths, and the tool-mark period follows from the positions of the spectral lines.

The table below compares parameters derived from the NIST and RIT Talystep data:

Parameter	NIST	RIT
a_1 (A)	103	115
a_2 (A)	166	193
σ (A)	145	170
d (μm)	3.47	3.50

the a 's are the amplitudes of the first two lines in the spectrum of the tool marks, sigma is the rms surface roughness integrated up to f = 3.0 μm (-1), and d is the period of the tool marks.

The RIT height values shown are $\sim$ 10-15% greater than the NIST values, which might suggest a difference in calibration.

The table below compares parameters from the earlier and current Linnik measurements, using the CCD (A) and photomultiplier-pinhole (B) detectors.

Although there are differences beteeen the height values, they generally overlap the range of the talystep data in the preceding table [11]. In general, the photomultiplier data are larger than the CCD data, but are subject to greater uncertainties because of the overlapping lineshaps as shown in Fig 11.

Parameter	Earlier (A)	Current (A)	Current (B)
a_1 (A)	101	101	113
a_2 (A)	141	156	170
σ (A)	127	137	152
d (μm)	3.33	3.41	(3.65)

The next table compares the Talystep and the current Linnik data in more detail:

Parameter	Talystep	A	B	A/P(f)	B/P(f)
a_1 (A)	98 - 115	101	113	101	113
a_2 (A)	164 - 193	156	170	156	170
a_3 (A)	35 -- 41	28	16	28	16
a_4 (A)	21 -- 25	11	26	12	28
a_5 (A)	26 -- 30	4	3	7	5
a_6 (A)	13 -- 15	3	2	19	12
a_7 (A)	10 -- 12	0.5	1	0	0

The Talystep column gives the range of values determined by taking 0.85 and 1.00 times the RIT measurements. This 15% spread is taken to be a measure of the possible uncertainty in the NIST and RIT data. The next two columns give the raw values of the parameters derived from the current Linnik measurements, and the last two give these results restored by dividing by P(f) calculated from Eq (5) at the appropriate frequencies.

Lines 1, 2 and 3 fall in region I and are seen to be in good agreement across the board. Lines 4, 5 and 6 fall in region II. Line 4 lies at low enough frequency that it is not attenuated appreciably, but the raw values of lines 5 and 6 are, as expected, much smaller than the Talystep values. The restored value of line 5 falls well below its Talystep value, while the restored value of line 6 appears to be in good agreement, although this is probably fortuitous.

The conspicuous attenuation of lines 3 and 5 in case B relative to case A shown in Fig 11 is reflected in the finish parameters. The B data also indicate the surface signal above the optical cut-off of the microscope (line 7).

8. NON LINEAR EFFECTS

One possible explanation of the presence of spectral lines above the optical cut-off is the existence of significant non-linear effects in the imaging system [1]. According to Pether [8], the measured and true profiles is given by the non-linear relationship:

$$\underset{MEAS}{Z(X)} = \frac{\lambda}{4\pi} ARG\left\{ P(X) * EXP\left[i \frac{4\pi}{\lambda} \underset{TRUE}{Z(X)} \right] \right\} \tag{14}$$

when the reflecting surface is a pure phase object.

When the surface is smooth enough, the exponential can be replaced by its first two terms, and Eq 14 reduces to the linear form

$$\underset{MEAS}{Z(X)} = P(X) * \underset{TRUE}{Z(X)} \tag{15}$$

where P(x) is the Fourier transform of Eq 5. In this case the system will not propagate any frequencies above the optical cut-off given by Eq 11. However, if the surface is so rough that the system operates in its non-linear regime, the cross modulation of the low-frequency components can generate signals above this cut-off.

The expansion parameter in Eq 14 is of the order

$$\frac{4\pi}{\lambda} \sigma \sim 0.3 \tag{16}$$

for the present test surface, which is not obviously small or large. To determine non-linear effects in this case we have evaluated the spectrum of Eq 14 when the 'true' profile is a typical RIT Talystep trace.

Figure 12

..... shows the results. The shaded curve on the top is the spectrum of the input Talystep trace, the solid line is the spectrum of the measured profile assuming the linearized transfer function in Eq 15, and the shaded curve on the bottom is the spectrum of the measured profile given by Eq 14.

The bottom curve does show the presence of signal above the linear cut-off, but at too low an intensity to account for such lines observed in Figs 9, 10 and 11.

9. SUMMARY

All in all the agreement between the optical and mechanical measurements appears to be excellent. This supports the physical idea that the surface being measured can be modelled as a single sharp and rigid interface between two homogeneous media.

As far as differences go, there appears to be a systematic 10 - 15% displacement between the two sets of mechanical measurements which could be due to calibration uncertainties. In the optical case the linearized phase-modulation transfer function does not appear to account quantitatively for small signals near the optical cut-off.

We have more data than have been analyzed here: more of the recent Talystep data, and Linnik measurements made in a confocal configuration and under other conditions. Future studies will include these in comparisons similar to those reported here, plus an determination of the homogeneity of the finish of the present test surface by an examination of the fluctuations of the shapes and parameters of individual profiles.

10. ACKNOWLEDGEMENTS

We thank J. M. Bennett for making the Talystep measurements of our test surface, and M. I. Pether for discussions of the imaging properties of the Linnik microscope.

11. DISCLAIMER

Certain items of commercial equipment have been mentioned in the text to specify experimental procedure, but this does not denote recognition or endorsement by the US Department of Defense or the US Department of Energy.

12. APPENDIX

The convolution of two unit cylinders, Cyl(Q) and Cyl(Q/S), normalized to unity at Q = 0 , is

$$P(Q) = \begin{cases} 1 & \text{for} \quad 0 < Q < 1 - S \\ 1 + C & \text{for} \quad 1 - S < Q < SQR(1 - S*S) \\ 0 + C & \text{for} \quad SQR(1 - S*S) < Q < 1 + S \\ 0 & \text{for} \quad 1 + S < Q < \end{cases} \tag{17}$$

where

$$C = (ATN(B/A) + S*S*ATN(B/(Q-A)) - B*Q)/(pi*S*S)$$

$$B = SQR(1-A*A) \tag{18}$$

and

$$A = (1+Q*Q-S*S)/(2*Q) \qquad (19)$$

When S = 0 and 1 these expressions reduce the the textbook forms of the coherent and incoherent transfer functions of a simple circular lens, as expected.

The one-dimensional version of P(Q) is obtained by convolving Rect(Q) and Rect(Q/S). In that case the second and third lines in Eq 17 are replaced by the single expression

$$P(Q) = (1 - Q + S)/(2*S) \qquad (20)$$

The configuration-space forms of these results, P(x), are obtained by taking the Fourier transforms of P(f). Because of the normalization P(f=0) = 1 , P(x) has unit area. In two dimensions P(x) is the product of two Jinc functions, and in one dimension, the product of two Sinc functions.

In the present application $Q = \lambda f/NA$.

13. REFERENCES AND FOOTNOTES

1. E. L. Church, J. C. Dainty, D. M. Gale and P. Z. Takacs, "Comparison of optical and mechanical measurements of surface finish", Proc SPIE 954 189-199 (1988).
2. E. L. Church, T. V. Vorburger and J. C. Wyant, "Direct comparison of mechanical and optical measurements of the finish of precision machined optical surfaces", Optical Engineering 24 388-395 (1985).
3. E. L. Church and P. Z. Takacs, "Effects of the optical transfer function on surface finish measurement", Proc SPIE 1164 46-59 (1989).
4. The unusual tool marks may result from a dynamic non-linear mechanical interaction between the diamond tool tip and the silicon surface during the machining process.
5. E. L. Church and P. Z. Takacs, "Effects of the non vanishing tip size in mechanical profile measurements", Proc SPIE 1332 504-514 (1990).
6. E. L. Church, M. R. Howells and T. V. Vorburger, "Spectral analysis of the finish of diamond-turned mirror surfaces", Proc SPIE 315 202-218 (1982).
7. D. M. Gale, M. I. Pether and F. C. Reavell, "Interference microscopy of surface relief structures", Proc SPIE 811 40-47 (1987).
8. M. I. Pether, "Image formation in the Linnik microscope", privately circulated notes, 1988.
9. E. L. Church and P. Z. Takacs, "Instrumental effects in surface finish measurements", Proc SPIE 1009 46-55 (1988).
10. E. L. Church and P. Z. Takacs, "The optimal estimation of finish parameters", Proc SPIE 1150 (1991), in publication.
11. The value of 141 Angstroms in Table I of Ref 1 quoted here appears to be in error. The identity of the earlier and current Linnik spectra in region I indicate that the correct value is approximately 156 Angstroms.

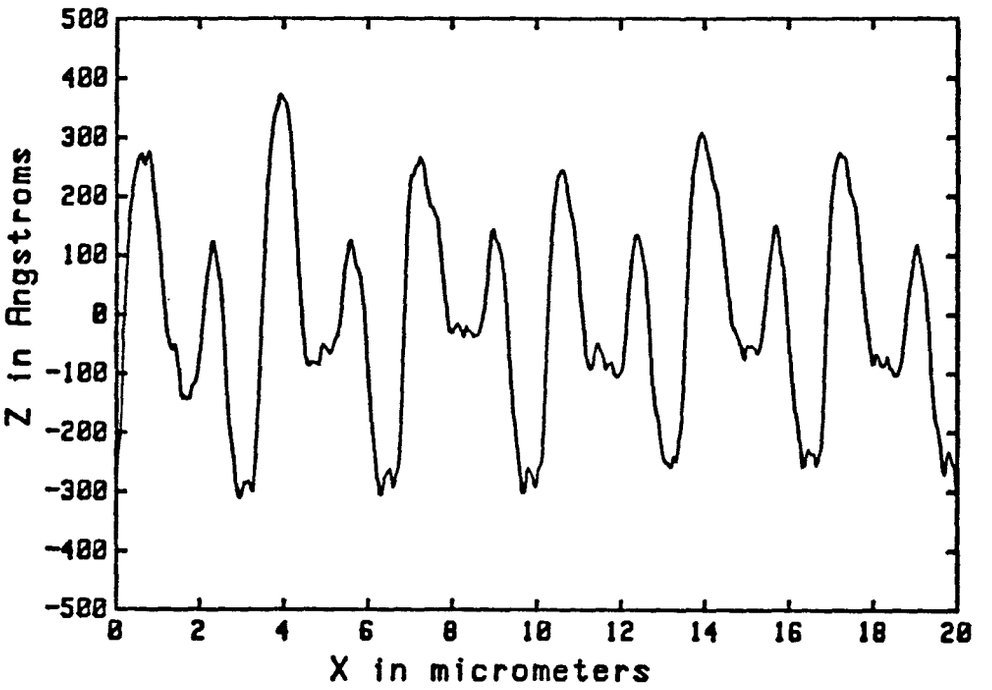

Figure 1. A 20-μm segment of an RIT Talystep profile of the silicon test surface.

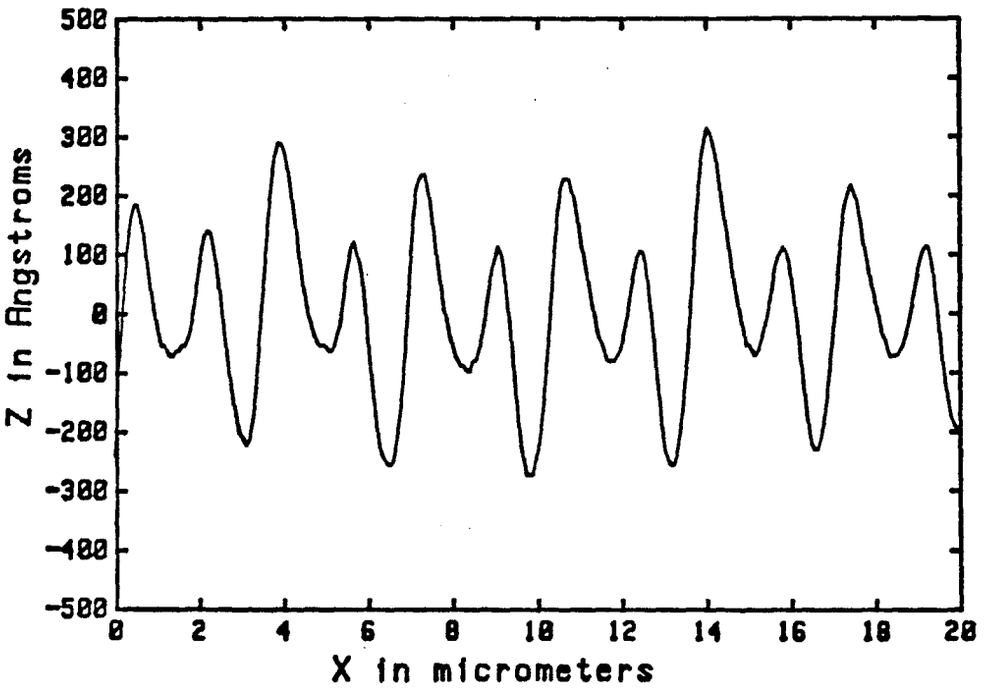

Figure 2. A 20-mu segment of a Linnik profile using the CCD detector in the 1991 configuration.

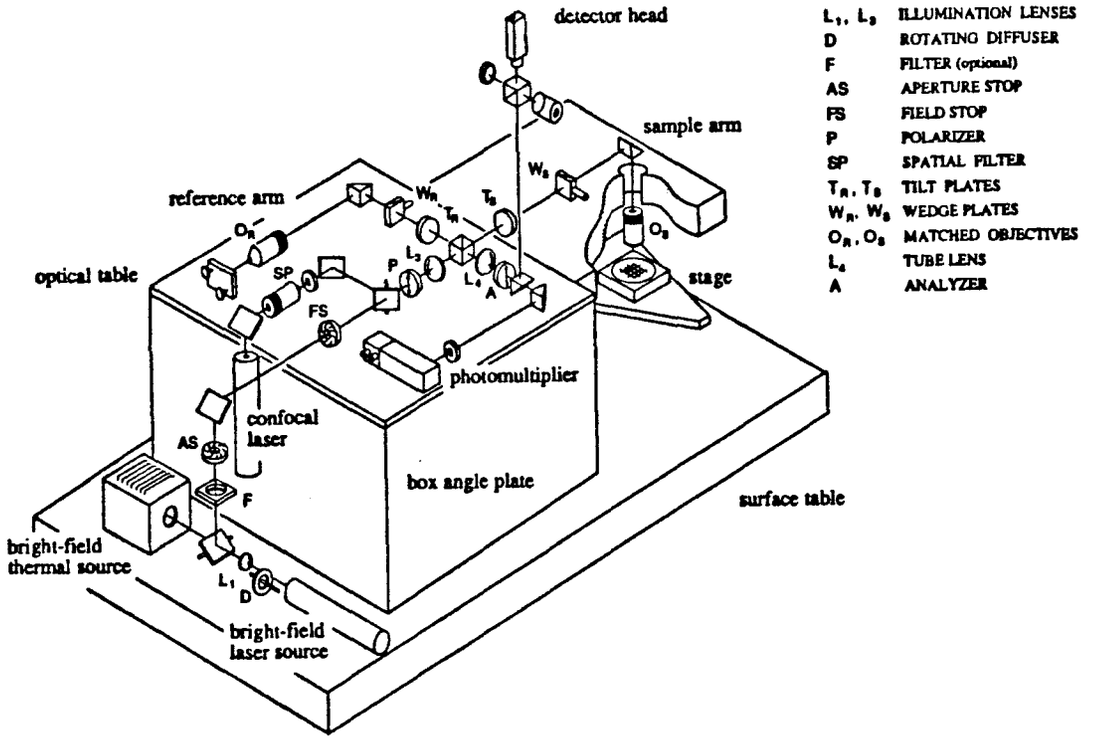

Figure 3. Layout of the current Linnik microscope.

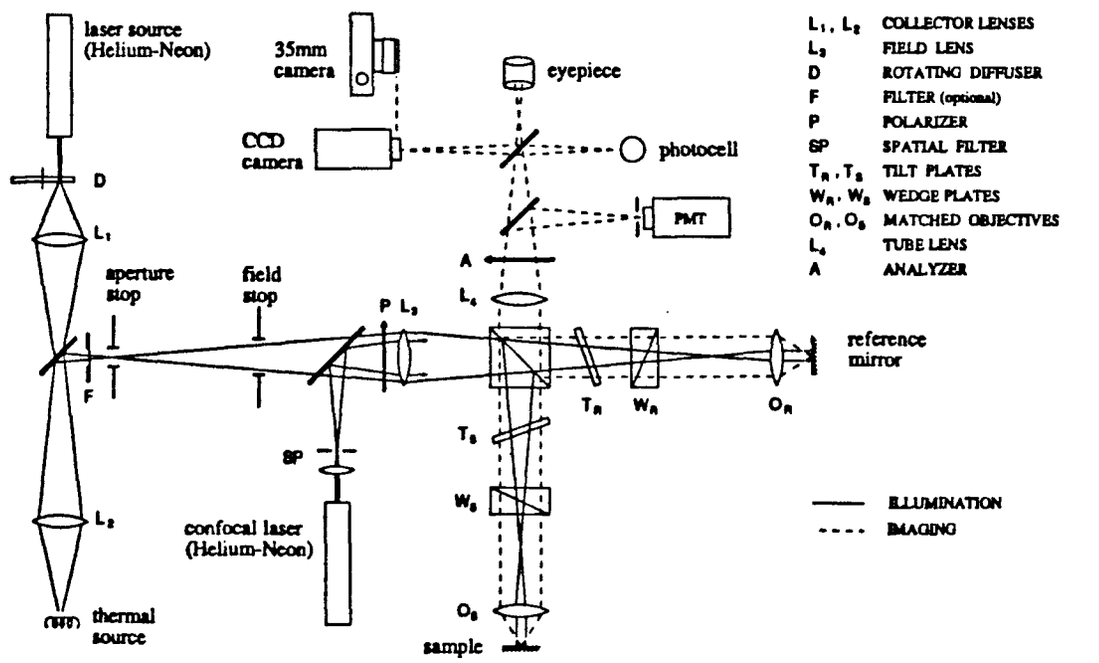

Figure 4. Schematic of the Linnik microscope.

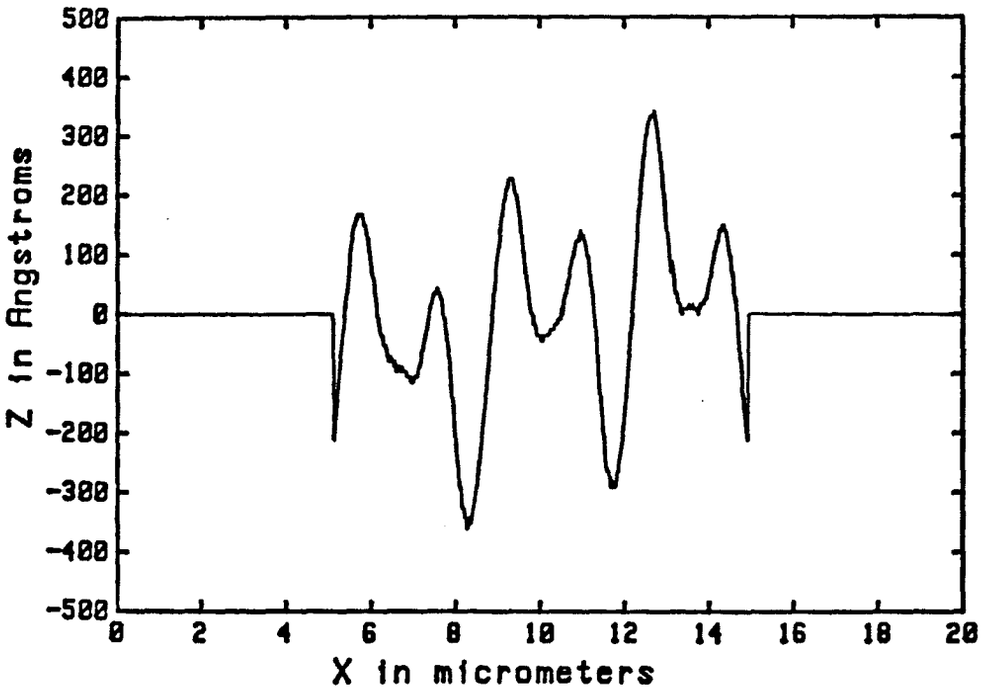

Figure 5. A Linnik profile obtained using the fixed photomultiplier detector with a pinhole aperture and scanning the test surface.

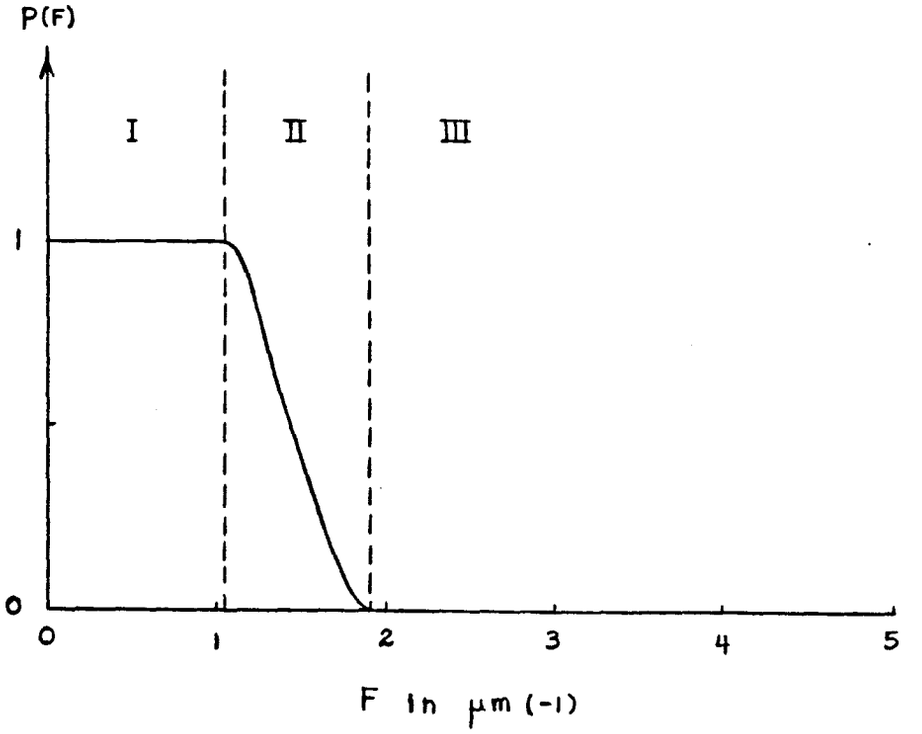

Figure 6. Sketch of the transfer function of a linear microscope with partially-coherent illumination. The dashed lines indicate the frequency limits in Eq 10.

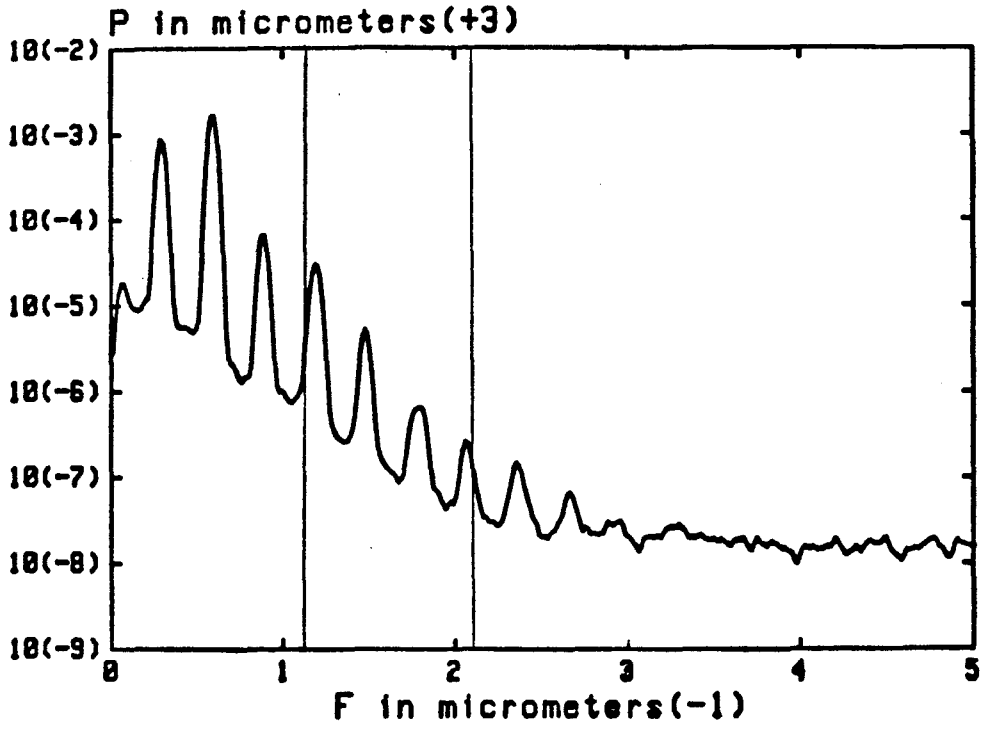

Figure 9. Power spectral density obtained by averaging the results of 50 Linnik profiles using the CCD camera in the 1988 configuration.

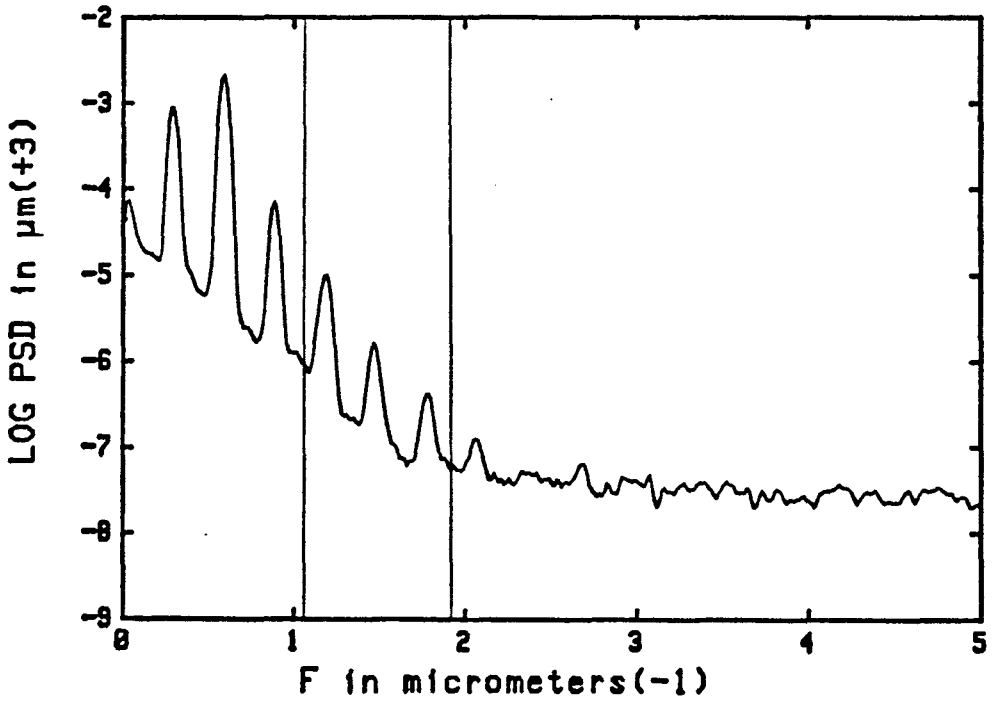

Figure 10. The same as Fig 9 but in the 1991 configuration.

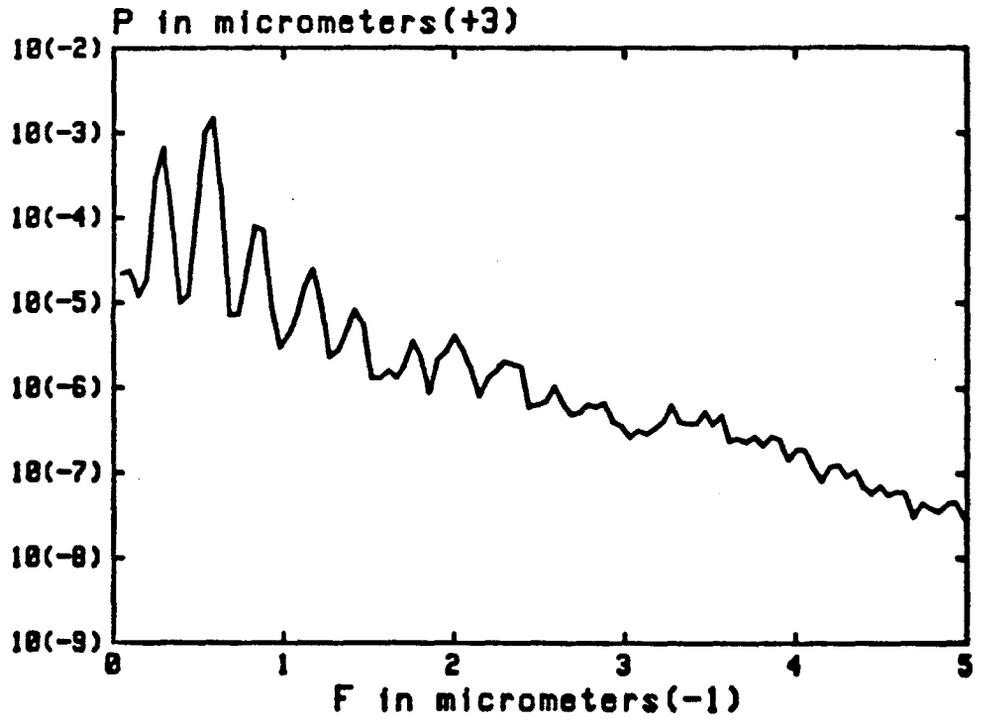

Figure 7. Power spectral density obtained by averaging the results of 30 NIST Talystep profiles.

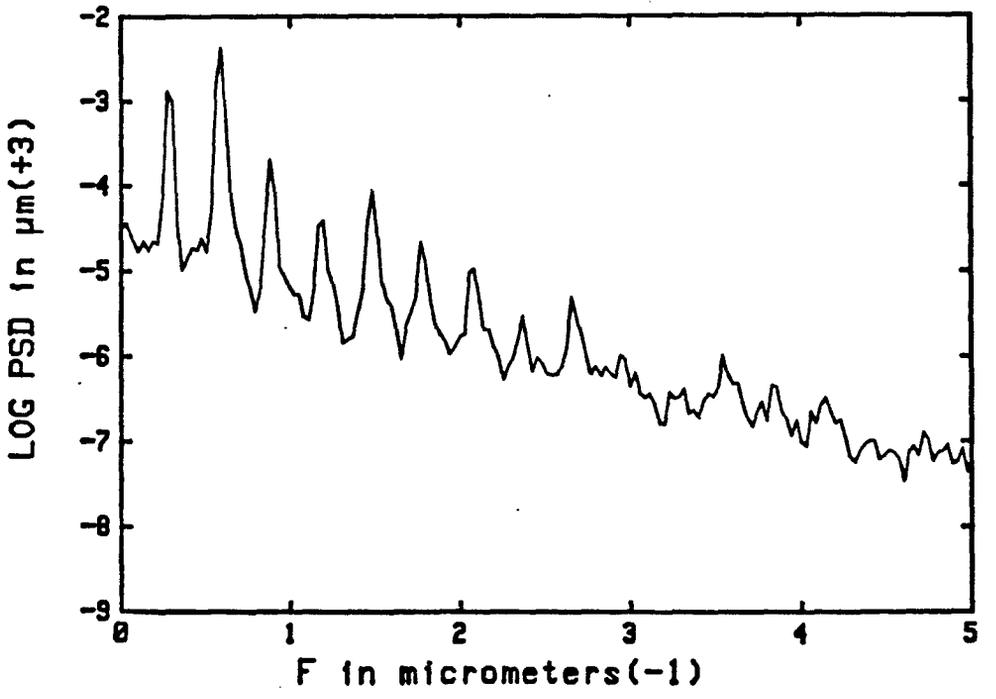

Figure 8. Power spectral density obtained by averaging the results of 10 RIT Talystep profiles.

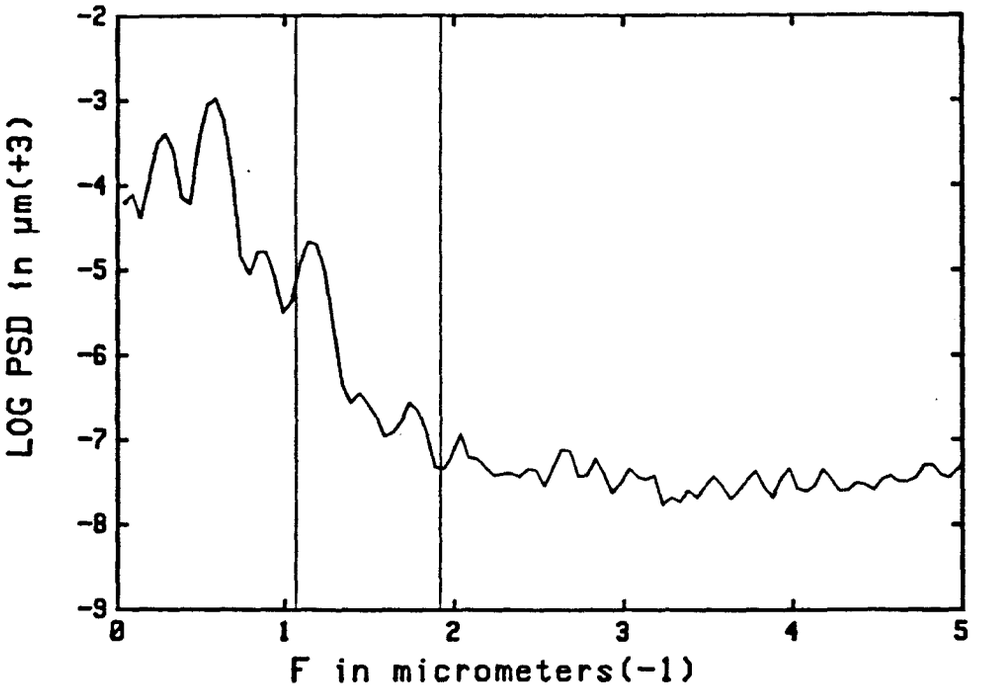

Figure 11. Power spectrum obtained by averaging the results of 48 Linnik profiles using the fixed photomultiplier detector with a pinhole aperture and scanning the test surface.

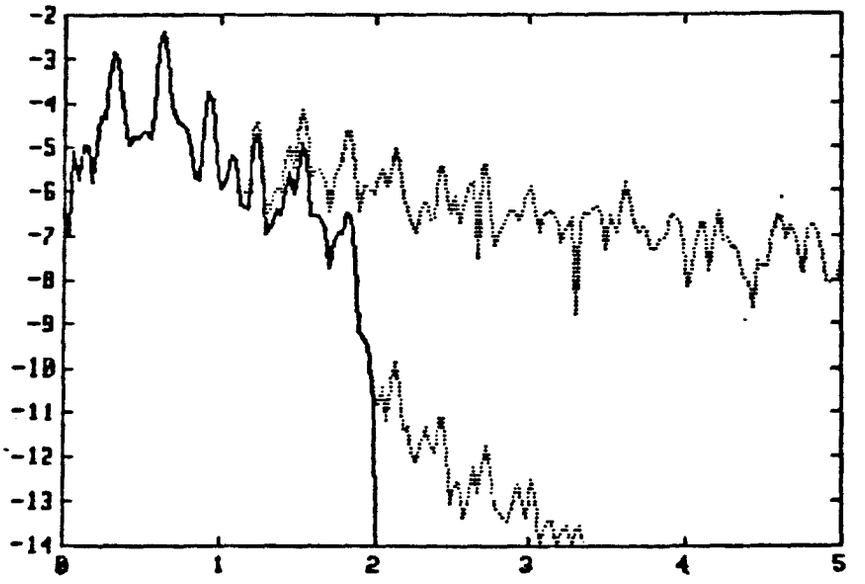

Figure 12. Spectra of a single RIT profile passed through different filters. Top line: unit filter. Solid line: linear microscope (Eq 15). Bottom line: non-linear microscope (Eq 14).