Development of a smooth-surface microroughness standard

Bradley W. Scheer [a] and John C. Stover [b]

[a] VLSI Standards, Incorporated, 3087 North First Street, San Jose, CA 95134
[b] ADE Optical Systems Corp., 9625 Southern Pine Blvd., Charlotte, NC 28273

ABSTRACT

Currently, there are several techniques available for measuring microroughness. However, the results tend to be qualitative. Until recently, there was no metrology standard available to correlate the accuracy of various instruments at extremely low levels of surface texture. This paper describes a metrology standard that is useful for calibrating instruments for the levels of microroughness encountered in semiconductor, disk drive, and related industries today. In advanced applications, this level is about 5 Å rms in a 0.01 - 1.0 μm^{-1} spatial bandwidth range. This standard uses a one-dimensional square wave to reduce the effects of instrument spatial bandwidth. The standard has a 20 μm pitch with feature depths as small as 10 Å. The overall theoretical design guidance for this standard has been described previously.[1]

Keywords: profilometry, PSD, rms, scatterometry, standards, roughness

1. INTRODUCTION

Microroughness is defined as "surface roughness components with spacings between irregularities (spatial wavelength) less than about 100 micrometers."[2] This definition differentiates microroughness from the larger scale surface variations of bow and warp.*

These very small levels of surface texture are becoming more problematical in a number of industries as the complexity of integrated circuits and the amount of information stored on disk drives increases. As an example, geometries in the integrated circuit industry are fast approaching molecular dimensions. The June 1994 report, the National Technology Roadmap for Semiconductors (NTRS), has published a requirement for gate oxide thicknesses approaching 4.5 nm ± 4%.[3] As a point of reference, the lattice constant for lightly doped (i.e., nearly pure) silicon is 0.543 nm. The gate dielectric molecule, silicon dioxide, is nominally 0.355 nm "diameter".[4,†] The ability of silicon dioxide or any film layer to function efficiently as an insulator depends partially on the underlying microroughness of the silicon surface. For oxides less than 10 nm, breakdown voltages are reduced commensurately with increased levels of microroughness. This can be readily understood by envisioning the "peaks" of the microroughness terrain as being much closer to the film surface than the overall average level of the peaks and valleys combined. Additionally, there are similar effects on film layers deposited in later processing steps, and an effect on bonding for silicon-on-insulator (SOI) applications.[5]

Currently, there are several techniques available for measuring microroughness. However, the results tend to be qualitative. Until recently, there was no metrology standard available to correlate the accuracy of various instruments.[‡] This becomes especially important when comparing instruments with differing spatial bandwidths, each possessing a unique transfer function. Due to the varying spatial bandwidths, different types of instruments can give rms microroughness values that differ by over an order of magnitude, even when measuring the same surface.[6]

The RQS (from the abbreviation for rms-roughness, R_q standard) uses a one-dimensional square wave to reduce the effects of bandwidth. Although the spatial bandwidth of the instrument is still important to know for the most accurate readings, the RQS makes use of the nature of the square wave to allow for fast, albeit less accurate, measurements. These concepts are discussed in detail in sections 2, *Instrument Considerations* and 4, *Determination of Specific Microroughness Values*.

---

* Bow and warp are terms associated with silicon wafer specifications. They are typically on a scale of millimeters of spatial wavelength.

† Obviously it is not possible to associate a diameter with an SiO_2 molecule; however, this number is based on a volume ratio which is 2.25 times that of a silicon atom. The linear dimension stated is taken as the cube root of this ratio.

‡ The term "standard" in this sense relates to a physical artifact that is used to verify the accuracy and precision of analytical test equipment. This is in contrast to the consensus, or "paper standard" which may specify a test procedure. VLSI Standards also manufactures an *isotropic Haze and Microroughness Standard* that may be used in conjunction with the RQS. This is especially useful when correlating surface texture with "haze", as reported by surface scanning inspection systems (SSIS).

The methodology used to produce this standard provides a known surface texture on a substrate with feature depths ranging from 1 nm to 10 nm.§ This is accomplished by precisely etching a square wave feature into silicon at known locations on a wafer (Figure 1). The imparted texture is fully quantifiable by angle resolved light scattering and measurements performed by atomic force microscopy. By having these data available and knowing the spatial bandwidth of a given instrument, it is possible to provide a direct method of quantifying a prescribed surface texture or microroughness value.

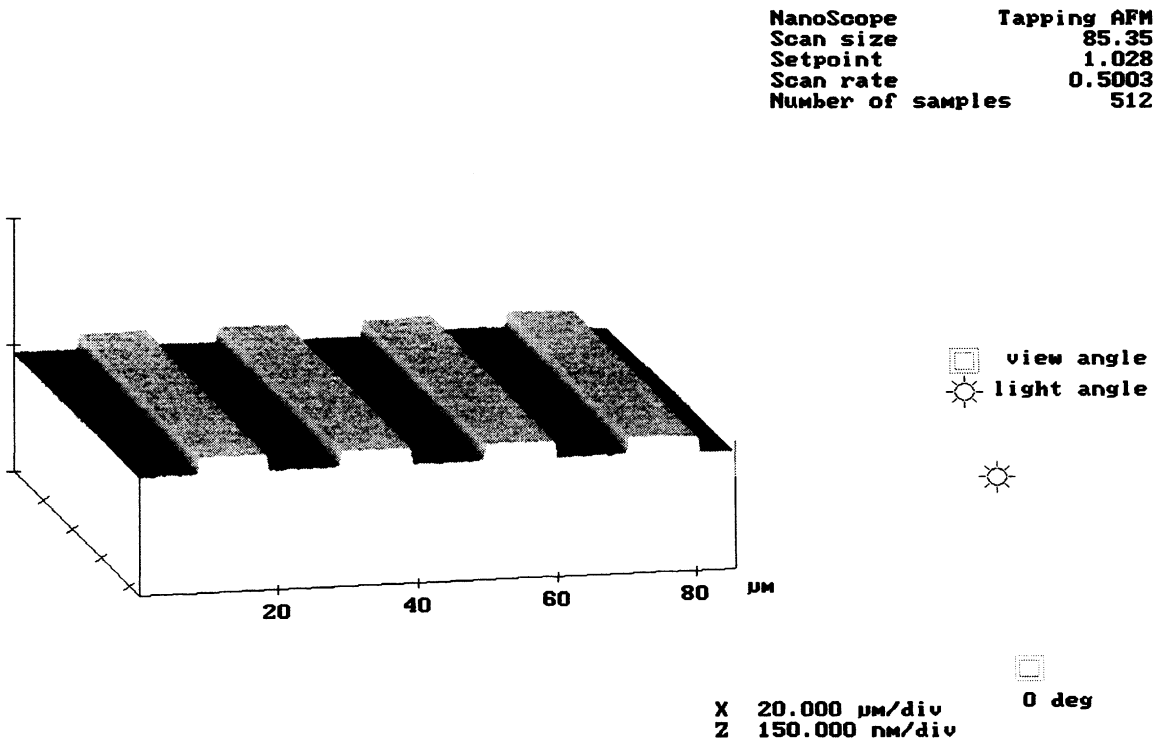


Figure 1. *Atomic force microscopy scan of RQS square wave structure (the period is 20 µm).*

In order to characterize surfaces, it is useful to calculate statistical properties from the measurement data obtained from various instruments. The two most widely used of these parameters are the centerline average roughness, R_a, and rms roughness, R_q (equation 1). Roughness values are calculated from the measured values of height variation, z_i, over a portion of a surface with a given number of sample points, N.[7]

$$R_a = \frac{1}{N}\sum_{i=1}^{N}|z_i| \qquad R_q = \sqrt{\frac{1}{N}\sum_{i=1}^{N}z_i^2} \tag{1}$$

However, these parameters do not uniquely define a given surface as surfaces with completely different profiles may have identical height averages, as shown in Figure 2. Roughness averages also depend on the spatial bandwidths over which they are taken. Surface texture is fractal in nature.

§ There is only one depth present on any given standard. The range indicated allows for various nominal values of microroughness based on a spatial bandwidth that includes at least the first peak in the PSD curve.

A measurement system that samples, say, every 10 micrometers of a surface would totally miss the high spatial frequency** variations on Surface I in Figure 2 and reduce its average roughness. This produces a microroughness value that is much lower than would be obtained had it sampled every 1 micrometer.

A more complete surface quantification that takes instrument spatial bandwidth into account is the power spectral density (PSD) curve. The PSD function is the frequency spectrum of the surface roughness measured in units of spatial frequency, typically inverse micrometers.[8] The PSD function provides information about both the amplitude and spatial wavelength ($1/f$) of the surface. From the PSD data, readings from various instruments may be correlated by incorporating their spatial bandwidths. This concept will be discussed in more detail later.

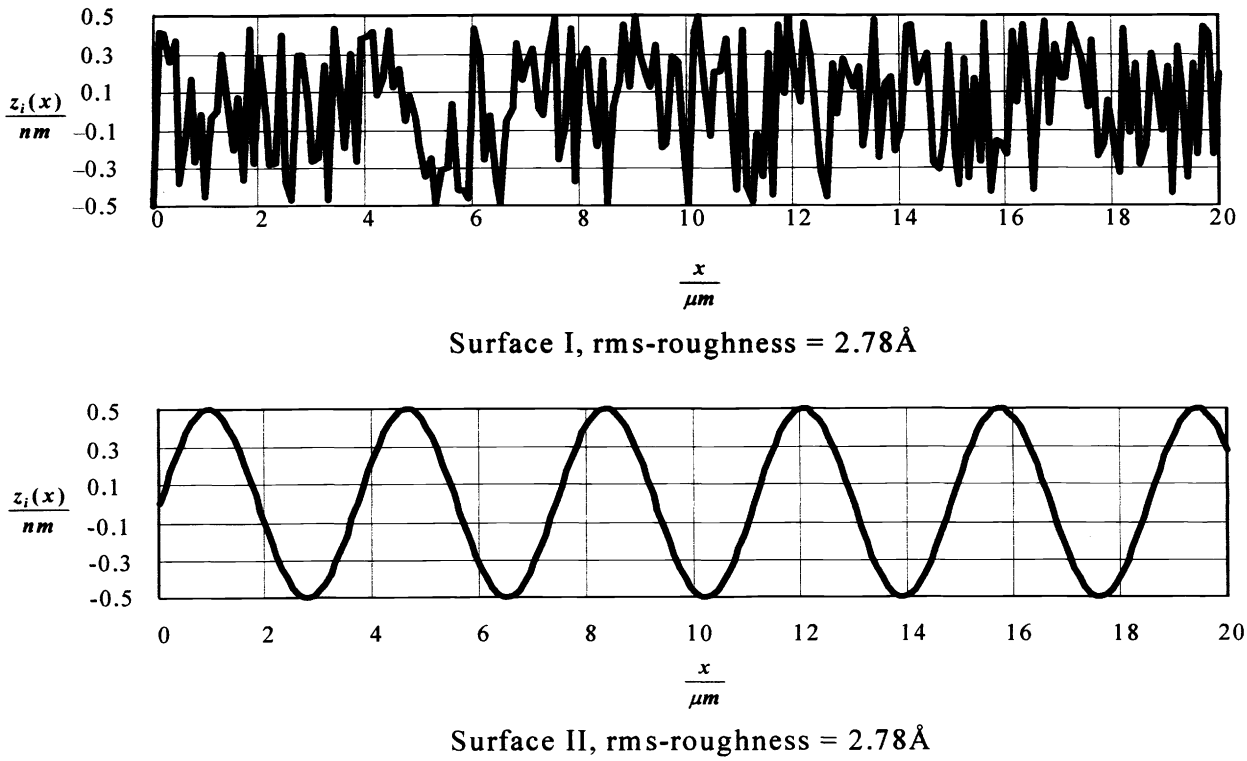


Figure 2. Comparison of two significantly different surfaces exhibiting the same measured rms roughness value.

2. INSTRUMENT CONSIDERATIONS IN DETERMINATION OF MICROROUGHNESS

Consider the surface from Figure 2, magnified to show a similar section 5 μm in length (Figure 3). If this surface is measured by an instrument which samples, say, every 0.5 μm as opposed to every 0.1 μm, then the reported surface tends to look smoother since the high spatial frequencies are missed. This is analogous to the electrical phenomenon of aliasing. In this case however, rather than sampling a time-varying function at too low of rate, a spatially-varying function is sampled too infrequently for a proper representation of the actual surface. In other words, the high end spatial frequency capability (measured in units of inverse micrometers) of the sampling instrument is too low. On the other extreme, if the warp in a 200 mm silicon wafer is measured by a stylus profilometer with a traversing length of 50 μm, it never senses the warped condition of the surface. Since the warp in the wafer tends to be near the wafer diameter, this implies that the spatial wavelength is 200 mm (or, conversely, a spatial frequency of $5(10^{-6})$ μm^{-1}).

** Spatial frequency is the inverse of spatial wavelength. A surface variation that changes every 10 micrometers would have a spatial frequency of 0.1 inverse micrometers.

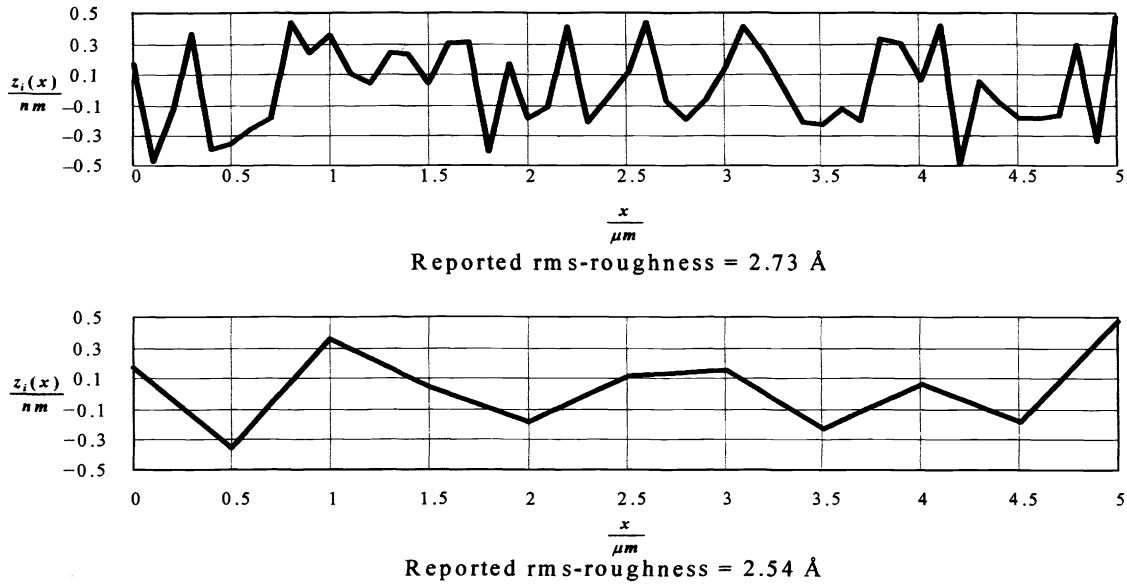


Figure 3. Reported rms microroughness of the same sample area by two different surface topography instruments.

The high frequency limit, f_{max}, of an instrument may be calculated directly by knowing the sampling distance, τ_o, where τ_o is greater than or equal to the lateral resolution of the instrument.

$$f_{max} = \frac{1}{2\tau_o} \tag{2}$$

The factor of 2 in the denominator assures that the minimum Nyquist length criterion is met. This helps to minimize aliasing effects.

The low frequency limit, f_{min}, is simply determined from the evaluation length, L, such that

$$f_{min} = \frac{1}{L} \tag{3}$$

where L is less than or equal to the traversing length, L_t. The spatial bandwidth (or bandpass) is then defined by these spatial frequency limits, or may be electronically narrowed (e.g., an electrical cut-off filter on a stylus profilometer).

In the case of a laser based instrument, such as an integrating scatterometer, the bandpass is defined somewhat differently. If the incident monochromatic laser beam, at incidence angle θ_i, specular beam, and scatter beams (at angles θ_s) are all in the same plane, the spatial frequency is related to the scatter angle by the one-dimensional grating equation

$$f = \frac{\sin(\theta_s)\cos(\phi_s) - \sin(\theta_i)}{\lambda} \tag{4}$$

where f is the spatial frequency and λ is the laser illumination wavelength. The $\cos\phi_s$ term takes on values of 1.0 and -1.0 for $\phi_s = 0°$ and 180° respectively.

Ideally, the instrument transfer function is flat in the bandpass region and zero elsewhere. This is never achieved in practice and the spatial bandpass must always be convolved with the transfer function (if available) in order to achieve true inter-instrument comparisons.

Referring back to Figure 3, the first surface is reported by an instrument with a high frequency limit of at least 5 μm^{-1} ($f_{max} = \frac{1}{2(0.1\mu m)}$). This instrument reports an rms roughness level of 2.73Å. The second instrument is reporting 2.54Å rms but has a high-end limit of 1 μm^{-1} ($f_{max} = \frac{1}{2(0.5\mu m)}$). However, this could be the same instrument with a low-pass filter

employed. It is shown later that this difference in reported roughness can be much more drastic with even more diverse instruments.

Figure 4 shows scans of an RQS standard made with an angle-resolved scatterometer. The format of this graph is a one-dimensional power spectral density plot which relates surface roughness power per unit of spatial frequency. The PSD plot is the square of the modulus of the Fourier transform. Recall that the Fourier transform of a square wave with a 50% duty-cycle is an infinite series of odd-order harmonics. In the case of an exact 50% duty-cycle, there is an added benefit in that Rq, or the rms-roughness value is identically equal to the arithmetic, or R_a, roughness. This value then equals half the overall height (peak-to-valley) of the square wave. Referring again to Figure 4, there are even order harmonics apparent here, but their amplitude is over two orders of magnitude lower than the primary peak at low spatial frequencies. Keep in mind that this is a log-log scale.

The most useful feature of the PSD function is that it relates information about the Fourier transform of the surface into a form that makes it possible to readily compare information generated from various instruments. The rms roughness may be calculated directly as the square root of the integral of the one-dimensional PSD curve. In the case of the RQS standard, an isotropic roughness value needs to be added in as well to account for the inherent roughness of the silicon; this is described in section 4.[††] Another compelling reason for making use of a square wave to produce a PSD plot is that over 90% of the power is contained in the first peak. This means that the standard is somewhat less sensitive to the exact instrument spatial bandwidth as long as at least the first peak is captured. Of course, it is important to completely integrate under the curve with the integration limits set equal to the instrument bandwidth for best results. Figure 5 shows some typical bandwidth limits for various pieces of surface texture measuring equipment.

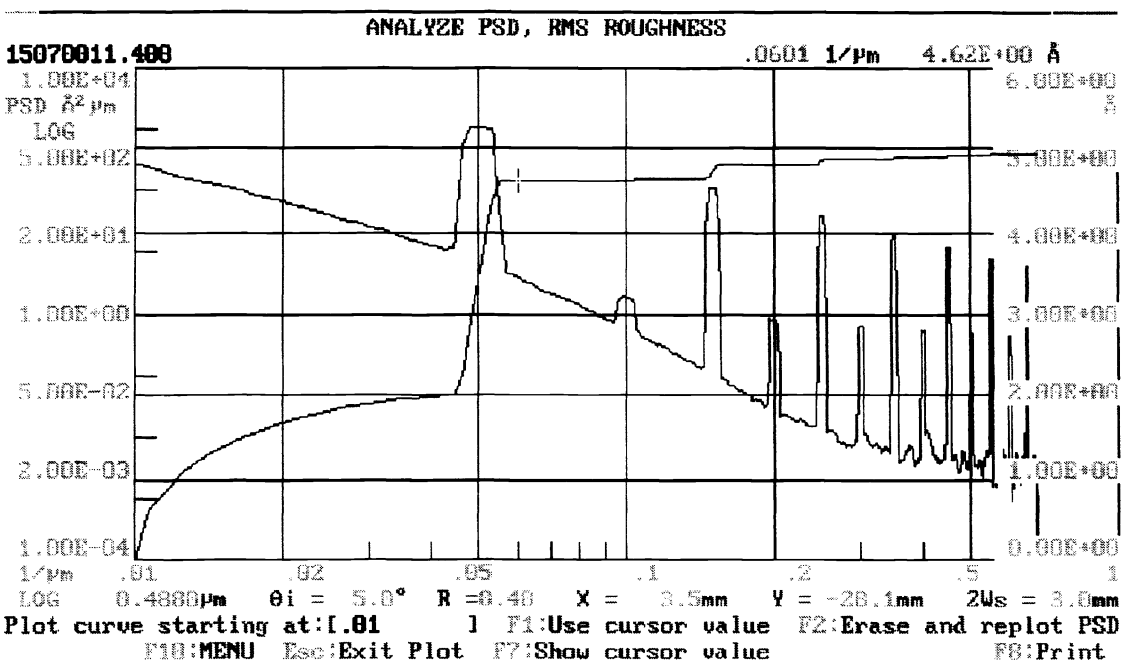


Figure 4. Power spectral density plots from angle-resolved light scattering scans of the RQS nominal 5 Å standard. The scale on the right side indicates rms-roughness. Notice the curve starting at the lower left hand corner by 0.01 μm^{-1}. This is the square root of the integrated value under the PSD plot – yielding rms-roughness. At the cursor location on the curve corresponding to 0.06 μm^{-1}, the rms roughness value is 4.62 Å (indicate in the upper right corner), or about 93% of the total 4.98 Å level at 0.6 μm^{-1}.

[††] A full description of the power spectral density is far beyond the scope of this paper but is discussed in more detail in section 4, *Determination of Specific Microroughness Levels*. The reader interested in a complete description is directed to the Stover reference, *Optical Scattering*.

Figure 4 was generated by a TMA CASI® angle-resolved light scattering (ARS) instrument. This is a specialized tool that allows for first-principles traceability based on the wavelength of light and optical geometries of the instrument. Details of the ARS procedure are presented in depth in section 4, *Determination of Specific Microroughness Levels.* For now, it is important to realize how the same surface can be measured by completely dissimilar instruments *and* give the same results within the same bandpass.

The realization of the graph being on a log-log scale emphasizes the assertions made earlier about the essential importance of knowing the spatial bandwidth limits of a given instrument. In the case of a typical measured surface (one that does not posses the unique features of a square wave), there will be a continuous power spectrum. In this case, integration between the limits of 0.01 μm^{-1} to 0.1 μm^{-1} (which may be typical for an optical profilometer) may result in a rms roughness value over an order of magnitude higher as compared with a 1 μm x 1 μm AFM scan (bandwidth of 1 μm^{-1} to 256 μm^{-1}; recall that the high spatial frequency, based on 512 sample points, is calculated from equation 2 as $\frac{1}{2 \cdot \tau_o} = \frac{1}{2 \cdot {}^{1\,\mu m}/_{512}} = 256\ \mu m^{-1}$).

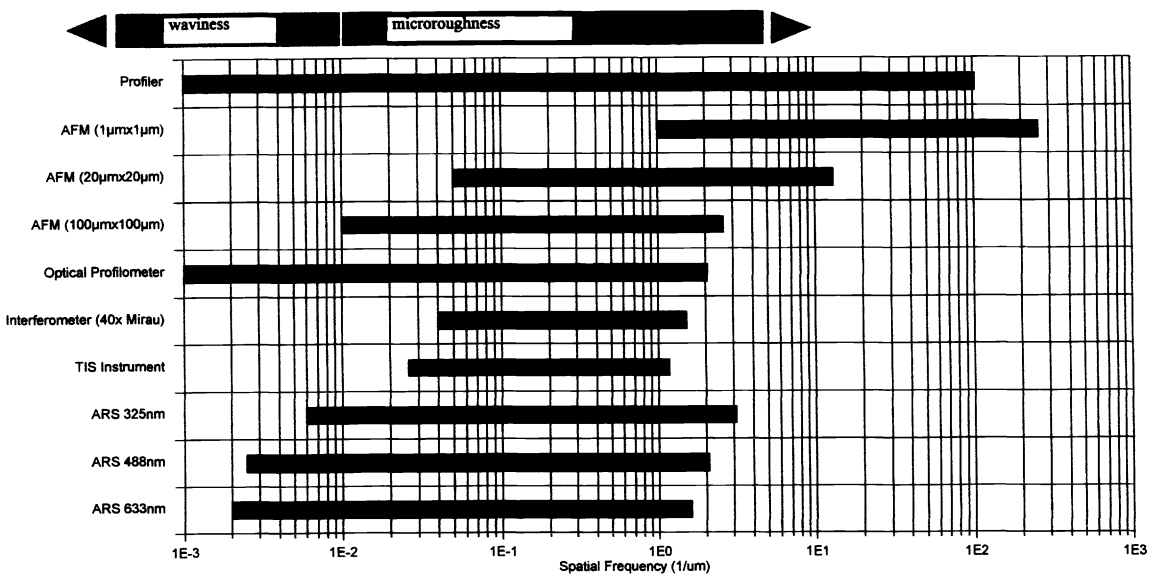


Figure 5. *Bandwidths of some surface texture instruments.*[9]

Knowing the PSD of a given surface allows determination of rms-roughness in a straightforward manner. The rms roughness, R_q, between spatial frequencies f_{min} and f_{max}, from a one-dimensional power spectral density function, PSD_{1D}, is

$$R_q = \sqrt{2 \int_{f_{min}}^{f_{max}} PSD_{1D}(f) df} \tag{5}$$

3. DESIGN CONSIDERATIONS

The relationship between the surface PSD and the resulting scatter was introduced by Church[x] into the optics literature in 1975 and has been extensively used to monitor micro-roughness via scatter measurement.[6] The general technique is explained in the next section, with the key relationship, sometimes called the "Golden Rule" appearing as Equation 9. In order for this relationship to be used to find the PSD from the measured scatter, virtually all of the scatter must come from surface topography. Non-topographic sources of scatter include surface bound particles, films, oxide layers and smooth surface index variations, such as those found across grain boundaries. Non-topographic sources of scatter follow other scatter laws besides that found in Equation 9. Sources of topographic scatter have the property that regardless of changes in wavelength , or incident angle the same PSD will be calculated. This property is called wavelength scaling in the literature. Significant scatter from particles and grain boundaries as well as the interference effects associated with surface films have all been shown to produce significant variations in the calculated PSD. In other words, they do not wavelength scale.

Clean, polished silicon wafers have been shown to scatter topographically from the near IR to the UV[6, 13]. They can be produced virtually free of contamination and films, and as single crystal surfaces, they are free of effects from grain boundaries. In general, these surfaces are very low scatter, and for the spatial bandwidths in question can be characterized by a low level isotropic background roughness. Because techniques are available for producing designed surface structures on silicon, this is an ideal material for a micro-roughness standard.

The basic design employs a nominally 50% duty-cycle square wave with a period of 20 μm. This produces a fundamental spatial frequency component of 0.05 μm^{-1} followed by odd harmonics at 0.15 μm^{-1}, 0.25 μm^{-1}, etc..

4. DETERMINATION OF SPECIFIC MICROROUGHNESS LEVELS

Since all of the data from angle-resolved scatterometry are binned into discrete points, the "integration" required to turn these data into a rms-roughness value merely becomes the square root of the summation of the one-dimensional PSD function, $PSD_{1D}(f)$, multiplied times the differential frequency step size within the appropriate limits of integration.

$$R_q = \left(2 \int_{f_l}^{f_h} PSD_{1D}(f) df \right)^{\frac{1}{2}} \quad \Rightarrow \quad R_q = \sqrt{2 \sum_{i=f_l}^{f_h} PSD_{1D} \delta f} \tag{6}$$

Note also that the equation signifies the limits of integration denoted f_l and f_h which are set according to the spatial bandwidth limits of the instrument under test. To account for the inherent isotropic roughness of the silicon, we also need to add in an additional term PSD_{iso} to equation 6. The final formula for the RQS standard takes the form of equation 7.

$$R_q = \left(2 \int_{f_l}^{f_h} PSD_{1D}(f) df + \int_{f_l}^{f_h} PSD_{iso}(f) df \right)^{\frac{1}{2}} \tag{7}$$

This is the equation that is used to certify a rms-roughness value for the standard for a given range of integration limits.

4.1 Development of the PSD Calibration Curve

The actual quantity measured through angle-resolved scatterometry is the Bi-directional Reflectance Distribution Function, BRDF

$$BRDF = \frac{P_s / \Omega_s}{P_i \cdot \cos\theta_s} \tag{8}$$

where P_S is the power of the scattered light collected over the solid angle Ω_S as a function of the angle θ_s. The factor P_i indicates the incident laser power at angle θ_i from the wafer normal.[10] Therefore, BRDF is physically nothing more than the redistributed energy scattered into a given solid angle. Recall from equation 4 that the spatial frequency, f, is related to the scatter angle by the one-dimensional grating equation

$$f = \frac{\sin\theta_s \cos(\phi_s) - \sin\theta_i}{\lambda}$$

The PSD function then is calculated from the BRDF and is a measure of the scattered power per unit of spatial frequency in units of $Å^2 \mu m^2$

$$PSD(f_x, f_y) = \frac{10^8 \lambda^4 BRDF}{16\pi^2 \cos\theta_i \cdot \cos\theta_s \, Q} \tag{9}$$

where, for s-polarization, the factor Q is approximated by the specular reflectance of the wafer surface. The reflectance is a function of wavelength, incidence angle, and polarization. If the surface is isotropic, $PSD(f_x, f_y)$ may be integrated around the azimuthal angle to obtain an isotropic PSD function

$$PSD_{iso}(f) = \int_0^{2\pi} PSD(f_x, f_y) f \, d\phi_s = 2\pi f \cdot PSD(f_x, f_y) \tag{10}$$

with units of Å²µm and f equals the root-sum-of-squares of f_x and f_y. It is this function, $PSD_{iso}(f)$ in equation 10, that is added to the one-dimensional PSD function and becomes the function from which rms roughness is calculated. The square wave is a one-dimensional scatterer, however, the roughness of the silicon "peaks and troughs" are isotropic scatterers. Therefore, two measurements are made – one with the direction of the square wave surface perpendicular to the plane-of-incidence of the laser (for the overall one-dimensional PSD curve generation), and one where the surface is parallel. The second scan measures the silicon roughness, independent of the etched features. This value is then converted to an isotropic value, equation 10, and then added as the root-sum-of squares to the one-dimensional scan (equation 7). The silicon roughness is really only significant at low nominal roughness values but is consistently measured as a matter of practice.

Figure 6 provides some reference as to what type of PSD may be expected for various surfaces typically encountered. The reader versed in signal theory will also note the analogy to Fourier analysis of communications signals.

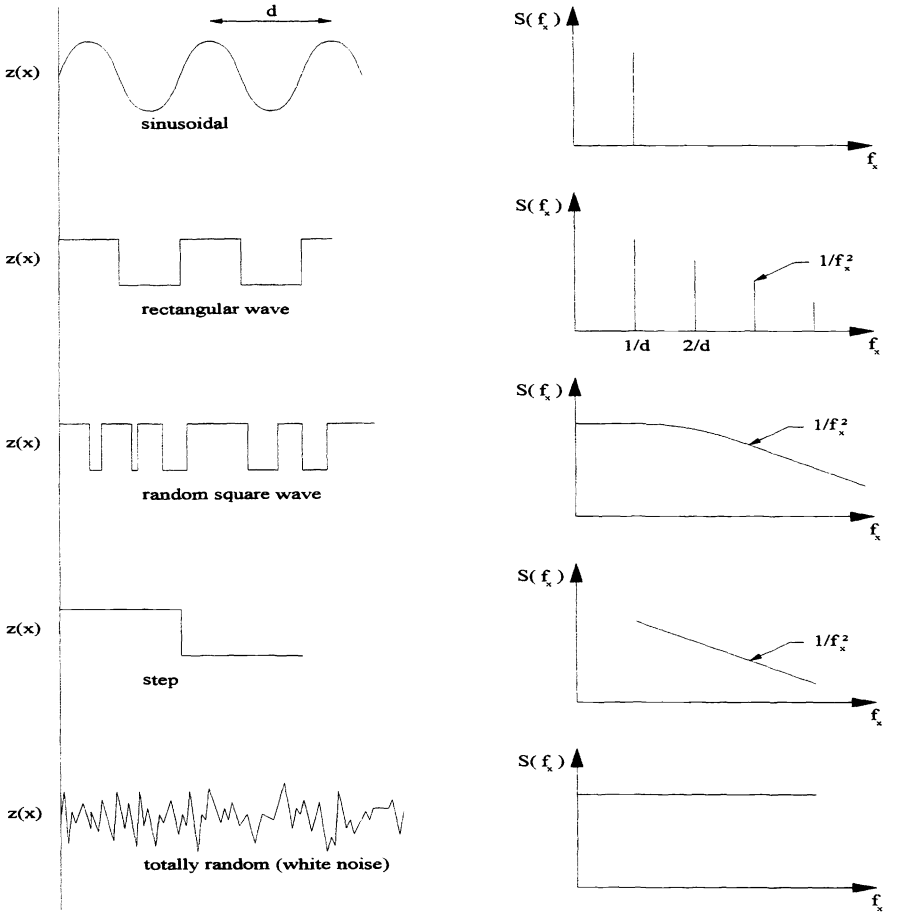
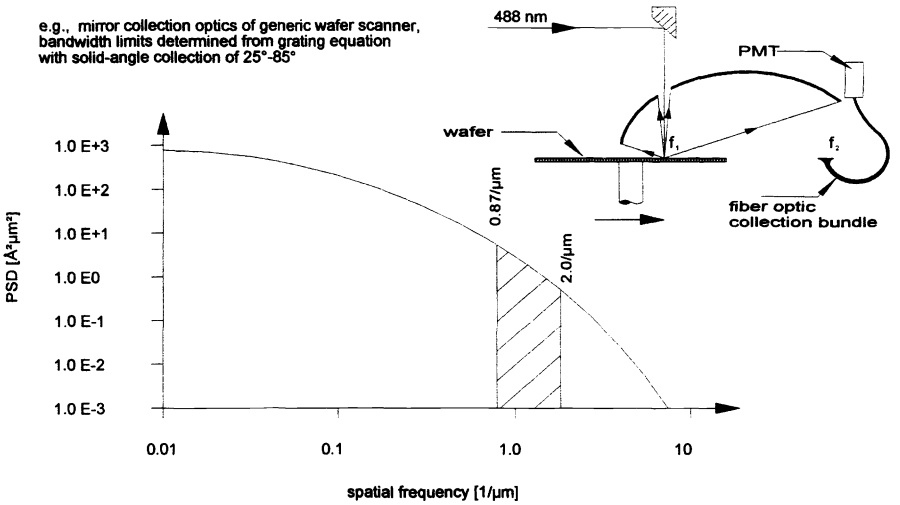


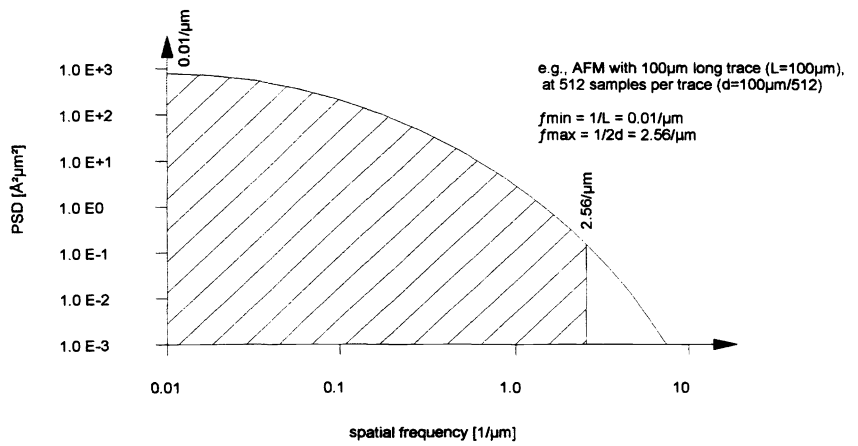
Figure 6. Examples of a how a given surface may be defined as a one-dimensional PSD function. (Adapted from reference 9).

Figures 7 and 8 summarize the determination of the power spectral density function and spatial frequency bandwidth for scattering instruments (e.g., total integrated scatterometers) and profilers (e.g., atomic force microscopes) respectively. Notice that the instruments may be correlated directly in the region of bandpass overlap.



$$PSD(f_x, f_y) = \frac{10^8 \lambda^4 BRDF}{16\pi^2 \cos\theta_i \cdot \cos\theta_s \, Q} \left[\frac{^\circ 2}{A} \; \mu m^2 \right]$$

Figure 7. Determination of the PSD and related spatial frequencies for a generic total integrated scattering inspection system. PSD is a measure of the scattered power per unit of spatial frequency.



The PSD is calculated as the square of the Fourier transform from a 2-D surface profile

$$PSD(f_x, f_y) = \frac{d^2}{MN} \left| \sum_{k=1}^{M} \sum_{j=1}^{N} z_{j,k} \cdot \exp\left\{ -i2\pi \left[f_x(j-1) + f_y(k-1) \right] d \right\} \right|^2 \left[\frac{^\circ 2}{A} \; \mu m^2 \right]$$

where f_x, f_y are the spatial frequencies such that $f = \kappa / L$, with κ being an integer **1,2, ...**, $N/2$, and d is the same in both the x- and y-directions.

Figure 8. Determination of the PSD and related spatial frequencies for a surface profiling system (e.g., an atomic force microscope).

5. SUMMARY

When making comparisons between different types of texture measurement tools, it is extremely important to report the spatial bandwidth along with the reported rms-roughness value. Even when comparing results between the same instrument type, the operating parameters can significantly affect the reported results. This variation can be well over an order of magnitude different. Ideally, the transfer function for a given instrument must also be known for the best quantitative comparison between toolsets.

Finally, a practical microroughness standard has been developed that is based upon the principles of optical scatterometry for certification. This standard allows for comparison between seemingly uncorrelatable tools, within the proper spatial bandwidth for each toolset. Using this methodology, it is possible to compare microroughness readings on a TIS system to the microroughness readings on an atomic force microscope.

6. REFERENCES

[1] Stover, John C., "Requirements and Suggestions for Industrial Smooth Surface Microroughness Standards", *Flatness, Roughness, and Discrete Defect Characterization for Computer Disks, Wafers, and Flat Panel Displays proceedings*, 8-9 August 1996, Denver, CO, SPIE vol 2862, pgs. 69-77.

[2] ASTM (American Society for Testing and Materials), F-1 Electronics Committee, 100 Barr Harbor Drive, West Conshohocken, PA 19428-2959.

[3] Semiconductor Industry Association, *National Technology Roadmap for Semiconductors*, June 1994.

[4] Fink, Donald G., and Donald Christiansen, eds. *Electrical Engineering Handbook*, Third Edition, McGraw-Hill, New York, 1989.

[5] Bullis, W. Murray, "Microroughness of Silicon Wafers", *Electrochemical Society Silicon Symposium*, San Francisco, May, 1994.

[6] Stover, J.C., *Optical Scattering Measurement and Analysis*, McGraw-Hill, New York, 1990.

[7] ANSI/ASME Standard B46.1, *Surface Roughness, Waviness, and Lay*, 1993.

[8] Bennett, Jean M. and Lars Mattsson, *Introduction to Surface Roughness and Scattering*, Optical Society of America, Washington, D.C., 1989.

[9] Adapted from W. Murray Bullis *ECS* presentation and *Particles, Haze and Microroughness Symposium*, San Jose, CA, September, 1994.

[10] Bullis, W. Murray, "Microroughness of Silicon Wafers", *Electrochemical Society Silicon Symposium*, San Francisco, May, 1994.