Volume 96, Number 3, May–June 1991
Journal of Research of the National Institute of Standards and Technology

Contents

Articles

Volume 96, Number 3, May–June 1991
Journal of Research of the National Institute of Standards and Technology

News Briefs

Volume 96, Number 3, May–June 1991
Journal of Research of the National Institute of Standards and Technology

Volume 96, Number 3, May–June 1991
Journal of Research of the National Institute of Standards and Technology

[J. Res. Natl. Inst. Stand. Technol. 96, 247 (1991)]

A High-Temperature Transient Hot-Wire Thermal Conductivity Apparatus for Fluids

Volume 96 | Number 3 | May–June 1991

R. A. Perkins and H. M. Roder

National Institute of Standards and Technology,
Boulder, CO 80303

and

C. A. Nieto de Castro[1]

Departamento de Quimica,
Universidade de Lisboa,
R. Ernesto Vasconcelos, Bloco C1,
1700 Lisboa, Portugal

A new apparatus for measuring both the thermal conductivity and thermal diffusivity of fluids at temperatures from 220 to 775 K at pressures to 70 MPa is described. The instrument is based on the step-power-forced transient hot-wire technique. Two hot wires are arranged in different arms of a Wheatstone bridge such that the response of the shorter compensating wire is subtracted from the response of the primary wire. Both hot wires are 12.7 μm diameter platinum wire and are simultaneously used as electrical heat sources and as resistance thermometers. A microcomputer controls bridge nulling, applies the power pulse, monitors the bridge response, and stores the results. Performance of the instrument was verified with measurements on liquid toluene as well as argon and nitrogen gas. In particular, new data for the thermal conductivity of liquid toluene near the saturation line, between 298 and 550 K, are presented. These new data can be used to illustrate the importance of radiative heat transfer in transient hot-wire measurements. Thermal conductivity data for liquid toluene, which are corrected for radiation, are reported. The precision of the thermal conductivity data is ±0.3% and the accuracy is about ±1%. The accuracy of the thermal diffusivity data is about ±5%. From the measured thermal conductivity and thermal diffusivity, we can calculate the specific heat, C_p, of the fluid, provided that the density is measured, or available through an equation of state.

Key words: argon; heat capacity; nitrogen; radiation correction; thermal conductivity; thermal diffusivity; toluene; transient hot-wire.

Accepted: March 5, 1991

1. Introduction

The transient hot-wire method is widely accepted as the most accurate technique for fluid thermal conductivity measurements at physical states removed from the critical region proper [1]. The method is very fast relative to steady state techniques. The duration of a typical experiment is about 1 s when 250 temperature rises are measured. Normally the experiment is completed before free convection can develop in the fluid. If free convection is present, it is easy to detect because it results in a pronounced curvature in the graph of temperature rise versus the logarithm of time.

In addition to the thermal conductivity, thermal diffusivity can be measured with transient hot-wire instruments. With an appropriate design of the instrument [2], measurements of fluid thermal diffusivity can be made with reasonable accuracy over wide ranges of density. The heat capacity of a fluid can then be obtained from the measurements of thermal conductivity and thermal diffusivity, provided that the density is known or available from an equation of state.

[1] Also Centro de Quimica Estrutural, Complexo I, IST, 1096 Lisboa Codex, Portugal.

2. Method

The transient hot-wire system is considered to be an absolute primary instrument [1]. The ideal working equation is based on the transient solution of Fourier's law for an infinite linear heat source [3]. The temperature rise of the fluid at the surface of the wire, where $r = r_0$, at time t is given by

$$\Delta T_{\text{ideal}}(r_0,t) = \frac{q}{4\pi\lambda}\ln\left(\frac{4at}{r_0^2 C}\right) = \frac{q}{4\pi\lambda}\ln\left(\frac{4a}{r_0^2 C}\right) + \frac{q}{4\pi\lambda}\ln(t). \quad (1)$$

In eq (1), q is the power input per unit length of wire, λ is the thermal conductivity, $a = \lambda/\rho C_p$ is the thermal diffusivity of the fluid, ρ is the density, C_p is the isobaric heat capacity, and $C = e^{\gamma} = 1.781...$ is the exponential of Euler's constant. We use eq (1) and deduce the thermal conductivity from the slope of a line fit to the ΔT_{ideal} versus $\ln(t)$ data. The working equation for the thermal diffusivity is

$$a = \frac{r_0^2 C}{4t'}\exp\left[\frac{4\pi\lambda\Delta T_{\text{ideal}}(r_0,t')}{q}\right]. \quad (2)$$

The thermal diffusivity is obtained from λ and a value of ΔT_{ideal}, from the fit line, at an arbitrary time t'. We normally select t' to be 1 s in our data analysis, as discussed in reference [2].

The thermal conductivity is reported at the reference temperature T_r and density ρ_r defined in eq (3) below. The thermal diffusivity calculated from eq (2) must be referred to zero time, that is, the equilibrium or cell temperature. In summary, the thermal conductivity and the thermal diffusivity evaluated by the data reduction program are related to the reference state variables and to the zero time cell variables as follows:

$$\begin{aligned}
&\lambda = \lambda(T_r,\rho_r),\\
&T_r = T_0 + 0.5\,(\Delta T_{\text{initial}} + \Delta T_{\text{final}}),\\
&\rho_r = \rho(T_r,P_0),\\
&a = a(\rho_0,T_0) = \frac{\lambda(T_0,\rho_0)}{\rho_0(C_p)_0},\\
&\rho_0 = \rho(T_0,P_0), \text{ and}\\
&(C_p)_0 = C_p(T_0,P_0),
\end{aligned} \quad (3)$$

where T_0 is the equilibrium temperature and P_0 is the equilibrium pressure at time $t = 0$.

The experimental apparatus is designed to approximate the ideal model as closely as possible. There are, however, a number of corrections which account for deviations between the ideal line-source heat transfer model and the actual experimental heat transfer situation. The ideal temperature rise is obtained by adding a number of corrections to the experimental temperature rise as

$$\Delta T_{\text{ideal}} = \Delta T_{\text{experimental}} + \sum_i \delta T_i. \quad (4)$$

These temperature rise corrections are described in references [2,4]. Our implementation of the corrections follows these two references with the exception of the thermal radiation correction. This correction is dependent on the optical properties of the fluid and the cell, and is discussed in more detail below.

2.1 The Radiation Correction

If the fluid is transparent to infrared radiation, then this correction is only a function of the cell geometry and the optical properties of the materials used in its construction. The radiation correction described in references [2,4] assumes that all of the surfaces in the cell are blackbodies. The blackbody radiation correction is given by

$$\delta T_{5T} = \frac{8\pi r_0 \sigma T_0^3 \Delta T^2}{q} \quad (5)$$

where σ is the Stefan-Boltzmann constant. In practice, many experimenters assume that this correction is negligible and neglect the correction. We have found that this correction changes the reported thermal conductivity of argon at 300 K by about 1% for our geometry, so it is not appropriate to ignore it. A more accurate correction can be obtained by considering the optical properties of the surfaces in the hot-wire cell.

For this analysis we consider the cell surfaces to be diffuse gray surfaces and follow the analysis presented in reference [5]. We consider the cell to be an infinitely long hot wire in a concentric cylindrical cavity. Thus, two surfaces are involved in the heat transfer. Surface 1 is the hot wire whose temperature is a function of time, and surface 2 is the cylindrical cavity surrounding the hot wire which remains at the initial equilibrium temperature. The net radiative heat flux for the hot wire, using the tabulated view factors in reference [5], is

$$Q_1 = \frac{A_1\sigma(T_1^4 - T_2^4)}{\frac{1}{\epsilon_1} + \frac{A_1}{A_2}\left(\frac{1}{\epsilon_2} - 1\right)}, \quad (6)$$

where A_i is the area, T_i is the temperature, and ϵ_i is the emissivity of surface i. The ratio of the surface areas A_1/A_2 which is present in the denominator of eq (6) is quite small since very thin hot wires are used. In our cell this surface area ratio is $A_1/A_2 = 0.001$. The inverse emissivity of the hot wire $1/\epsilon_1$ varies from 10 to 25 for platinum and $1/\epsilon_2$ is approximately 2. Therefore, the second term in the denominator of eq (5) is negligible to within 0.1% in Q_1, and we are left with

$$Q_1 = A_1 \epsilon_1 \sigma (T_1^4 - T_2^4). \tag{7}$$

Because the surface area of the cavity surrounding the hot wire is so much larger than the surface area of the hot wire, to a first approximation the heat transfer is not a function of the emissivity of the cavity.[2] The cavity appears to be a blackbody, and the heat transfer is only a function of the emissivity of the platinum hot wire. Following the analysis of reference [4], the resulting correction to the experimental temperature rise in a transparent fluid is

$$\delta T_{5T} = \frac{8\pi r_0 \epsilon_{\text{platinum}} \sigma T_0^3 \Delta T^2}{q}. \tag{8}$$

The emissivity of platinum, $\epsilon_{\text{platinum}}$, is a function of temperature and is tabulated in reference [6]. At 300 K the emissivity of platinum is 0.0455 relative to an emissivity of 1 for a blackbody. The blackbody radiation correction of eq (5) is roughly 20 times larger than the real case, eq (8), when platinum hot wires are used.

For fluids which absorb infrared radiation, the technique described in reference [7] works well. The technique is based on the numerical simulations of transient conduction and radiative heat transfer from a hot wire in an absorbing medium. Since the emissivity of the platinum hot wire is so small, the radiative heat flux from the wire is negligible in the simulations. The primary mechanism for radiative losses is from emission from the fluid at the boundary of the expanding conduction front. This analysis [7] yields a radiation correction for absorbing media which is given by

$$\delta T_{5A} = -\frac{qB}{4\pi\lambda}\left[\frac{r_0^2}{4a}\ln\left(\frac{4at}{r_0^2 C}\right) + \frac{r_0^2}{4a} - t\right]. \tag{9}$$

[2] This is possible because, as shown later, $Q_1/Lq \approx 2 \times 10^{-2}$ for the transient hot-wire instrument and, therefore, an error of 0.1% in Q_1 produces an error of 0.002% in q, well beyond the experimental accuracy.

The radiation parameter B is related to the fluid properties by

$$B = \frac{16 K n^2 \sigma T_0^3}{\rho C_p}, \tag{10}$$

where K is the mean extinction coefficient of the fluid and n is its refractive index. These fluid properties are a function of the fluid density and temperature and are not generally available. The procedure described in reference [7] allows B to be estimated from the experimental temperature rise data. Equation (9) indicates that the radiation correction introduces a term which is a direct function of time into the temperature rise equation. When the radiation correction is added to the ideal temperature rise, we obtain

$$\Delta T = \frac{q}{4\pi\lambda}\left[1 + \frac{Br_0^2}{4a}\right]\ln\left(\frac{4at}{r_0^2 C}\right) - \frac{Bqt}{4\pi\lambda} + \frac{Bqr_0^2}{16\pi a\lambda} + \ldots . \tag{11}$$

Thus, we correct the experimental data with all the other corrections and fit the resulting temperature rise to a function of the form

$$\Delta T = C_1 \ln(t) + C_2 t + C_3. \tag{12}$$

The experimental radiation parameter B is determined from coefficient C_2 using

$$B = C_2\left(\frac{-4\pi\lambda}{q}\right). \tag{13}$$

Once B is determined, we use eq (9) to correct for radiation in the absorbing fluid. This technique allows us, as shown later, to use our experimental data to determine whether there is a significant thermal radiation correction in an absorbing fluid and to correct for the radiation. No prior knowledge of the optical properties of the fluid is required.

3. Apparatus

The apparatus is quite similar to a previously described low temperature system [8] which is used from 80 to 320 K. The new apparatus is designed to operate from 220 to 750 K at pressures to 70 MPa. A preliminary version of the new instrument has been described elsewhere [9]. Improvements have been incorporated into the new system to improve the precision and accuracy of the thermal conduc-

tivity measurement and to enable measurement of the thermal diffusivity. They were based on modifications introduced in the low temperature system which are fully described in references [10] and [11].

3.1 Hot Wires

The hot wires are selected to conform to the ideal line-source model as closely as possible. The line-source model assumes that the wire has no heat capacity and that it is infinitely long, so there is no axial heat conduction. The wire diameter is 12.7 μm in this instrument to minimize effects due to its finite heat capacity while retaining good tensile strength and uniformity. A two-wire compensating system is used in order to eliminate effects due to axial heat conduction. The arrangement of the two wires is shown in figure 1. The two wires have different lengths and are arranged in a modified Wheatstone bridge where the thermal response of the short wire is subtracted from the response of the long wire. The resulting response from a finite length of wire approximates that of an infinitely long hot wire. The length of the equivalent wire is the difference in the lengths of the long and short hot wires.

The hot wires are used simultaneously as electrical heat sources and as resistance thermometers.

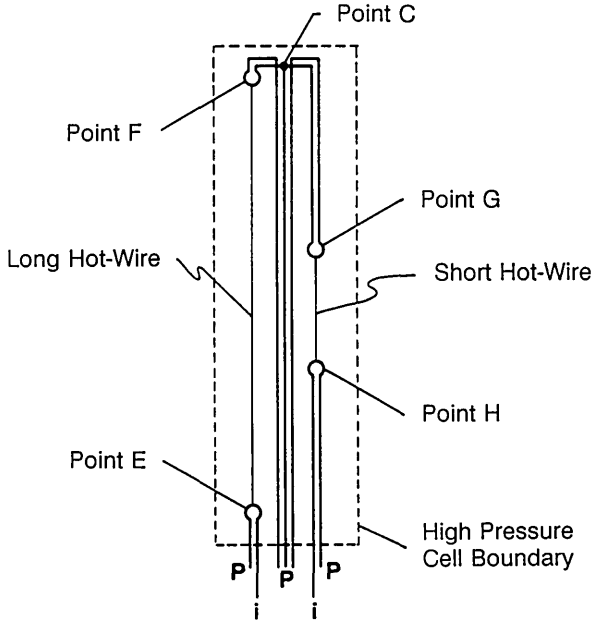

Figure 1. Arrangement of current leads (i) and potential taps (P) within the pressure cell. Bridge points correspond to those in figure 3.

Platinum wire is used in this instrument because its mechanical and electrical properties are well known over a wide temperature range, and it is resistant to corrosion up to 750 K. As shown above, platinum has the added advantage of low emissivity. The length of the long hot wire is about 19 cm. The length of the short hot wire is about 5 cm. The platinum hot wires are annealed after they are installed, so that their resistance will be stable during high temperature operation. The resistance of the annealed hot wires is about 20% less than the hard-drawn platinum wire. The resistance of the hot wires is calibrated *in situ* as a function of temperature and pressure [12].

The wires are welded to rigid upper suspension stirrups and weighted lower suspension stirrups. The floating lower weights are used to tension the wires and to allow for thermal expansion. There are fine copper wires welded between the bottom weights and the massive bottom leads. These fine wire leads are flexible so that they do not introduce significant stress on the platinum hot wires. This arrangement provides both current and potential leads to both ends of each hot wire. Thus, four-terminal resistance measurements can be made on both the long and short hot wires, eliminating uncertainty due to lead resistances.

3.2 Hot-Wire Cell

The two platinum hot wires are contained in a pressure vessel which is designed for 70 MPa at 750 K. The cell is connected with a capillary tube to a sample-handling manifold. This sample-handling manifold allows evacuation of the cell, charging and pressurization of liquids with a screw pump, and pressurization of gases with a diaphragm compressor. There are seven electrical leads into the pressure vessel to enable four-terminal resistance measurements of both hot wires. The electrical leads pass through a 6.25 mm O.D. pressure tube which connects the bottom of the pressure vessel to the lead pressure seal. The pressure seal for the electrical leads is made at ambient temperature for improved reliability. The vessel access tube is located on the bottom of the vessel so that there is always a positive temperature gradient with respect to height to eliminate free convective driving forces. The entire pressure system is constructed of 316 stainless steel for corrosion resistance.

The thermal conductivity cell is shown in its temperature control environment in figure 2. The cell pressure vessel is surrounded by a 12 mm thick cylindrical aluminum heat shield. The aluminum

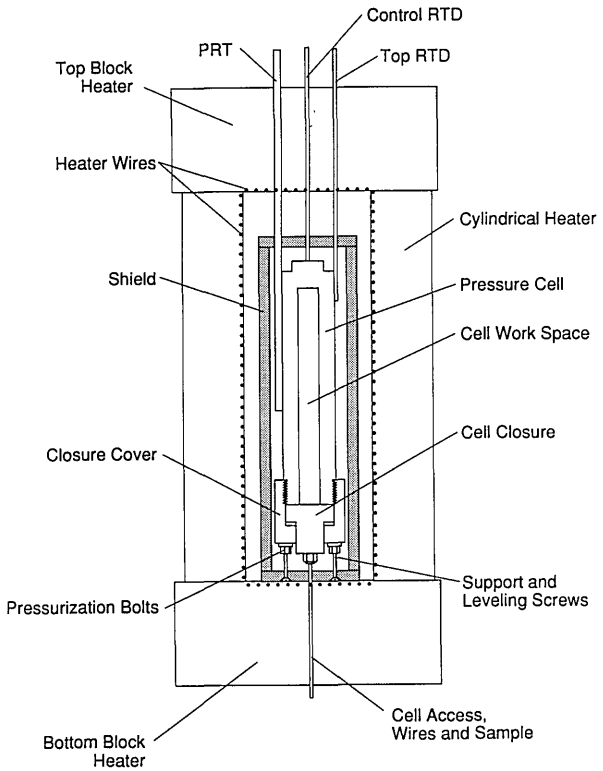

Figure 2. High-pressure cell, shield, and furnace.

has a high thermal conductivity and provides a nearly isothermal environment for the pressure vessel. There is an air gap between the vessel and the heat shield. This air gap isolates the pressure vessel from temperature fluctuations in the heat shield. Tubes are silver-soldered to the outside of the pressure vessel which enclose the reference standard platinum resistance thermometer (PRT) and two smaller platinum resistance probes (RTDs). The two RTDs can be moved axially along the vessel to detect temperature gradients. Normally, one RTD is located near the top of the vessel, and the other RTD is located near the bottom of the vessel. This configuration allows us to measure the cell temperature at the center of the vessel with the reference standard PRT and temperature gradient in the cell with the two RTDs for each thermal conductivity measurement.

For experiments from ambient temperature to 750 K, the vessel and heat shield are placed in a cylindrical furnace constructed of heating elements cast in fibrous ceramic insulation. These heating elements are shown in figure 2 and are separated from the aluminum heat shield by a second air gap. An additional platinum RTD is located on the top of the aluminum heat shield. This probe provides the feedback signal for the furnace temperature control system. The main power supply is under computer control and is connected to the bottom end heating element and the tubular heating elements. The second trim power supply is manually controlled to eliminate axial gradients in the thermal conductivity cell. The heating elements are driven with dc power supplies to minimize electromagnetic noise in the thermal conductivity instrument. Temperature fluctuations in the cell are normally less than 0.01 K.

For experiments between 220 and 300 K, the electrical heaters are replaced by a copper cooling coil enclosed in polystyrene insulation. A refrigerant with a low freezing point is pumped through the cooling coil by a recirculating temperature control bath. This recirculating bath controls the fluid temperature to within 0.01 K. The aluminum heat shield and air gap further reduce the temperature fluctuations in the cell to less that 0.01 K.

3.3 Wheatstone Bridge Circuit

This instrument uses a Wheatstone bridge circuit to monitor the resistance changes of the hot wires during the step-power pulse. The two hot wires are set up in opposing legs of the Wheatstone bridge as shown in figure 3. The drive voltage is applied across points A and B. The bridge response is monitored by a high speed digital multimeter across points C and D. The bridge is initially balanced with a 100 mV drive voltage. There is negligible heating of the hot wires with this small balance voltage. The four legs of the Wheatstone bridge are designated R1, R2, R3, and R4. Each of the four legs contains a variable decade resistor. The smallest step on these decade resistors is 0.01 Ω. These four decade resistors are adjusted so that the bridge imbalance signal is 0 and the total resistance of each leg is the same.

There are two current paths between points A and B. Each current path contains a calibrated 100 Ω standard resistor in order to determine the current flowing through that path during the balancing procedure. Figure 3 shows a number of voltage taps on the Wheatstone bridge which allow the multiplexed digital multimeter to measure the voltage drops across all of the resistances in the bridge. Using the current, provided by the voltage drop across the standard resistors, we can obtain the resistance of all of the components of the bridge.

These resistances must be known very precisely, and the bridge must be balanced very closely, in

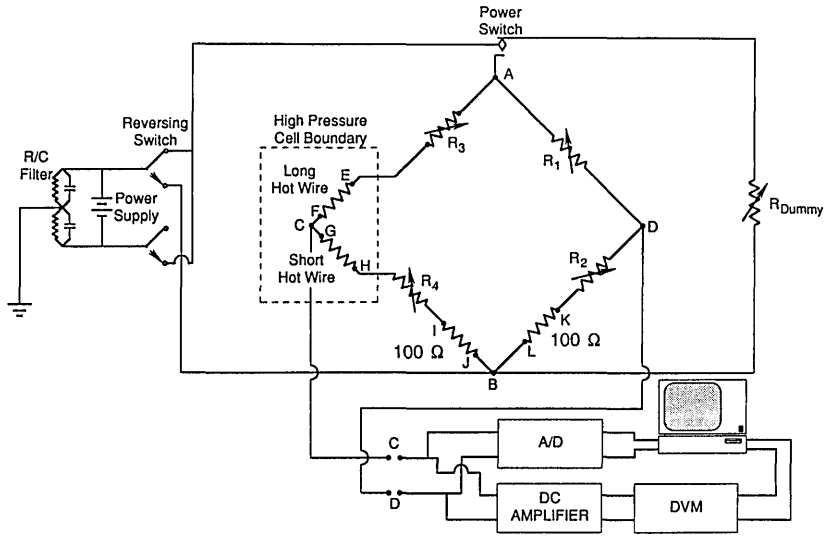

Figure 3. The Wheatstone bridge schematic for the transient hot-wire apparatus. Potential taps are indicated by points A–L.

order to obtain accurate thermal diffusivities from the experiment. Thermal voltages from the components of the bridge have a significant impact on the balancing of the bridge. In order to eliminate errors from thermal voltages, the bridge is alternately measured with a positive and negative drive voltage with a reversing relay. During the balancing procedure, 10 alternating drive voltage cycles are measured. During each cycle the digital multimeter monitors the voltage across all of the voltage taps. These values are subsequently averaged and displayed by the system computer.

When a satisfactory bridge balance is obtained, we are ready to begin the transient hot-wire experiment. The power supply is switched to a dummy resistor and the drive voltage is set to a level which will produce the desired heating of the hot wires. The experiment begins when the power supply is switched from the dummy resistor to the Wheatstone bridge. During the experiment the multimeter records the bridge voltage as a function of time across points C and D. This signal is proportional to the differential resistance change of the two hot wires. This differential resistance change of the two wires is related to the temperature changes of the two hot wires by the wire calibration which is described below. The experiment normally lasts 1 s with a bridge response voltage recorded every 4 ms.

3.4 Data Acquisition and Control

Data acquisition and control are coordinated by a personal computer. The computer controls the cell temperature, synchronizes the experimental timing, records the data, and provides a graphical display of the data. The computer has an analog-to-digital interface board which generates the timing signals based on the computer's internal quartz crystal oscillator and controls the system voltage multiplexers. The computer is also equipped with an IEEE-488 interface which allows communication with a dedicated digital temperature controller, a digital nanovoltmeter, and the high speed digital multimeter.

The cell PRT and the two gradient RTDs are connected in series with a standard resistor and a precision 1 mA current source. The computer controls a multiplexer which allows the nanovoltmeter to measure the voltage drops across the three resistance thermometers and the calibrated standard resistor. Using the current which is determined by the voltage drop across the standard resistor, we can obtain the resistances of the three thermometers.

A second multiplexer is connected to the input of the high speed digital multimeter. This multiplexer allows sampling of all the voltage taps on the Wheatstone bridge during bridge balancing. Since standard resistors are included in both current paths of the bridge, we can obtain accurate measurements of all the resistances in the bridge. The resistance of the two hot wires is used in conjunction with the PRT temperature to obtain the calibrations for the hot wires. In addition, the multiplexer allows us to measure the drive voltage and the resistance of the power switching relay for

an accurate determination of the power applied to the hot wires.

During the experiment, there are two parallel systems measuring the bridge response. A 16 bit analog-to-digital converter directly monitors the bridge response, while the high speed digital multimeter monitors the response of an instrumentation amplifier which is also connected across points C and D. The instrumentation amplifier has a fixed gain of 100 and also has an analog filter built in. This filter significantly reduces the noise of the bridge response but introduces a time lag which we must account for. The noise of the raw signal is 25 μV but is reduced to 3 μV by the filter. The experimental timing is fixed by the raw signal which is monitored by the analog-to-digital converter. The relatively noisy raw signal is used to adjust the timing of the filtered bridge response which is recorded by the high speed digital multimeter.

4. Hot-Wire Calibration

The electrical resistance of pure platinum as a function of temperature is very well characterized because of its widespread use in thermometry. In most thermometry applications the platinum is maintained at ambient pressure. In transient hot-wire instruments, however, the platinum is immersed directly in the fluid of interest. Roder et al. [12] showed that the effect of pressure on the resistance of the platinum hot wires must be accounted for. The functional form of our calibration is given by

$$R(T,P)=A+B\,T+C\,T^2+(D+E\,T)P, \qquad (14)$$

where R is the wire resistance, T is the temperature, and P is the applied pressure.

We have found that an *in situ* calibration provides the most reliable measurements possible. In practice, we obtain the resistance of both hot wires at the cell temperature and pressure for every experiment. The calibration process is an integral part of balancing the bridge. As described above, we have the capability to make a four-terminal resistance measurement of each hot wire without errors from the temperature-dependent lead resistance. When we have completed all measurements on a given fluid, we do a surface fit of the resistance of each wire using the functional form above. Examining trends in deviations from this surface fit helps us to detect inconsistent data. Slow changes in the calibration usually indicate changes in the physical condition of the hot wires, such as continued annealing of the platinum at high temperatures. Sudden changes in the wire calibration provide an indication of mechanical damage to the wires. In addition, the capability to generate an *in situ* calibration provides freedom to use materials other than platinum for the hot wires.

5. Performance Verification

Toluene was selected to verify the instrument performance in the liquid phase since it has been recently recommended as a thermal conductivity reference standard [13]. Argon and nitrogen were selected to verify performance of the apparatus in the gas phase since they have been widely studied with both steady-state techniques and transient hot-wire instruments. In addition, they have been studied with our low temperature instrument so that discrepancies between the two instruments can be detected and resolved.

5.1 Toluene

The thermal conductivity of liquid toluene has been widely studied with both steady-state and transient hot-wire instruments for a number of years. Early steady-state experiments on toluene were often plagued by free convection. Free convection is easily avoided in a transient hot-wire instrument, but, if present, is easily detected due to deviations from the ideal line-source model. The contribution of thermal radiation to the apparent thermal conductivity of toluene has also been of much concern since toluene is not transparent in the infrared. Nieto de Castro et al. [7] have made an extensive study of thermal radiation and concluded that the radiative contribution to heat transfer is very small for toluene at temperatures up to 370 K. Above 370 K, it was estimated that the contribution of heat transport by radiation to the measured value of thermal conductivity would increase with temperature resulting in nonzero values of the quantity B in eq (13). Toluene was selected to verify both the performance of the new instrument in the liquid phase and the size and effect of the radiative contribution at the higher temperatures.

The spectroscopic grade toluene used in our verification measurements was dried over calcium hydride and distilled to remove a trace of benzene impurity. The purified toluene was analyzed by gas chromatography and found to have less than 50 parts per billion (ppb) benzene and less than 100 ppb water. The results of the saturated liquid toluene tests are provided in table 1. In order to

Table 1. Thermal conductivity, thermal diffusivity, and heat capacity of liquid toluene from 300 to 550 K

Run Pt.	Pressure MPa	Temperature K	Density mol/L	Power W/m	Thermal conductivity W/(m·K) STAT	Cell temperature K	Thermal diffusivity m^2/s DSTAT	Specific heat J/(mol·K)
1101	0.090	302.441	9.3100	0.94731	0.12880 0.000	297.811	0.866×10^{-7} 0.003	160.6
1102	0.090	302.128	9.3131	0.87943	0.12918 0.000	297.840	0.880×10^{-7} 0.003	158.4
1103	0.090	301.835	9.3160	0.81210	0.12885 0.000	297.872	0.861×10^{-7} 0.004	161.3
1104	0.090	301.549	9.3188	0.74835	0.12865 0.000	297.869	0.852×10^{-7} 0.004	162.7
1105	0.090	301.265	9.3215	0.68819	0.12868 0.001	297.891	0.858×10^{-7} 0.005	161.5
1201	0.088	328.877	9.0466	0.96177	0.12145 0.000	323.969	0.791×10^{-7} 0.003	170.8
1202	0.088	328.528	9.0501	0.88883	0.12132 0.000	324.039	0.789×10^{-7} 0.003	170.9
1203	0.088	328.174	9.0537	0.81994	0.12144 0.000	324.005	0.792×10^{-7} 0.004	170.2
1204	0.088	327.841	9.0571	0.75378	0.12146 0.000	323.979	0.788×10^{-7} 0.004	171.1
1205	0.088	327.506	9.0605	0.69063	0.12166 0.001	323.937	0.790×10^{-7} 0.005	170.7
1206	0.089	327.180	9.0638	0.62968	0.12167 0.001	323.951	0.786×10^{-7} 0.006	171.5
1207	0.088	326.927	9.0664	0.57154	0.12171 0.001	323.951	0.784×10^{-7} 0.007	171.9
1208	0.088	326.638	9.0693	0.51693	0.12210 0.001	323.927	0.795×10^{-7} 0.007	170.0
1301	0.086	347.212	8.8572	0.45110	0.11626 0.001	344.798	0.744×10^{-7} 0.008	177.1
1302	0.086	349.371	8.8346	0.86369	0.11571 0.000	344.881	0.745×10^{-7} 0.003	176.8
1303	0.086	349.037	8.8381	0.79636	0.11573 0.000	344.778	0.748×10^{-7} 0.003	176.0
1304	0.086	348.691	8.8417	0.73227	0.11595 0.000	344.785	0.752×10^{-7} 0.004	175.3
1305	0.086	348.357	8.8452	0.67060	0.11595 0.000	344.788	0.749×10^{-7} 0.004	175.8
1306	0.086	348.045	8.8485	0.61146	0.11607 0.001	344.805	0.737×10^{-7} 0.005	178.7
1307	0.086	347.756	8.8515	0.55506	0.11610 0.001	344.821	0.734×10^{-7} 0.006	179.3
1308	0.086	347.484	8.8544	0.50142	0.11613 0.001	344.838	0.737×10^{-7} 0.006	178.6
1401	0.129	372.502	8.5863	0.77696	0.10982 0.000	368.375	0.701×10^{-7} 0.003	183.7
1402	0.129	372.132	8.5904	0.71422	0.10971 0.000	368.341	0.692×10^{-7} 0.004	185.7
1403	0.129	371.809	8.5940	0.65399	0.10981 0.001	368.273	0.689×10^{-7} 0.004	186.4
1404	0.129	371.503	8.5973	0.59648	0.10998 0.001	368.297	0.694×10^{-7} 0.005	185.3
1405	0.129	371.206	8.6006	0.54161	0.11008 0.001	368.308	0.695×10^{-7} 0.006	185.0
1406	0.129	370.932	8.6036	0.48963	0.11010 0.001	368.298	0.689×10^{-7} 0.006	186.3
1407	0.129	370.634	8.6068	0.43960	0.11006 0.001	368.382	0.685×10^{-7} 0.008	187.4
1408	0.129	370.375	8.6097	0.39278	0.11009 0.001	368.395	0.675×10^{-7} 0.009	190.0
1501	0.391	405.204	8.2181	0.71525	0.10192 0.000	401.334	0.646×10^{-7} 0.004	193.4
1502	0.402	404.847	8.2225	0.65718	0.10181 0.001	401.354	0.632×10^{-7} 0.005	197.2
1503	0.410	404.522	8.2265	0.60151	0.10213 0.000	401.323	0.648×10^{-7} 0.004	192.6
1504	0.415	404.211	8.2302	0.54869	0.10213 0.001	401.382	0.640×10^{-7} 0.006	194.9
1505	0.425	403.927	8.2338	0.49827	0.10236 0.001	401.345	0.650×10^{-7} 0.006	192.2
1506	0.433	403.654	8.2371	0.45015	0.10227 0.001	401.301	0.647×10^{-7} 0.006	192.8
1507	0.439	403.391	8.2403	0.40460	0.10247 0.001	401.345	0.658×10^{-7} 0.007	189.6
1508	0.446	405.536	8.2152	0.77476	0.10197 0.000	401.315	0.651×10^{-7} 0.003	192.0
1601	0.554	426.783	7.9589	0.81341	0.09762 0.000	422.459	0.621×10^{-7} 0.003	199.3
1602	0.558	426.430	7.9634	0.75186	0.09735 0.000	422.522	0.608×10^{-7} 0.003	202.5
1603	0.561	426.075	7.9679	0.69337	0.09752 0.000	422.539	0.616×10^{-7} 0.003	199.9
1604	0.561	425.732	7.9722	0.63755	0.09761 0.000	422.454	0.612×10^{-7} 0.004	201.3
1605	0.562	425.405	7.9763	0.58387	0.09758 0.000	422.551	0.610×10^{-7} 0.004	201.7
1606	0.564	425.096	7.9802	0.53246	0.09763 0.001	422.582	0.606×10^{-7} 0.005	202.9
1607	0.566	424.796	7.9840	0.48350	0.09791 0.001	422.497	0.618×10^{-7} 0.005	199.3
1608	0.569	424.523	7.9875	0.43695	0.09802 0.001	422.474	0.620×10^{-7} 0.006	198.9
1701	0.768	453.898	7.6094	0.86040	0.09206 0.000	449.481	0.584×10^{-7} 0.003	209.1
1702	0.770	453.514	7.6147	0.80056	0.09215 0.000	449.510	0.586×10^{-7} 0.003	208.5
1703	0.772	453.140	7.6199	0.74281	0.09221 0.000	449.496	0.589×10^{-7} 0.003	207.2
1704	0.776	452.793	7.6247	0.68722	0.09222 0.000	449.568	0.586×10^{-7} 0.004	207.9
1705	0.779	452.439	7.6297	0.63370	0.09241 0.000	449.538	0.589×10^{-7} 0.003	207.2
1706	0.781	452.119	7.6341	0.58236	0.09224 0.001	449.544	0.583×10^{-7} 0.005	208.6
1707	0.782	451.799	7.6385	0.53352	0.09243 0.001	449.527	0.586×10^{-7} 0.005	207.6
1708	0.782	451.521	7.6423	0.48659	0.09244 0.001	449.507	0.584×10^{-7} 0.006	208.2

Table 1. Thermal conductivity, thermal diffusivity, and heat capacity of liquid toluene from 300 to 550 K—Continued

Run Pt.	Pressure MPa	Temperature K	Density mol/L	Power W/m	Thermal conductivity W/(m·K) STAT	Cell temperature K	Thermal diffusivity m^2/s DSTAT	Specific heat J/(mol·K)
1801	1.023	480.654	7.2293	0.84664	0.08720 0.000	473.650	0.560×10^{-7} 0.003	217.4
1802	1.018	480.279	7.2349	0.78957	0.08709 0.000	473.684	0.561×10^{-7} 0.003	216.6
1803	1.016	479.914	7.2405	0.73476	0.08728 0.000	473.684	0.565×10^{-7} 0.004	215.2
1804	1.014	479.555	7.2459	0.68157	0.08738 0.000	473.713	0.568×10^{-7} 0.003	213.9
1805	1.014	479.212	7.2512	0.63048	0.08739 0.000	473.808	0.568×10^{-7} 0.004	213.7
1806	1.009	478.875	7.2562	0.58183	0.08742 0.001	473.789	0.559×10^{-7} 0.004	217.0
1807	1.008	478.564	7.2609	0.53447	0.08754 0.001	473.801	0.569×10^{-7} 0.005	213.1
1808	1.008	478.262	7.2655	0.48937	0.08764 0.001	473.753	0.564×10^{-7} 0.005	215.3
1901	1.681	504.349	6.8718	0.88755	0.08337 0.000	497.373	0.554×10^{-7} 0.003	221.5
1902	1.673	503.948	6.8784	0.82922	0.08338 0.000	497.373	0.560×10^{-7} 0.003	218.8
1903	1.669	503.572	6.8847	0.77381	0.08352 0.000	497.404	0.566×10^{-7} 0.003	216.3
1904	1.662	503.194	6.8909	0.71986	0.08347 0.000	497.340	0.565×10^{-7} 0.003	216.4
1905	1.659	502.816	6.8973	0.66827	0.08356 0.000	497.278	0.551×10^{-7} 0.003	221.6
1906	1.654	502.464	6.9031	0.61856	0.08366 0.000	497.298	0.554×10^{-7} 0.004	220.4
1907	1.651	502.127	6.9088	0.57068	0.08375 0.001	497.277	0.553×10^{-7} 0.004	220.8
1908	1.647	501.825	6.9138	0.52450	0.08380 0.001	497.309	0.558×10^{-7} 0.005	218.5
2001	2.293	526.378	6.4964	0.87427	0.08046 0.000	519.727	0.542×10^{-7} 0.003	230.8
2002	2.295	525.971	6.5049	0.81718	0.08046 0.000	519.611	0.542×10^{-7} 0.003	230.5
2003	2.297	525.577	6.5132	0.76206	0.08063 0.000	519.568	0.549×10^{-7} 0.003	227.8
2004	2.299	525.193	6.5212	0.70904	0.08059 0.000	519.685	0.547×10^{-7} 0.004	228.1
2005	2.301	524.830	6.5287	0.65793	0.08079 0.000	519.738	0.547×10^{-7} 0.004	228.2
2006	2.302	524.479	6.5359	0.60865	0.08078 0.000	519.685	0.545×10^{-7} 0.004	228.5
2007	2.305	524.144	6.5429	0.56137	0.08089 0.001	519.685	0.550×10^{-7} 0.004	226.4
2008	2.306	523.830	6.5493	0.51601	0.08103 0.001	519.706	0.558×10^{-7} 0.006	223.1
2101	2.682	554.337	5.8477	0.86413	0.07653 0.001	548.030	0.516×10^{-7} 0.007	254.8
2102	2.683	553.911	5.8611	0.80784	0.07659 0.001	548.130	0.517×10^{-7} 0.007	253.6
2103	2.684	553.515	5.8736	0.75366	0.07672 0.001	548.063	0.523×10^{-7} 0.008	250.8
2104	2.686	553.111	5.8862	0.70126	0.07680 0.001	548.109	0.524×10^{-7} 0.006	250.2
2105	2.686	552.722	5.8981	0.65069	0.07686 0.001	548.140	0.522×10^{-7} 0.009	250.7
2106	2.688	552.367	5.9089	0.60212	0.07703 0.001	548.129	0.533×10^{-7} 0.009	245.6
2107	2.688	552.041	5.9187	0.55544	0.07702 0.001	548.132	0.532×10^{-7} 0.010	245.6
2108	2.691	551.695	5.9294	0.51041	0.07709 0.001	548.141	0.534×10^{-7} 0.011	244.1
2109	19.346	553.364	6.9566	0.86455	0.09108 0.000	548.247	0.632×10^{-7} 0.003	211.7
2110	19.353	552.979	6.9613	0.80814	0.09119 0.000	548.194	0.633×10^{-7} 0.004	211.2
2111	19.357	552.629	6.9654	0.75383	0.09120 0.000	548.217	0.631×10^{-7} 0.004	211.5
2112	19.358	552.276	6.9695	0.70146	0.09107 0.001	548.226	0.619×10^{-7} 0.004	214.9
2113	14.335	553.521	6.7600	0.86425	0.08774 0.000	548.096	0.607×10^{-7} 0.003	218.3
2114	14.335	553.155	6.7648	0.80806	0.08775 0.000	548.075	0.608×10^{-7} 0.003	217.4
2115	14.335	552.794	6.7695	0.75378	0.08789 0.001	548.132	0.613×10^{-7} 0.004	215.5
2116	14.336	552.424	6.7744	0.70141	0.08778 0.001	548.153	0.603×10^{-7} 0.005	218.6
2117	9.599	553.745	6.5167	0.86414	0.08408 0.000	548.099	0.586×10^{-7} 0.004	224.4
2118	9.599	553.363	6.5227	0.80789	0.08418 0.000	548.088	0.592×10^{-7} 0.004	221.9
2119	9.599	553.003	6.5283	0.75370	0.08423 0.001	548.045	0.598×10^{-7} 0.004	219.2
2120	9.598	552.630	6.5341	0.70139	0.08420 0.001	548.065	0.592×10^{-7} 0.004	221.1
2121	6.941	553.927	6.3343	0.86406	0.08179 0.000	548.065	0.579×10^{-7} 0.004	226.7
2122	6.941	553.531	6.3415	0.80788	0.08192 0.000	548.078	0.587×10^{-7} 0.004	223.6
2123	6.941	553.133	6.3487	0.75367	0.08202 0.001	548.045	0.589×10^{-7} 0.005	222.4
2124	6.942	552.777	6.3552	0.70128	0.08199 0.001	548.065	0.588×10^{-7} 0.005	222.5
2125	4.512	554.119	6.1081	0.86378	0.07885 0.000	548.033	0.550×10^{-7} 0.004	237.6
2126	4.512	553.711	6.1173	0.80730	0.07890 0.001	548.023	0.556×10^{-7} 0.004	234.7
2127	4.512	553.323	6.1261	0.75307	0.07904 0.000	548.033	0.567×10^{-7} 0.004	230.2
2128	4.512	552.936	6.1348	0.70093	0.07910 0.001	548.034	0.565×10^{-7} 0.004	230.4

Table 2. Thermal conductivity, thermal diffusivity, and heat capacity of argon gas near 300 K

Run Pt.	Pressure MPa	Temperature K	Density mol/L	Power W/m	Thermal conductivity W/(m·K) STAT	Cell temperature K	Thermal diffusivity m^2/s DSTAT	Specific heat J/(mol·K)
1001	65.224	301.925	19.9271	0.33737	0.05378 0.000	298.178	0.795×10^{-7} 0.004	33.9
1002	65.224	301.453	19.9550	0.29703	0.05381 0.000	298.145	0.772×10^{-7} 0.004	34.9
1003	65.224	301.021	19.9805	0.25949	0.05392 0.001	298.188	0.768×10^{-7} 0.005	35.1
1004	65.223	300.621	20.0041	0.22432	0.05404 0.001	298.152	0.767×10^{-7} 0.006	35.2
1005	60.534	301.982	19.1757	0.33733	0.05128 0.000	298.079	0.772×10^{-7} 0.004	34.6
1006	60.533	301.554	19.2011	0.29699	0.05131 0.001	298.069	0.774×10^{-7} 0.005	34.5
1007	60.533	301.101	19.2281	0.25926	0.05136 0.001	298.067	0.771×10^{-7} 0.005	34.6
1008	60.531	300.698	19.2520	0.22406	0.05141 0.001	298.105	0.768×10^{-7} 0.006	34.7
1009	56.254	301.663	18.4516	0.29684	0.04899 0.000	298.073	0.753×10^{-7} 0.004	35.1
1010	56.254	301.195	18.4801	0.25917	0.04905 0.001	298.029	0.758×10^{-7} 0.005	34.9
1011	56.254	300.770	18.5058	0.22406	0.04907 0.001	298.081	0.753×10^{-7} 0.006	35.1
1012	56.256	300.350	18.5315	0.19174	0.04925 0.001	298.081	0.757×10^{-7} 0.007	35.1
1013	52.289	301.803	17.6955	0.29714	0.04683 0.001	298.017	0.750×10^{-7} 0.004	35.2
1014	52.289	301.333	17.7241	0.25931	0.04687 0.001	298.049	0.755×10^{-7} 0.005	34.9
1015	52.289	300.877	17.7519	0.22425	0.04696 0.001	298.047	0.750×10^{-7} 0.006	35.2
1016	52.288	300.462	17.7771	0.19162	0.04702 0.001	298.051	0.746×10^{-7} 0.007	35.4
1017	48.788	301.940	16.9726	0.29698	0.04486 0.000	298.026	0.749×10^{-7} 0.004	35.1
1018	48.789	301.451	17.0025	0.25927	0.04492 0.001	298.000	0.752×10^{-7} 0.004	35.0
1019	48.789	300.978	17.0314	0.22414	0.04496 0.001	298.019	0.749×10^{-7} 0.005	35.1
1020	48.788	300.553	17.0573	0.19154	0.04504 0.001	298.031	0.751×10^{-7} 0.007	35.0
1021	45.435	302.110	16.2241	0.29672	0.04306 0.000	298.010	0.769×10^{-7} 0.003	34.3
1022	45.435	301.595	16.2553	0.25911	0.04304 0.001	297.992	0.763×10^{-7} 0.004	34.5
1023	45.435	301.120	16.2842	0.22393	0.04302 0.001	298.020	0.762×10^{-7} 0.005	34.5
1024	45.435	300.652	16.3128	0.19141	0.04308 0.001	297.990	0.765×10^{-7} 0.006	34.4
1025	42.251	301.754	15.4902	0.25937	0.04117 0.001	298.003	0.749×10^{-7} 0.005	35.2
1026	42.251	301.239	15.5212	0.22413	0.04119 0.001	297.997	0.742×10^{-7} 0.005	35.5
1027	42.249	300.762	15.5498	0.19151	0.04127 0.001	298.024	0.736×10^{-7} 0.006	35.9
1028	42.249	300.317	15.5767	0.16149	0.04131 0.001	297.949	0.727×10^{-7} 0.008	36.4
1029	39.526	301.865	14.7903	0.25936	0.03955 0.000	298.007	0.749×10^{-7} 0.004	35.4
1030	39.526	301.334	14.8219	0.22432	0.03959 0.001	297.965	0.751×10^{-7} 0.005	35.3
1031	39.526	300.853	14.8508	0.19179	0.03965 0.001	297.975	0.751×10^{-7} 0.006	35.3
1032	39.525	300.389	14.8785	0.16183	0.03975 0.001	298.012	0.748×10^{-7} 0.008	35.5
1033	36.708	302.037	14.0145	0.25981	0.03800 0.000	298.005	0.774×10^{-7} 0.004	34.7
1034	36.708	301.493	14.0464	0.22461	0.03791 0.001	298.010	0.762×10^{-7} 0.005	35.1
1035	36.710	300.960	14.0781	0.19185	0.03794 0.001	298.041	0.759×10^{-7} 0.005	35.3
1036	36.710	300.488	14.1059	0.16175	0.03798 0.001	298.019	0.756×10^{-7} 0.007	35.4
1037	33.968	302.077	13.2166	0.25240	0.03622 0.001	297.990	0.773×10^{-7} 0.005	35.0
1038	33.968	301.609	13.2432	0.22419	0.03627 0.001	297.950	0.777×10^{-7} 0.005	34.9
1039	33.969	301.175	13.2683	0.19786	0.03626 0.001	297.929	0.779×10^{-7} 0.005	34.8
1040	33.969	300.784	13.2907	0.17318	0.03627 0.001	297.930	0.773×10^{-7} 0.007	35.1
1041	31.748	302.007	12.5367	0.23797	0.03491 0.000	297.938	0.806×10^{-7} 0.004	34.1
1042	31.748	301.535	12.5629	0.21097	0.03489 0.001	297.946	0.795×10^{-7} 0.005	34.6
1043	31.748	301.096	12.5873	0.18554	0.03492 0.001	297.926	0.795×10^{-7} 0.006	34.6
1044	31.749	300.678	12.6109	0.16143	0.03497 0.001	297.941	0.796×10^{-7} 0.007	34.5
1045	29.280	301.947	11.7377	0.22396	0.03347 0.000	297.936	0.839×10^{-7} 0.004	33.6
1046	29.281	301.476	11.7630	0.19797	0.03353 0.001	297.935	0.849×10^{-7} 0.005	33.2
1047	29.281	301.044	11.7859	0.17332	0.03355 0.001	297.909	0.840×10^{-7} 0.006	33.5
1048	29.282	300.612	11.8092	0.15021	0.03347 0.001	297.956	0.826×10^{-7} 0.007	34.0
1049	27.318	301.897	11.0699	0.21088	0.03216 0.000	297.971	0.836×10^{-7} 0.004	34.3
1050	27.319	301.427	11.0940	0.18535	0.03212 0.001	297.945	0.830×10^{-7} 0.005	34.5
1051	27.319	300.963	11.1178	0.16128	0.03217 0.001	297.930	0.832×10^{-7} 0.006	34.4
1052	27.320	300.558	11.1389	0.13912	0.03225 0.001	297.957	0.836×10^{-7} 0.008	34.3
1053	25.276	302.085	10.3335	0.21027	0.03091 0.001	297.951	0.890×10^{-7} 0.004	33.1

Table 2. Thermal conductivity, thermal diffusivity, and heat capacity of argon gas near 300 K—Continued

Run Pt.	Pressure MPa	Temperature K	Density mol/L	Power W/m	Thermal conductivity W/(m·K) STAT	Cell temperature K	Thermal diffusivity m^2/s DSTAT	Specific heat J/(mol·K)
1054	25.276	301.603	10.3566	0.18491	0.03092 0.001	297.979	0.884×10^{-7} 0.005	33.3
1055	25.276	301.123	10.3796	0.16112	0.03089 0.001	297.932	0.879×10^{-7} 0.006	33.4
1056	25.276	300.690	10.4005	0.13915	0.03096 0.001	297.956	0.881×10^{-7} 0.007	33.4
1057	23.207	302.038	9.5687	0.19758	0.02966 0.000	298.009	0.928×10^{-7} 0.004	32.8
1058	23.207	301.529	9.5916	0.17297	0.02969 0.001	297.980	0.936×10^{-7} 0.005	32.6
1059	23.208	301.072	9.6124	0.14995	0.02968 0.001	298.006	0.933×10^{-7} 0.006	32.7
1060	23.208	300.634	9.6321	0.12854	0.02967 0.001	298.007	0.926×10^{-7} 0.008	32.9
1061	21.499	301.961	8.9175	0.18527	0.02868 0.001	297.969	0.100×10^{-6} 0.005	31.4
1062	21.499	301.451	8.9387	0.16130	0.02862 0.001	297.983	0.991×10^{-7} 0.006	31.8
1063	21.499	301.206	8.9490	0.14993	0.02864 0.001	297.953	0.998×10^{-7} 0.006	31.6
1064	21.499	300.536	8.9771	0.11843	0.02867 0.001	297.978	0.100×10^{-6} 0.009	31.5
1065	19.660	301.868	8.1962	0.17302	0.02753 0.001	297.975	0.101×10^{-6} 0.005	32.6
1066	19.660	301.342	8.2163	0.14986	0.02753 0.001	297.986	0.100×10^{-6} 0.006	32.9
1067	19.660	300.856	8.2350	0.12845	0.02753 0.001	297.981	0.989×10^{-7} 0.008	33.4
1068	19.662	300.399	8.2532	0.10874	0.02756 0.001	298.003	0.976×10^{-7} 0.009	33.8
1069	17.864	302.035	7.4644	0.17310	0.02643 0.001	298.005	0.111×10^{-6} 0.006	31.4
1070	17.864	301.489	7.4833	0.15007	0.02642 0.001	297.966	0.109×10^{-6} 0.006	31.7
1071	17.864	300.981	7.5010	0.12874	0.02644 0.001	297.933	0.109×10^{-6} 0.007	32.0
1072	17.864	300.516	7.5173	0.10900	0.02647 0.001	297.960	0.108×10^{-6} 0.010	32.3
1073	16.141	302.241	6.7481	0.17324	0.02545 0.001	297.971	0.121×10^{-6} 0.005	30.5
1074	16.141	301.664	6.7659	0.15023	0.02542 0.001	297.956	0.119×10^{-6} 0.006	30.9
1075	16.141	301.119	6.7828	0.12876	0.02544 0.001	297.989	0.120×10^{-6} 0.007	30.8
1076	16.141	300.607	6.7988	0.10905	0.02546 0.001	297.956	0.119×10^{-6} 0.009	31.0
1077	14.442	302.139	6.0414	0.16149	0.02446 0.001	297.970	0.133×10^{-6} 0.005	29.7
1078	14.442	301.576	6.0568	0.13912	0.02445 0.001	297.955	0.133×10^{-6} 0.006	29.7
1079	14.442	301.049	6.0712	0.11858	0.02442 0.001	297.977	0.132×10^{-6} 0.007	30.1
1080	14.442	300.563	6.0846	0.09963	0.02446 0.001	297.960	0.132×10^{-6} 0.010	30.1
1081	12.754	302.079	5.3298	0.14979	0.02351 0.001	297.927	0.151×10^{-6} 0.006	28.6
1082	12.754	301.496	5.3436	0.12843	0.02348 0.001	297.950	0.149×10^{-6} 0.006	29.0
1083	12.754	300.934	5.3569	0.10882	0.02352 0.001	298.002	0.149×10^{-6} 0.008	29.0
1084	12.754	300.446	5.3685	0.09075	0.02349 0.001	298.003	0.148×10^{-6} 0.011	29.2
1085	10.898	302.004	4.5421	0.13917	0.02253 0.001	297.935	0.172×10^{-6} 0.006	28.1
1086	10.898	301.417	4.5539	0.11851	0.02251 0.001	297.945	0.171×10^{-6} 0.007	28.4
1087	10.898	300.864	4.5648	0.09964	0.02251 0.001	297.996	0.169×10^{-6} 0.009	28.7
1088	10.898	300.355	4.5748	0.08236	0.02251 0.001	297.951	0.166×10^{-6} 0.012	29.3
1089	9.295	302.244	3.8552	0.13887	0.02171 0.001	297.960	0.205×10^{-6} 0.006	26.8
1090	9.295	301.611	3.8654	0.11841	0.02167 0.001	297.949	0.202×10^{-6} 0.007	27.2
1091	9.295	301.033	3.8748	0.09956	0.02166 0.001	297.938	0.200×10^{-6} 0.009	27.5
1092	9.294	300.501	3.8832	0.08235	0.02169 0.001	297.979	0.199×10^{-6} 0.012	27.7
1093	7.510	302.191	3.0984	0.12862	0.02085 0.001	297.991	0.247×10^{-6} 0.006	26.5
1093	7.510	301.544	3.1064	0.10882	0.02082 0.001	297.941	0.243×10^{-6} 0.008	27.0
1095	7.510	300.939	3.1138	0.09061	0.02083 0.001	297.995	0.245×10^{-6} 0.010	26.8
1096	7.509	300.403	3.1203	0.07423	0.02084 0.001	297.975	0.244×10^{-6} 0.014	27.0
1097	5.829	302.136	2.3897	0.11818	0.02008 0.001	297.999	0.324×10^{-6} 0.007	25.2
1098	5.828	301.479	2.3955	0.09942	0.02008 0.001	297.970	0.324×10^{-6} 0.009	25.3
1099	5.827	300.879	2.4009	0.08221	0.02004 0.001	297.970	0.319×10^{-6} 0.012	25.7
1100	5.827	300.298	2.4060	0.06671	0.02009 0.002	297.950	0.307×10^{-6} 0.015	26.8
1101	4.375	302.061	1.7826	0.10864	0.01960 0.001	297.985	0.429×10^{-6} 0.008	25.0
1102	4.375	301.373	1.7872	0.09059	0.01952 0.001	297.980	0.413×10^{-6} 0.011	25.9
1103	4.375	300.743	1.7915	0.07418	0.01951 0.001	297.997	0.403×10^{-6} 0.014	26.5
1104	4.375	300.195	1.7952	0.05939	0.01948 0.002	298.010	0.388×10^{-6} 0.019	27.6
1105	2.601	302.360	1.0495	0.10871	0.01893 0.001	297.970	0.693×10^{-6} 0.010	25.3
1106	2.601	301.634	1.0522	0.09071	0.01887 0.001	297.950	0.676×10^{-6} 0.011	25.9

Table 2. Thermal conductivity, thermal diffusivity, and heat capacity of argon gas near 300 K—Continued

Run Pt.	Pressure MPa	Temperature K	Density mol/L	Power W/m	Thermal conductivity W/(m·K) STAT	Cell temperature K	Thermal diffusivity m^2/s DSTAT	Specific heat J/(mol·K)
1107	2.600	300.980	1.0544	0.07435	0.01886 0.001	297.999	0.655×10^{-6} 0.014	26.8
1108	2.600	300.356	1.0567	0.05952	0.01885 0.002	297.976	0.631×10^{-6} 0.018	27.9
1109	65.509	302.887	19.9141	0.41761	0.05374 0.000	298.353	0.761×10^{-7} 0.004	35.4
1110	65.499	302.474	19.9368	0.38071	0.05378 0.000	298.313	0.759×10^{-7} 0.004	35.4
1111	65.484	302.092	19.9570	0.34638	0.05388 0.000	298.261	0.778×10^{-7} 0.004	34.7
1112	65.474	301.728	19.9769	0.31306	0.05387 0.001	298.282	0.778×10^{-7} 0.004	34.6
1113	65.468	301.382	19.9964	0.28165	0.05389 0.001	298.266	0.780×10^{-7} 0.005	34.5
1114	65.454	301.034	20.0147	0.25238	0.05396 0.001	298.257	0.765×10^{-7} 0.006	35.2
1115	65.447	300.704	20.0331	0.22404	0.05408 0.001	298.216	0.767×10^{-7} 0.007	35.1
1116	65.437	300.433	20.0478	0.19763	0.05405 0.001	298.226	0.757×10^{-7} 0.008	35.6
1117	65.430	300.168	20.0624	0.17284	0.05403 0.001	298.269	0.780×10^{-7} 0.009	34.5
1118	65.422	299.903	20.0769	0.14987	0.05407 0.001	298.303	0.764×10^{-7} 0.011	35.2
1119	65.414	299.654	20.0905	0.12847	0.05410 0.002	298.231	0.748×10^{-7} 0.014	36.0
1120	65.407	299.444	20.1018	0.10873	0.05413 0.002	298.267	0.733×10^{-7} 0.017	36.7
1121	65.353	299.421	20.0951	0.10852	0.05406 0.002	298.212	0.756×10^{-7} 0.016	35.6
1122	65.348	299.641	20.0812	0.12835	0.05405 0.002	298.269	0.773×10^{-7} 0.014	34.8
1123	65.338	299.869	20.0662	0.14973	0.05401 0.001	298.227	0.766×10^{-7} 0.011	35.1
1124	65.331	300.121	20.0501	0.17277	0.05405 0.001	298.222	0.783×10^{-7} 0.009	34.4
1125	65.322	300.404	20.0319	0.19739	0.05395 0.001	298.222	0.780×10^{-7} 0.008	34.5
1126	65.314	300.682	20.0143	0.22362	0.05389 0.001	298.219	0.776×10^{-7} 0.007	34.7
1127	65.306	301.000	19.9942	0.25163	0.05376 0.001	298.224	0.770×10^{-7} 0.006	34.9
1128	65.299	301.327	19.9738	0.28120	0.05370 0.001	298.217	0.770×10^{-7} 0.005	34.9
1129	65.290	301.675	19.9519	0.31240	0.05367 0.001	298.244	0.770×10^{-7} 0.005	34.9
1130	65.282	302.034	19.9296	0.34525	0.05364 0.000	298.232	0.771×10^{-7} 0.004	34.8
1131	65.273	302.431	19.9049	0.37948	0.05352 0.001	298.212	0.767×10^{-7} 0.005	35.0
1132	65.265	302.841	19.8796	0.41638	0.05343 0.000	298.206	0.762×10^{-7} 0.004	35.2
1133	65.220	302.836	19.8731	0.41590	0.05346 0.000	298.170	0.775×10^{-7} 0.004	34.6
1134	65.213	302.439	19.8953	0.37977	0.05347 0.000	298.182	0.776×10^{-7} 0.004	34.6
1135	65.207	302.037	19.9180	0.34562	0.05357 0.000	298.201	0.777×10^{-7} 0.004	34.5
1136	65.200	301.673	19.9383	0.31300	0.05366 0.001	298.244	0.776×10^{-7} 0.005	34.6
1137	65.191	301.313	19.9583	0.28163	0.05372 0.001	298.227	0.777×10^{-7} 0.005	34.6
1138	65.185	300.992	19.9762	0.25205	0.05372 0.001	298.244	0.772×10^{-7} 0.006	34.8
1139	65.176	300.680	19.9934	0.22382	0.05377 0.001	298.224	0.771×10^{-7} 0.007	34.8
1140	65.170	300.369	20.0108	0.19759	0.05385 0.001	298.162	0.769×10^{-7} 0.008	35.0
1141	65.164	300.101	20.0259	0.17300	0.05388 0.001	298.207	0.765×10^{-7} 0.009	35.1
1142	65.159	299.841	20.0405	0.14996	0.05389 0.001	298.195	0.760×10^{-7} 0.011	35.3
1143	65.154	299.607	20.0537	0.12861	0.05404 0.002	298.214	0.760×10^{-7} 0.014	35.5
1144	65.146	299.385	20.0658	0.10884	0.05402 0.002	298.173	0.733×10^{-7} 0.017	36.7

Table 3. Thermal conductivity, thermal diffusivity, and heat capacity of nitrogen gas near 425 K

Run Pt.	Pressure MPa	Temperature K	Density mol/L	Power W/m	Thermal conductivity W/(m·K) STAT	Cell temperature K	Thermal diffusivity m^2/s DSTAT	Specific heat J/(mol·K)
4001	67.472	426.243	13.0983	0.41718	0.06062 0.001	421.862	0.122×10^{-6} 0.008	37.8
4002	67.472	425.770	13.1098	0.37241	0.06062 0.001	421.855	0.121×10^{-6} 0.009	38.1
4003	67.471	425.332	13.1205	0.33045	0.06063 0.001	421.844	0.122×10^{-6} 0.011	37.8
4004	67.470	424.918	13.1305	0.29096	0.06059 0.002	421.887	0.120×10^{-6} 0.014	38.1
4005	58.114	426.577	11.9010	0.41738	0.05671 0.001	421.881	0.125×10^{-6} 0.008	37.8
4006	58.113	426.072	11.9128	0.37279	0.05672 0.001	421.854	0.125×10^{-6} 0.010	37.8
4007	58.113	425.597	11.9239	0.33059	0.05668 0.001	421.851	0.124×10^{-6} 0.011	38.1

Table 3. Thermal conductivity, thermal diffusivity, and heat capacity of nitrogen gas near 425 K

Run Pt.	Pressure MPa	Temperature K	Density mol/L	Power W/m	Thermal conductivity W/(m·K)	STAT	Cell temperature K	Thermal diffusivity m^2/s	DSTAT	Specific heat J/(mol·K)
4008	58.112	425.159	11.9341	0.29111	0.05670	0.001	421.932	0.124×10^{-6}	0.013	38.0
4009	53.330	426.824	11.2302	0.41741	0.05472	0.001	421.932	0.129×10^{-6}	0.009	37.5
4010	53.331	426.299	11.2424	0.37276	0.05475	0.001	421.896	0.129×10^{-6}	0.008	37.5
4011	53.331	425.804	11.2537	0.33063	0.05473	0.001	421.910	0.128×10^{-6}	0.011	37.6
4012	53.330	425.351	11.2640	0.29114	0.05469	0.001	421.893	0.128×10^{-6}	0.012	37.8
4013	48.692	427.059	10.5387	0.41713	0.05282	0.001	421.976	0.134×10^{-6}	0.009	36.9
4014	48.692	426.504	10.5510	0.37264	0.05285	0.001	421.986	0.134×10^{-6}	0.009	37.0
4015	48.694	426.006	10.5622	0.33066	0.05278	0.001	421.980	0.133×10^{-6}	0.010	37.2
4016	48.695	425.535	10.5728	0.29113	0.05275	0.001	421.999	0.132×10^{-6}	0.012	37.4
4017	43.950	427.314	9.7851	0.41691	0.05090	0.001	421.968	0.141×10^{-6}	0.008	36.3
4018	43.950	426.757	9.7970	0.37232	0.05088	0.001	422.007	0.141×10^{-6}	0.008	36.5
4019	43.951	426.234	9.8081	0.33020	0.05082	0.001	421.962	0.140×10^{-6}	0.010	36.5
4020	43.952	425.736	9.8188	0.29078	0.05084	0.001	421.988	0.142×10^{-6}	0.012	36.2
4021	39.833	427.590	9.0889	0.41683	0.04918	0.001	422.036	0.150×10^{-6}	0.008	35.6
4022	39.833	426.999	9.1006	0.37243	0.04916	0.001	422.048	0.149×10^{-6}	0.009	35.7
4023	39.833	426.441	9.1120	0.33052	0.04915	0.001	422.045	0.149×10^{-6}	0.010	35.8
4024	39.833	425.924	9.1224	0.29100	0.04915	0.001	422.039	0.150×10^{-6}	0.012	35.6
4025	35.790	427.900	8.3645	0.41754	0.04757	0.001	422.152	0.160×10^{-6}	0.008	34.9
4026	35.790	427.305	8.3758	0.37311	0.04752	0.001	422.153	0.160×10^{-6}	0.008	35.0
4027	35.790	426.735	8.3864	0.33098	0.04749	0.001	422.110	0.160×10^{-6}	0.010	35.0
4028	35.790	426.179	8.3970	0.29138	0.04746	0.001	422.150	0.158×10^{-6}	0.012	35.3
4029	32.279	427.535	7.7127	0.37281	0.04613	0.001	422.149	0.171×10^{-6}	0.009	34.5
4030	32.278	426.922	7.7235	0.33073	0.04609	0.001	422.178	0.169×10^{-6}	0.010	34.7
4031	32.277	426.336	7.7338	0.29120	0.04609	0.001	422.166	0.166×10^{-6}	0.012	35.5
4032	32.277	425.813	7.7430	0.25423	0.04601	0.001	422.149	0.163×10^{-6}	0.013	36.0
4033	28.758	427.433	7.0199	0.37227	0.04475	0.001	421.825	0.182×10^{-6}	0.009	34.5
4034	28.758	426.811	7.0301	0.33032	0.04468	0.001	421.816	0.179×10^{-6}	0.010	34.9
4035	28.758	426.211	7.0400	0.29088	0.04468	0.001	421.790	0.179×10^{-6}	0.011	35.0
4036	28.758	425.639	7.0495	0.25398	0.04463	0.001	421.822	0.176×10^{-6}	0.013	35.6
4037	25.269	427.643	6.2935	0.37238	0.04336	0.001	421.789	0.200×10^{-6}	0.009	33.9
4038	25.269	426.975	6.3034	0.33045	0.04330	0.001	421.785	0.197×10^{-6}	0.010	34.3
4039	25.269	426.356	6.3127	0.29108	0.04329	0.001	421.807	0.196×10^{-6}	0.011	34.5
4040	25.268	425.804	6.3208	0.25421	0.04325	0.001	421.784	0.195×10^{-6}	0.013	34.6
4041	22.030	427.617	5.5903	0.35533	0.04207	0.001	421.789	0.222×10^{-6}	0.009	33.3
4042	22.029	427.074	5.5975	0.32252	0.04206	0.001	421.800	0.222×10^{-6}	0.010	33.3
4043	22.029	426.577	5.6040	0.29091	0.04201	0.001	421.786	0.221×10^{-6}	0.011	33.5
4044	22.028	426.084	5.6105	0.26121	0.04198	0.001	421.793	0.218×10^{-6}	0.013	33.8
4045	19.120	427.585	4.9318	0.33856	0.04095	0.001	421.825	0.246×10^{-6}	0.010	33.0
4046	19.120	427.026	4.9385	0.30638	0.04090	0.001	421.830	0.246×10^{-6}	0.011	33.0
4047	19.120	426.514	4.9446	0.27586	0.04086	0.001	421.829	0.244×10^{-6}	0.012	33.3
4048	19.120	426.029	4.9504	0.24689	0.04086	0.001	421.844	0.245×10^{-6}	0.013	33.3
4049	16.150	427.870	4.2299	0.33858	0.03985	0.001	421.893	0.281×10^{-6}	0.009	32.8
4050	16.150	427.308	4.2357	0.30642	0.03982	0.001	421.888	0.279×10^{-6}	0.011	33.1
4051	16.150	426.767	4.2413	0.27587	0.03982	0.001	421.892	0.280×10^{-6}	0.012	33.0
4052	16.150	426.256	4.2466	0.24689	0.03979	0.001	421.882	0.277×10^{-6}	0.013	33.3
4053	13.434	427.860	3.5681	0.32224	0.03885	0.001	421.902	0.332×10^{-6}	0.011	32.1
4054	13.433	427.263	3.5731	0.29090	0.03880	0.001	421.891	0.329×10^{-6}	0.012	32.4
4055	13.433	426.739	3.5777	0.26115	0.03879	0.001	421.927	0.328×10^{-6}	0.012	32.4
4056	13.433	426.225	3.5820	0.23294	0.03873	0.001	421.889	0.326×10^{-6}	0.014	32.7
4057	10.616	428.169	2.8565	0.32218	0.03788	0.001	421.961	0.404×10^{-6}	0.011	32.0
4058	10.615	427.574	2.8603	0.29100	0.03784	0.001	421.930	0.403×10^{-6}	0.012	32.1
4059	10.615	427.001	2.8643	0.26120	0.03782	0.001	421.955	0.405×10^{-6}	0.014	31.9
4060	10.614	426.476	2.8678	0.23298	0.03782	0.001	421.964	0.407×10^{-6}	0.015	31.8

Table 3. Thermal conductivity, thermal diffusivity, and heat capacity of nitrogen gas near 425 K—Continued

Run Pt.	Pressure MPa	Temperature K	Density mol/L	Power W/m	Thermal conductivity W/(m·K)	STAT	Cell temperature K	Thermal diffusivity m^2/s	DSTAT	Specific heat J/(mol·K)
4061	8.185	428.204	2.2268	0.30620	0.03705	0.001	421.986	0.511×10^{-6}	0.012	31.8
4062	8.184	427.588	2.2299	0.27571	0.03701	0.001	421.956	0.506×10^{-6}	0.012	32.1
4063	8.184	427.013	2.2329	0.24678	0.03700	0.001	421.992	0.509×10^{-6}	0.014	31.9
4064	8.184	426.459	2.2359	0.21939	0.03696	0.002	421.986	0.504×10^{-6}	0.016	32.2
4065	5.684	428.304	1.5628	0.29076	0.03627	0.001	422.071	0.732×10^{-6}	0.012	30.9
4066	5.684	427.662	1.5650	0.26101	0.03625	0.001	422.046	0.733×10^{-6}	0.014	30.9
4067	5.683	427.047	1.5671	0.23291	0.03624	0.001	421.984	0.714×10^{-6}	0.016	31.8
4068	5.682	426.467	1.5692	0.20636	0.03621	0.002	422.052	0.714×10^{-6}	0.017	31.7
4069	3.262	428.449	0.9050	0.27577	0.03563	0.001	422.058	0.122×10^{-5}	0.015	31.5
4070	3.260	427.778	0.9061	0.24681	0.03562	0.001	422.127	0.122×10^{-5}	0.016	31.5
4071	3.259	427.155	0.9071	0.21944	0.03561	0.002	422.067	0.121×10^{-5}	0.018	31.8
4072	3.258	426.571	0.9082	0.19367	0.03557	0.002	422.071	0.122×10^{-5}	0.020	31.6

obtain the isobaric heat capacity from the measured thermal diffusivity, we have calculated the density with the equation of state of Goodwin [14].

Figure 4 shows a typical deviation plot of the experimental temperature rises from the full heat transfer model for a liquid phase toluene point (number 1202) at a temperature of 324 K. The deviations from linearity are less than 0.04%. The deviations show that much of the noise is due to 60 Hz electromagnetic interference, but the noise is acceptably small. Table 1 shows two additional statistics which reflect nonlinearity of each data set relative to the ideal line source model, eq (1), after correcting according to eq (4). The first term is "STAT" which reflects the uncertainty in the slope of the regression line at a confidence level of 2 times the standard deviation (2σ). The term "DSTAT" reflects the uncertainty in the intercept of the regression line at a 2σ confidence level. For instance, a value of "STAT" or "DSTAT" of 0.001 indicates the 2σ uncertainty is 0.1%. As discussed earlier, we expect the thermal radiation correction to affect the measured thermal conductivity of toluene more and more as the temperature is increased above 370 K. The effect can be seen in the statistic "STAT" which is a numerical description of a deviation plot such as figure 4. Graphically, the deviation plots are no longer random but become systematically curved, as predicted by eq (11). Consequently, the thermal conductivities ob-

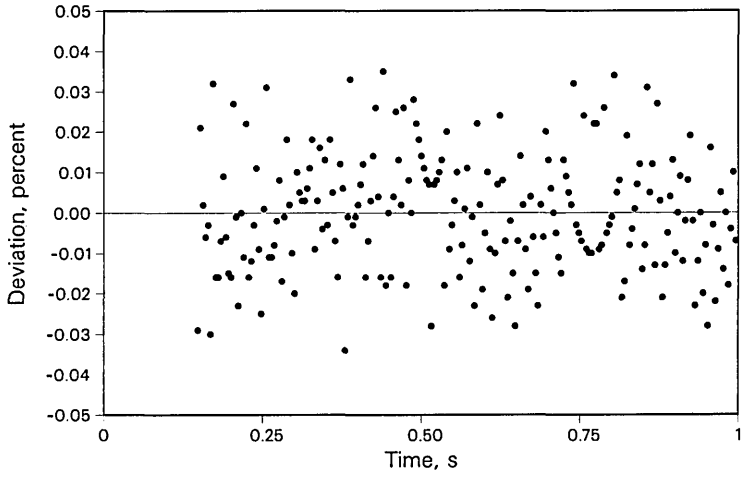

Figure 4. Typical deviations of experimental temperature rises from the calculated straight line versus the log of time for liquid toluene data point 1202 at $T_0 = 324.039$ K and $P = 0.088$ MPa.

tained from the usual linear fit are larger than they should be. To obtain correct results, we apply eq (12) to the experimentally measured temperature rises and evaluate B for every individual point. Next, the experimentally determined values for B are fit to a linear function in temperature. The resulting expression is

$$B = -0.0685 + 2.310 \times 10^{-4}\, T_0 \qquad (15)$$

where B is in s^{-1} and T_0 is in K. The values given by eq (15) are used to re-evaluate the radiation correction, δT_5, for each data point. The results corrected in this fashion are given in table 1.

Figure 5 shows the deviation plot for the temperature rises for a toluene data point (2105) at $T_0 = 548.140$ K and $P = 2.686$ MPa, before and after the radiation correction δT_5 has been applied. The deviation "STAT" has decreased from 0.002 to 0.001 and the curvature has been eliminated. These

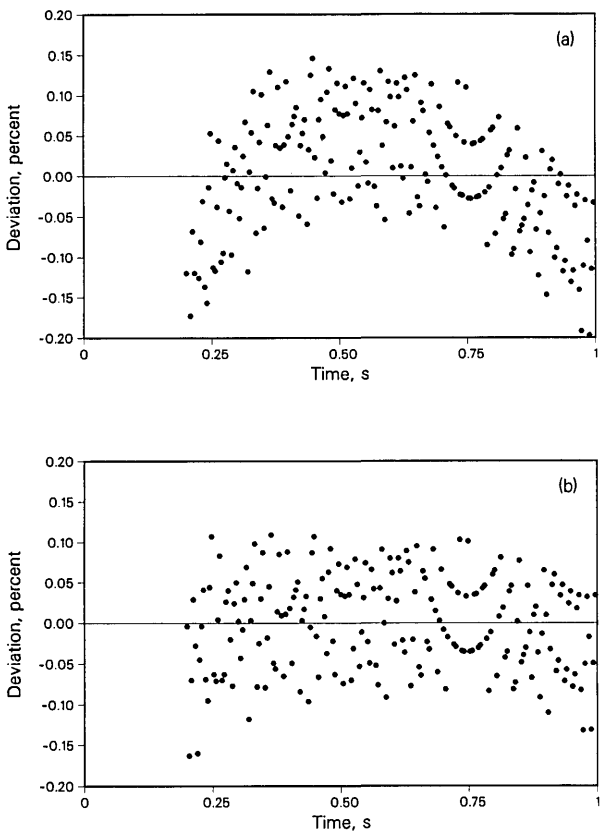

Figure 5. Liquid toluene data point 2105 at $T_0 = 548.140$ K and $P = 2.686$ MPa.

a) before application of the radiation correction, eq (9), "STAT" is 0.002.

b) after application of the radiation correction, eq (9), "STAT" is 0.001.

results support the model developed by Nieto de Castro et al. [7] to account for the effect of radiation in absorbing media, and suggest that the instrument with a revised δT_5 is operating in accordance with its mathematical model.

Figure 6 shows both the uncorrected and the radiation corrected thermal conductivity values of toluene near the saturation line as a function of temperature. The standard reference data correlation of Nieto de Castro et al. [13], which is valid to 360 K, is a line shown in figure 8. The measurements of Fischer and Obermeier [15] are also displayed. These were obtained with a rotating concentric-cylinder apparatus, operating in steady-state mode, for different gaps between the cylinders. We have included their extrapolation to zero gap, which is considered to be their radiation-corrected thermal conductivity. Figure 6 shows that our transient hot-wire instrument has a smaller radiation contribution than the steady-state measurements. However, the transient hot-wire radiation contribution becomes significant at elevated temperatures, 3.1% at 550 K. The larger radiation contribution in steady-state methods produces much larger uncertainty in the extrapolated radiation-corrected thermal conductivity data obtained with steady-state instruments. The temperature dependence along the saturation boundary, shown in figure 6, is similar to the trend reported in reference [13] with respect to the thermal conductivity data of Nieto de Castro et al. [7]. The data above 370 K show the presence of radiative effects. Also shown in figure 6, as an insert, are the compressed-liquid data at 550 K, which correspond to the shaded area of the diagram.

Deviations between the toluene thermal conductivity data and the correlation by Nieto de Castro et al. [13] are shown in figure 7 for temperatures up to 380 K. All of the data are within 1% of the correlation from 300 to 372 K; however, the deviations are systematic. We suggest that a higher-order temperature-dependent term might be added to the correlation in order to extend its temperature range.

Figure 8 displays the deviations between the heat capacity of toluene obtained from the measured thermal diffusivity and thermal conductivity using the density from the equation of state of Goodwin [14], versus the C_p value calculated by this equation of state. The data, uncorrected for radiation, show systematic departures from the equation-of-state prediction above 370 K, with deviations of 30% at 550 K. After the adjusted radiation correction δT_5 is applied, the deviations decrease to less than 10% at the highest temperature, falling in a band of ±5%

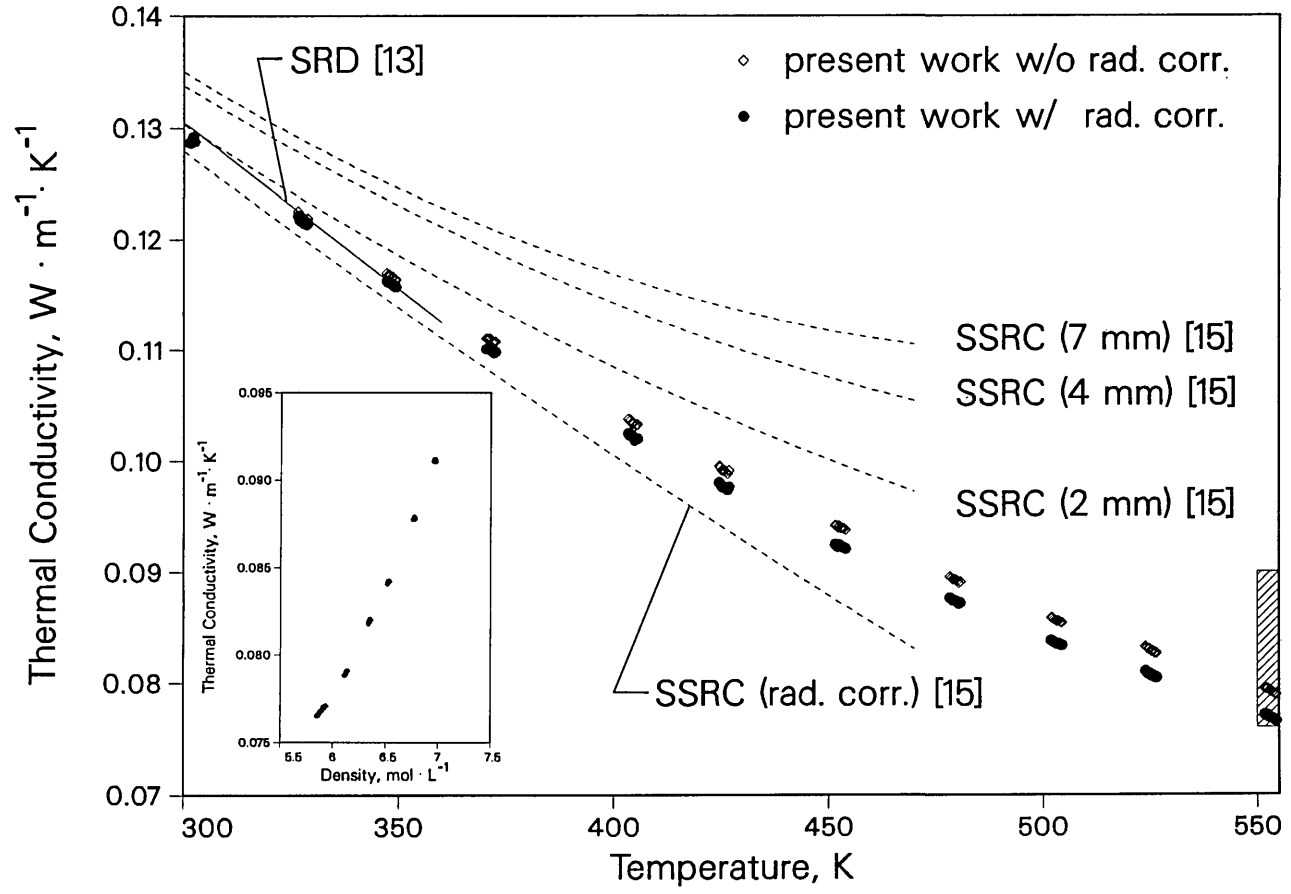

Figure 6. The thermal conductivity of liquid toluene near the saturation line. Dashed lines show steady-state rotating-cylinder (SSRC) results at various spacings along with the extrapolated radiation-corrected results. Solid line is the SRD correlation. The inset represents the thermal conductivity of toluene as a function of density near 550 K. The region of the inset is shaded on the main figure.

up to 500 K. The larger deviations above 500 K are still within the combined uncertainties of our diffusivity measurements and the equation of state of Goodwin [14].

Figures 6 and 8 demonstrate the performance of the instrument for the measurement of both thermal conductivity and thermal diffusivity at high temperatures in infrared absorbing fluids when the radiation correction, given by eqs (9) to (13) and (15), is applied.

5.2 Argon

We have previously reported two sets of transient hot-wire measurements of argon's thermal conductivity near 300 K [16,17]. Both of these data sets were made with the low temperature instrument described by Roder [8]. Thermal conductivity measurements on argon have also been reported by a number of other researchers [18–22]. Table 2 provides the results for the present measurements near 300 K. Younglove's equation of state [23] is used to obtain the densities reported in the table 2. The purity of the argon used in these measurements is better than 99.999%. Argon is transparent to thermal radiation, and the radiation correction at 300 K is negligible.

Deviations between the present thermal conductivity data and the new surface fit of Perkins et al. [24] as a function of density are shown in figure 9. The maximum deviation between our present measurements and the correlation is 1.2% at the highest densities. The present data were not, however, used in the development of the thermal conductivity surface [24]. The same trend of deviations relative to the correlation is exhibited by the other available data. Our thermal conductivity data agree with the results of the other data within ±1%. All of the other data were made with transient hot-wire instruments, with the exception of data from Michels et al. [19], which was obtained with a steady-state parallel-plate instrument.

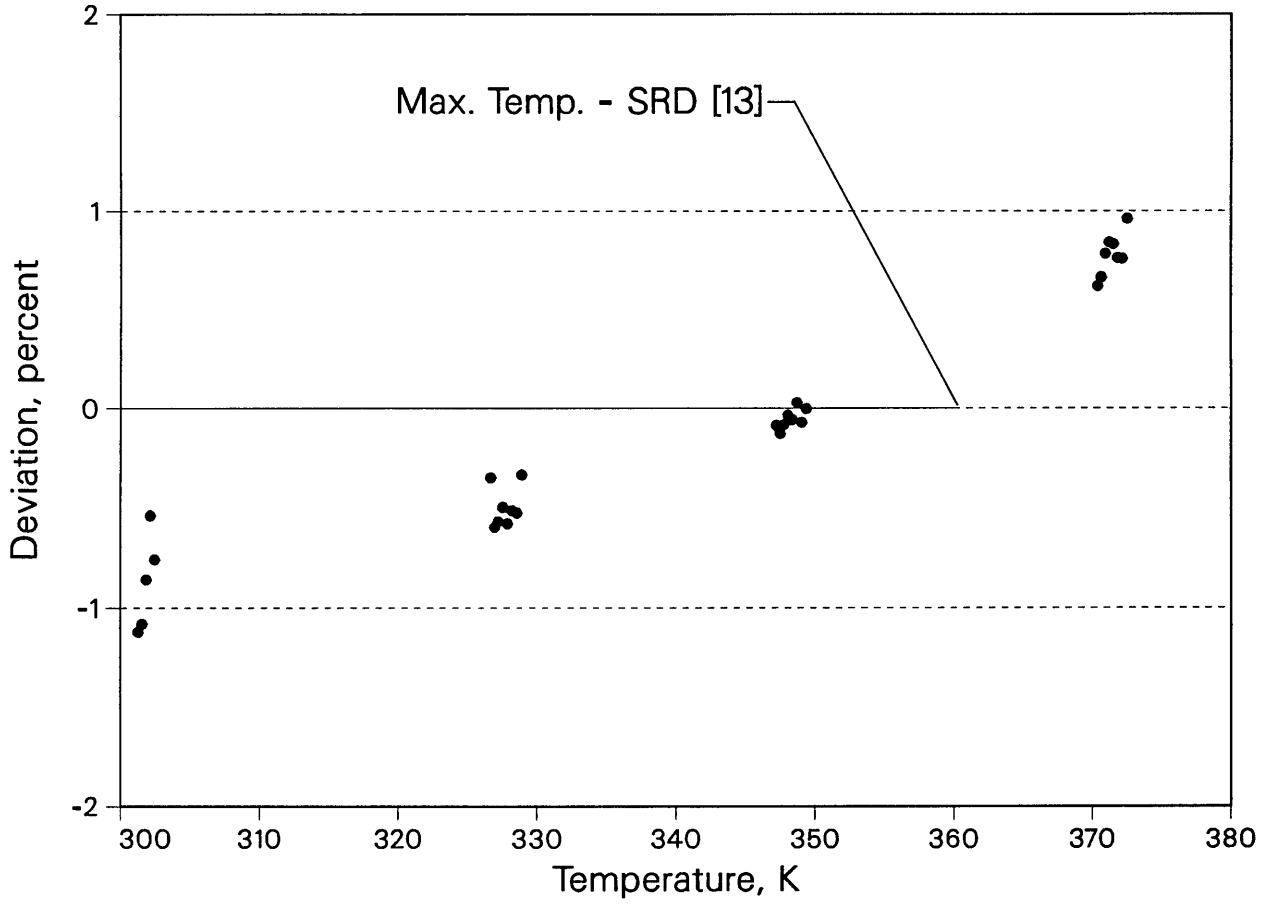

Figure 7. Deviation of liquid toluene thermal conductivity near saturation pressure relative to the correlation of Nieto de Castro et al. [13].

5.3 Nitrogen

For the present instrument, table 3 provides the results on nitrogen for temperatures near 425 K. Younglove's equation of state [23] is used to obtain the densities reported in the table 3. The purity of the nitrogen used in these measurements is better than 99.999%. Nitrogen is transparent to thermal radiation, and the radiation correction at 425 K is negligible.

Deviations between our thermal conductivity data and the correlation of Stephan et al. [25] as a function of density are shown in figure 10. The maximum deviation between our measurements and the correlation is 2%. Nitrogen thermal conductivity measurements have also been reported by several other researchers [21,22,26]. The same trend of deviations relative to the correlation is exhibited by the other available data. Our thermal conductivity data agree with those results to 1%, except for values from reference [22] for densities above 9 mol·L^{-1}. All of the other data were obtained with transient hot-wire instruments, with the exception of data from le Neindre [22], which were obtained with a steady-state concentric-cylinder instrument. The dilute gas value of Millat and Wakeham [27] is also plotted in this figure and agrees with the extrapolation of the present data within 0.5%. There is both theoretical [27] and experimental [28] evidence that the low density values of the Stephan et al. correlation [25] need to be revised. The correlation given by Younglove [23] has a completely different curvature as already shown in reference [28].

Figure 11 shows heat capacities of nitrogen given in table 3 for the isotherm at 425 K. The values are derived from the measured values of thermal conductivity and thermal diffusivity taking the densities from the equation of state [23]. They are compared to values calculated from the equation of state, and they are systematically higher than the equation-of-state predictions by about 5% except for the highest densities. We assign an estimated error of ±5% to our measured heat capacities; the

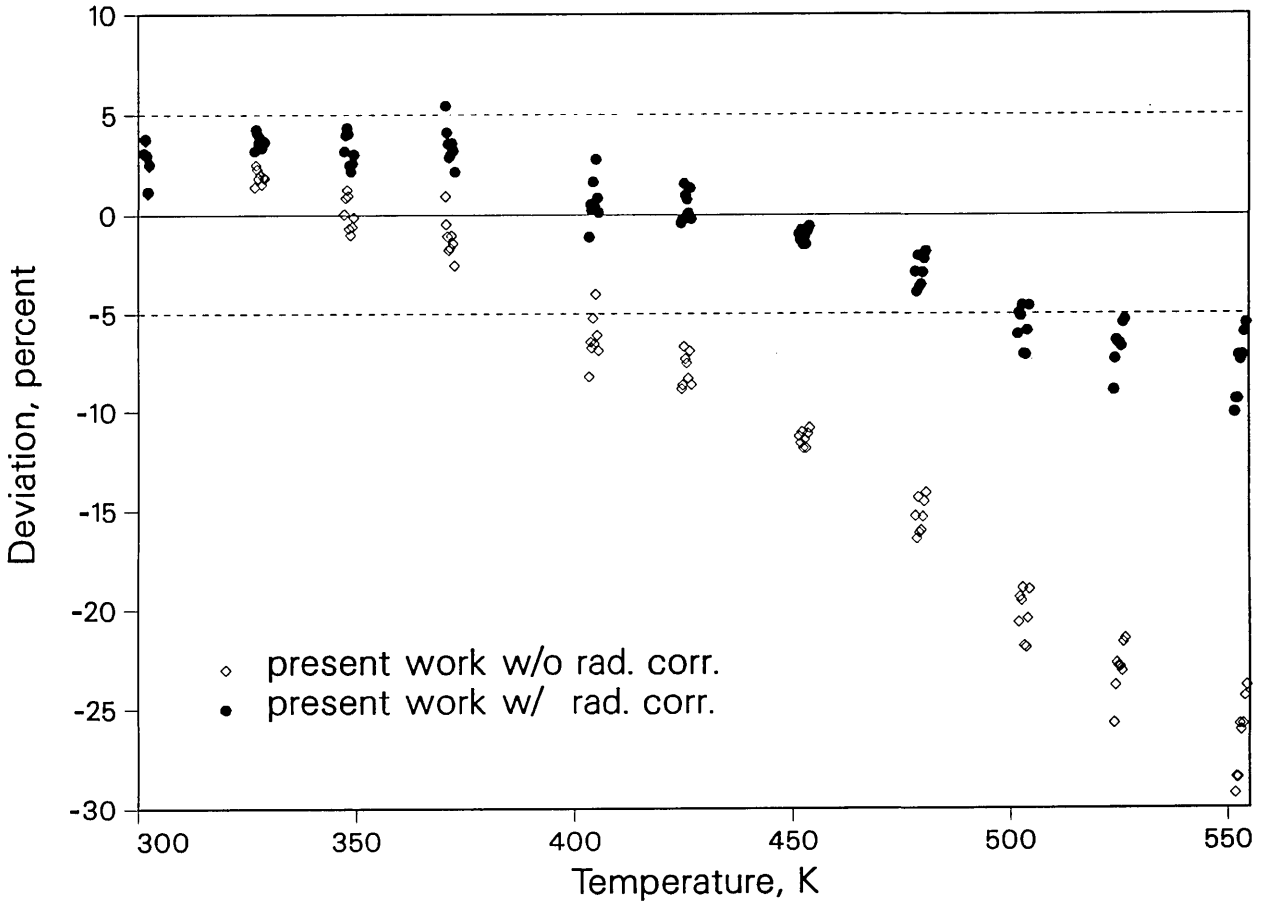

Figure 8. Deviation of liquid toluene isobaric heat capacity near saturation pressure relative to the C_p calculated using the equation of state of Goodwin [14].

error estimated for the specific heats from the equation of state is also 5%. Thus, the agreement between the two sources is within their mutual uncertainties even at the higher densities.

5.4 Repeatability Tests

In addition to comparisons of our thermal conductivity data with the data and correlations of other researchers, we have made many measurements to assess the repeatability of the instrument. The temperature assigned to a given thermal conductivity measurement is a function of the fluid temperature rise during the experiment. As a result, each power represents a different and independent temperature rise and experimental temperature. For a given cell temperature, we routinely make measurements at many powers not only to verify the instrument performance but also to check on the presence of convection. To check repeatability, results at different powers are compared in terms of deviations from a correlation of the thermal conductivity surface. Figure 12 shows deviations of the liquid toluene thermal conductivity data for four cell temperatures as a function of the applied power. There are from five to eight different powers for each cell temperature. The maximum difference between the deviations for each cell temperature is about 0.3%, which is equivalent to the experimental precision in λ. The deviations do not appear to have any power dependence.

The power dependence of the isobaric heat capacity of liquid toluene is shown in figure 13. The maximum difference between the deviations for each temperature is 2.6%. Again there is no discernable trend in the deviations of the heat capacity with respect to the applied power.

Figure 14 shows a deviation plot of 40 argon thermal conductivity data points relative to the correlation of Younglove et al. [29]. The applied power ranges from 0.11 to 0.42 W/m for a range of final temperature rises from 0.8 to 5 K. The data were obtained in four different sequences over 2 days. The four measurement sequences are shown

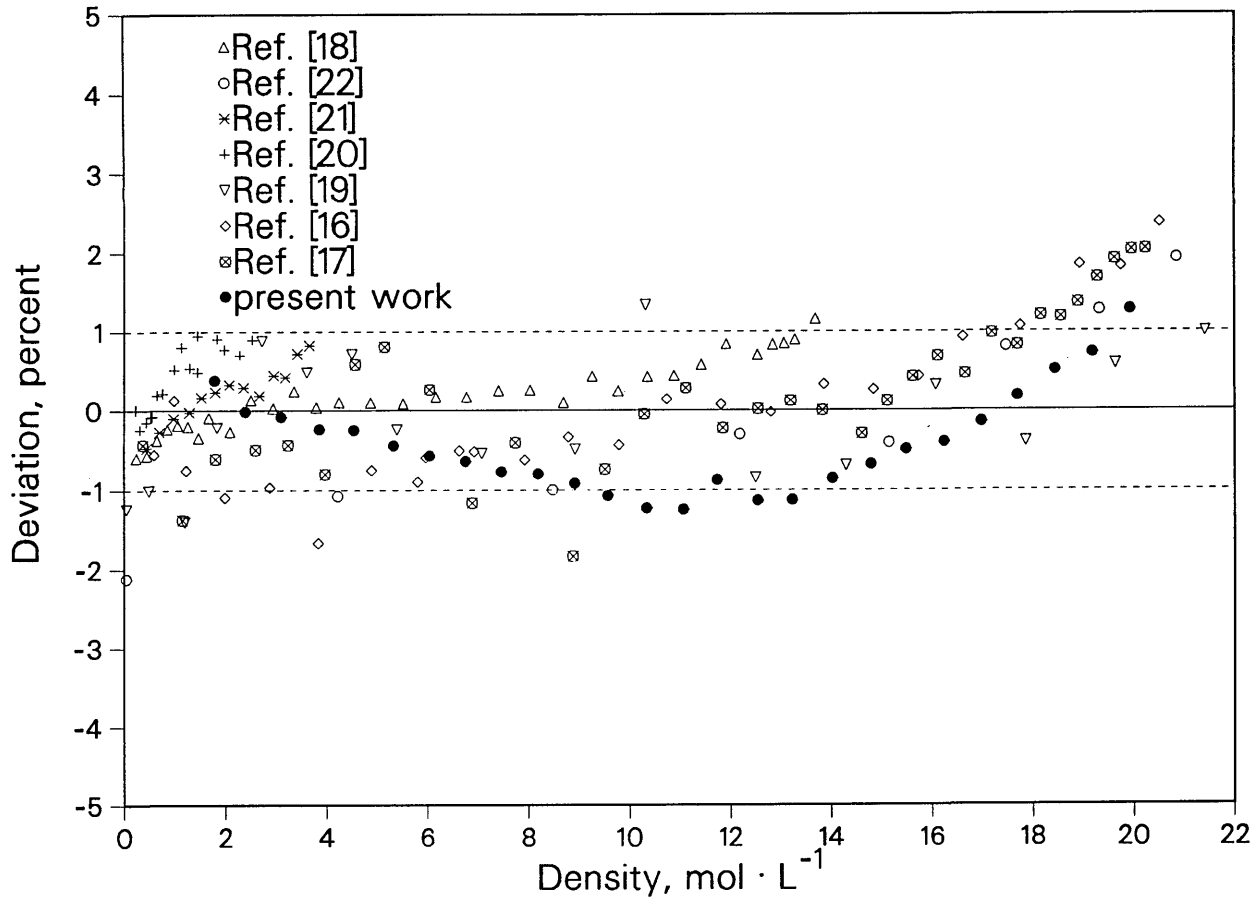

Figure 9. Deviations of argon thermal conductivity data near 300 K relative to the correlation of Perkins et al. [24].

with different plot symbols. The deviations from the correlation range from about −0.1% to −0.7%. Thus, the set of 40 measurements are consistent with each other and fall within a band of ±0.3%. The instrument's response is shown to be independent of applied power over a very wide range of temperature rises. The instrument's performance is also very repeatable over an extended period.

6. Summary

A new transient hot-wire thermal conductivity instrument for use at high temperatures is described. This instrument has an operating range from 220 to 750 K at pressures to 70 MPa. Thermal conductivity can be measured over a wide range of fluid density, from the dilute gas to the compressed liquid. The thermal conductivity data have a precision of ±0.3% and an accuracy of ±1%. The instrument is also capable of measuring the thermal diffusivity with a precision of ±3% and an accuracy of ±5%. Given accurate fluid densities, we can obtain isobaric heat capacities from the data. This instrument complements our low temperature instrument [8] which has a temperature range from 80 to 325 K at pressures to 70 MPa. A detailed analysis of the influence of radiative heat transfer in the transient hot-wire experiment has been performed, and radiation-corrected thermal conductivities are reported for liquid toluene near saturation at temperatures between 300 and 550 K. In addition, new measurements of the thermal conductivity and thermal diffusivity of argon and nitrogen verify the performance of the apparatus.

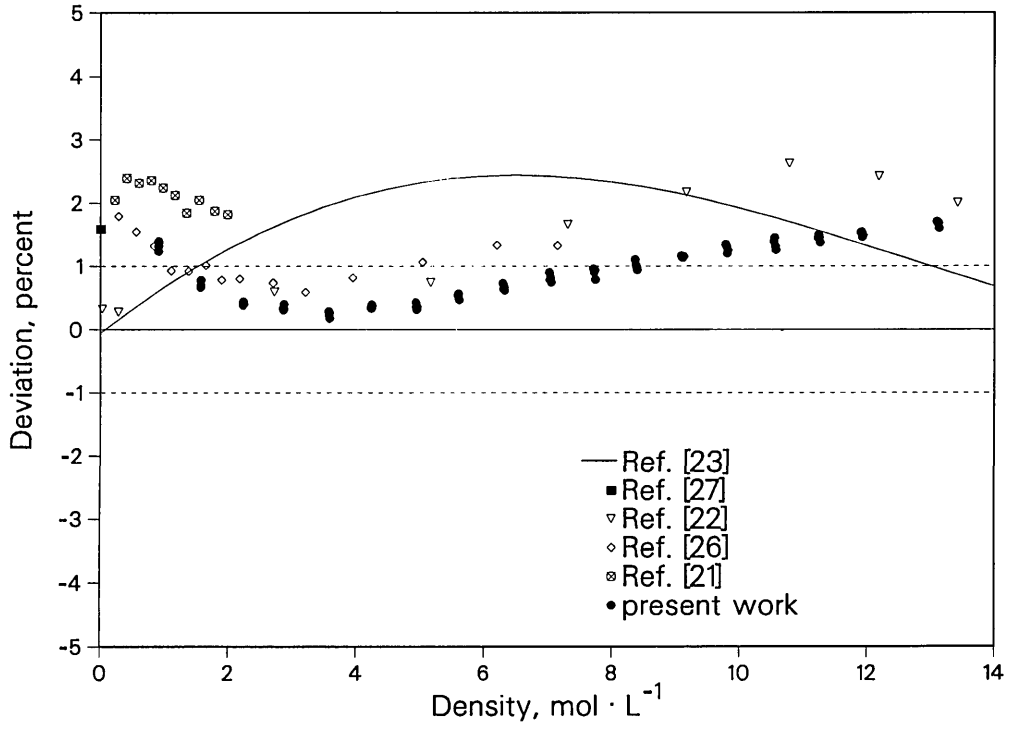

Figure 10. Deviations of nitrogen thermal conductivity near 428 K relative to the correlation of Stephan et al. [25].

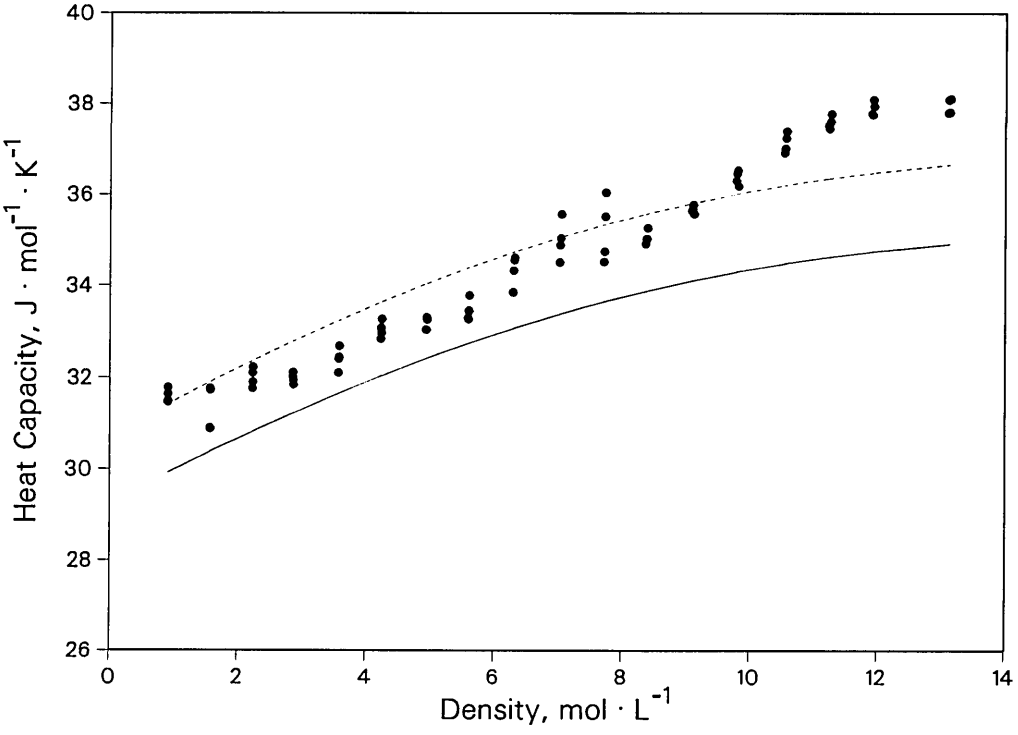

Figure 11. Nitrogen isobaric heat capacity relative to values calculated (solid line) from the equation of state of Younglove [23]. Dashed line is a 5% offset from [23].

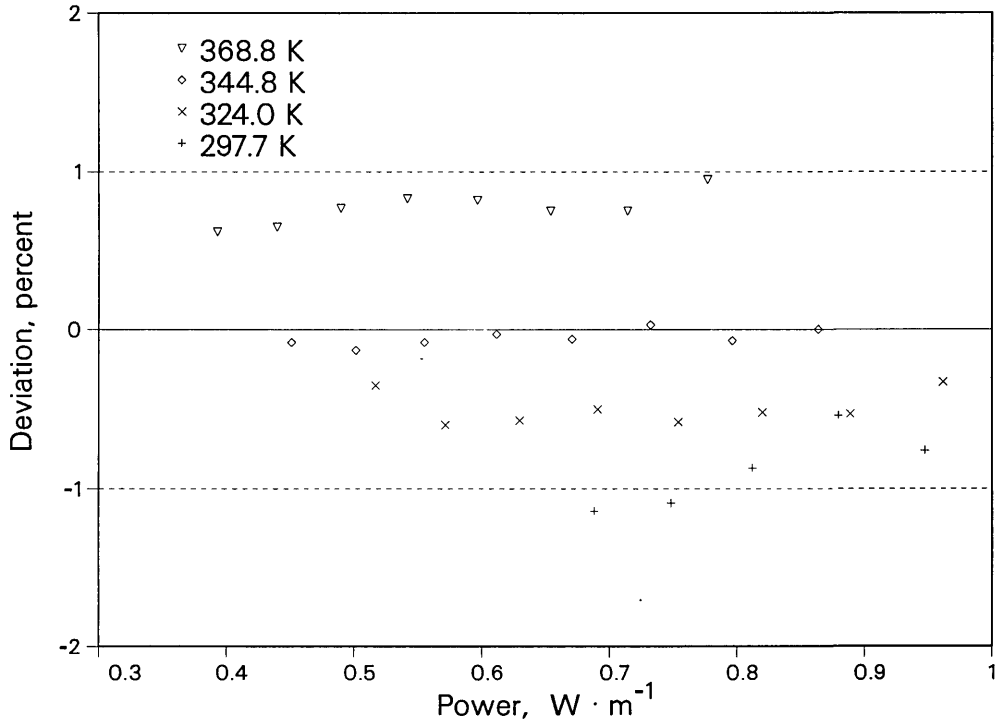

Figure 12. Deviations in the thermal conductivity of liquid toluene as a function of applied power. Baseline is the correlation of Nieto de Castro et al. [13]. Dashed lines show 95% uncertainty band.

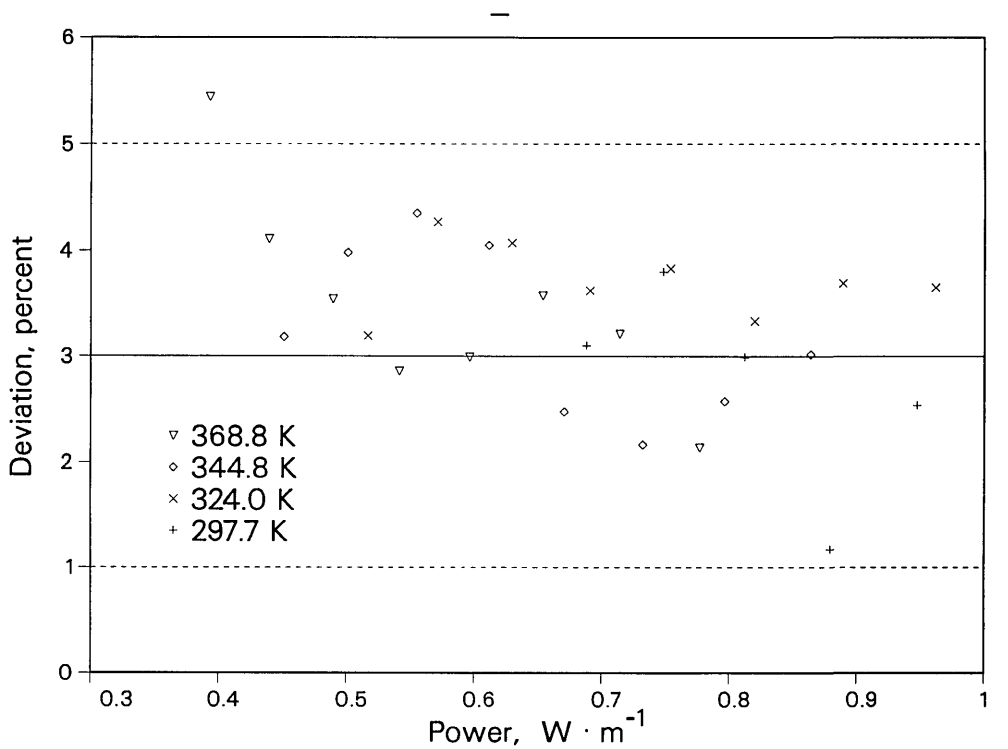

Figure 13. Deviations in the isobaric heat capacity of liquid toluene as a function of applied power. Baseline is the equation of state of Goodwin [14]. Dashed lines show 95% uncertainty band.

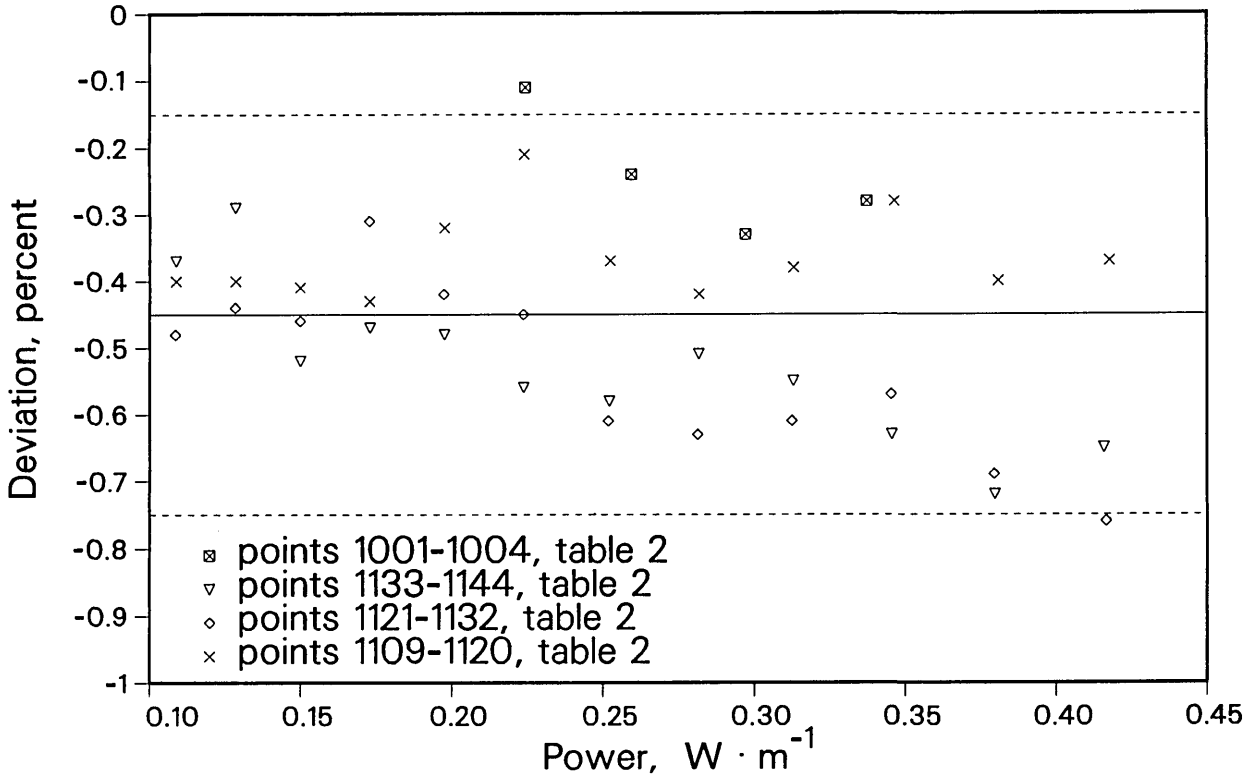

Figure 14. Deviations in the thermal conductivity of argon gas as a function of applied power. Baseline is the correlation of Younglove et al. [29]. Dashed lines show 95% uncertainty band.

Acknowledgments

We gratefully acknowledge the financial support of this work by the United States Department of Energy, Division of Chemical Sciences, Office of Basic Energy Sciences. One of the authors (C.A.N.C) thanks the Faculty of Sciences of the University of Lisbon for a leave of absence and the Thermophysics Division of the National Institute of Standards and Technology for a Guest Researchership.

7. References

[1] Nieto de Castro, C. A., JSME Inter. J. **31**, 387 (1988).
[2] Nieto de Castro, C. A., Taxis, B., Roder, H. M., and Wakeham, W. A., Int. J. Thermophys. **9**, 293 (1988).
[3] Carslaw, H. S., and Jaeger, J. C., Conduction of heat in solids, 2nd ed., Oxford University Press, London (1959).
[4] Healy, J. J., de Groot, J. J., and Kestin, J., Physica **82C**, 392 (1976).
[5] Siegel, R., and Howell, J. R., Thermal radiation heat transfer, 2nd ed., Hemisphere Publishing, New York (1981).
[6] Touloukian, Y. S., and DeWitt, D. P., Thermal radiative properties—metallic elements and alloys, Plenum Press, New York (1970).
[7] Nieto de Castro, C. A., Li, S. F., Maitland, G. C., and Wakeham, W. A., Int. J. Thermophys. **4**, 311 (1983).
[8] Roder, H. M., J. Res. Natl. Bur. Stand. (U.S.) **86**, 457 (1981).
[9] Roder, H. M., in Proc. 5th Symposium on Energy and Engineering Sciences, Dept. of Energy (U.S.) CONF-870687 (1987) p. 61.
[10] Roder, H. M., and Nieto de Castro, C. A., Cryogenics **27**, 312 (1987).
[11] Roder, H. M., and Nieto de Castro, C. A., in Thermal Conductivity 20, Hasselman, D. P. H., and Thomas, J.R., Jr., eds., Plenum Press, New York (1989) p. 173.
[12] Roder, H. M., and Perkins, R. A., J. Res. Natl. Inst. Stand. Technol. **94**, 1 (1989).
[13] Nieto de Castro, C. A., Li, S. F. Y., Nagashima, A., Trengove, R. D., and Wakeham, W. A., J. Phys. Chem. Ref. Data **15**, 1073 (1986).
[14] Goodwin, R. D., J. Phys. Chem. Ref. Data **18**, 1565 (1989).
[15] Fischer, S., and Obermeier, E., High Temp. High Press. **17**, 699 (1985).
[16] Nieto de Castro, C. A., and Roder, H. M., J. Res. Natl. Bur. Stand. (U.S.) **86**, 193 (1981).
[17] Roder, H. M., Perkins, R. A., and Nieto de Castro, C. A., Int. J. Thermophys. **10**, 1141 (1989).
[18] Kestin, J., Paul, R., Clifford, A. A., and Wakeham, W. A., Physica **100A**, 349 (1980).
[19] Michels, A., Sengers, J. V., and Van de Klundert, L. J. M., Physica **29**, 149 (1963).
[20] Millat, J., Ross, M. J., and Wakeham, W. A., Physica **159A**, 28 (1989).

[21] Haran, E. N., Maitland, G. C., Mustafa, M., and Wakeham, W. A., Ber. Bunsenges. Phys. Chem. **87**, 657 (1983).
[22] le Neindre, B., Int. J. Heat Mass Transfer **15**, 1 (1972).
[23] Younglove, B. A., J. Phys. Chem. Ref. Data **11**, Suppl. 1 (1982).
[24] Perkins, R. A., Roder, H. M., Friend, D. G., and Nieto de Castro, C. A., Int. J. Thermophys. (1991) in press.
[25] Stephan, K., Krauss, R., and Laesecke, A., J. Phys. Chem. Ref. Data **16**, 993 (1987).
[26] Johns, A. I., Rashid, S., Watson, J. T. R., and Clifford, A. A., J. Chem. Soc. Faraday Trans. I **82**, 2235 (1986).
[27] Millat, J., and Wakeham, W. A., J. Phys. Chem. Ref. Data **18**, 565 (1989).
[28] Perkins, R. A., Roder, H. M., Friend, D. G., and Nieto de Castro, C. A., Physica A, In Press (1991).
[29] Younglove, B. A., and Hanley, H. J. M., J. Phys. Chem. Ref. Data **15**, 1323 (1986).

About the authors: *Richard A. Perkins is a chemical engineer in the Thermophysics Division of the NIST Chemical Science and Technology Laboratory. Hans M. Roder is a retired physicist from the NIST Thermophysics Division. Carlos A. Nieto de Castro is the president of the chemistry department at University of Lisbon, Portugal, and was a Guest Researcher at NIST.*

Volume 96, Number 3, May-June 1991
Journal of Research of the National Institute of Standards and Technology

[J. Res. Natl. Inst. Stand. Technol. 96, 271 (1991)]

Standard Reference Specimens in Quality Control of Engineering Surfaces

Volume 96 | Number 3 | May-June 1991

J. F. Song[1] and
T. V. Vorburger

National Institute of Standards and Technology,
Gaithersburg, MD 20899

In the quality control of engineering surfaces, we aim to understand and maintain a good relationship between the manufacturing process and surface function. This is achieved by controlling the surface texture. The control process involves: 1) learning the functional parameters and their control values through controlled experiments or through a long history of production and use; 2) maintaining high accuracy and reproducibility with measurements not only of roughness calibration specimens but also of real engineering parts. In this paper, the characteristics, utilizations, and limitations of different classes of precision roughness calibration specimens are described. A measuring procedure of engineering surfaces, based on the calibration procedure of roughness specimens at NIST, is proposed. This procedure involves utilization of check specimens with waveform, wavelength, and other roughness parameters similar to functioning engineering surfaces. These check specimens would be certified under standardized reference measuring conditions, or by a reference instrument, and could be used for overall checking of the measuring procedure and for maintaining accuracy and agreement in engineering surface measurement. The concept of "surface texture design" is also suggested, which involves designing the engineering surface texture, the manufacturing process, and the quality control procedure to meet the optimal functional needs.

Key words: calibration; ISO; quality control; roughness; roughness average; specimen; SRM; standard; standard reference material; step height; stylus; surface.

Accepted: February 25, 1991

Contents

[1]Guest Researcher from the ChangCheng Institute of Metrology and Measurement, Beijing, China.

1. Introduction

1.1 Review of Quality Control for Engineering Surfaces

At the turn of this century, it became apparent that the surfaces of mechanical parts could be an important factor in determining how well the parts functioned. This seemed most true in cases where two components were in static or dynamic contact. Even earlier came the realization that no function of a part could be guaranteed unless the method of manufacture could be controlled. The control was usually achieved in those days by including many details of the manufacturing process, such as "rough machining," "medium machining," and "fine machining," or equivalent symbols on the engineering drawing. By carefully following the procedure laid down, some degree of control of the manufacture, and hence the function, could be achieved [1,2].

This early engineering surface quality control involved a control loop consisting of surface manufacture, surface texture, and surface function. The control process depended on the limitations of the machining process involved, as well as the arbitrary opinions of operator and inspector which, all too often, did not coincide. The resultant problems became increasingly acute as the demand for more comprehensive specifications increased to keep pace with technological development [2].

It was not until the early 1930s, however, that any serious attempt was made to quantify the magnitude and nature of the relationship between surface texture and function [1]. The stylus instrument, a technology now more than half a century old [3], laid the ground work in quantifying the surface texture, and made engineering surface quality control possible by enabling the measurement and control of the values of various surface parameters. Both the manufacturer and the consumer benefited from this control. One of the examples was from the automobile industry. Automobiles no longer require running-in, their engines and transmissions are reliable and long lived, and their fuel economy has been improved appreciably by reduced internal friction [4].

Digital methods and computer techniques have further revolutionized surface metrology because in principle, a digitizing surface instrument with a digital computer is infinitely flexible and can produce almost any analysis of the surface geometry that is required [5]. As a result, there has been a significant increase in standards, definitions, algorithms and parameters (dozens of them) which, on the one hand, describe the surface geometry from different views, but, on the other hand, force the engineer to cope with a variety of national and international standards whose interrelationships are obscure [6].

Meanwhile, other surface measurement techniques based on different principles, such as optical interference [7–8], optical scattering, [8–10], capacitance [11], and ultrasound [12] were developed to detect the surface texture. These techniques provided more means of surface measurements for industry, such as on-line, area-based scattering measurement [13] and non-contact, optical profiling and mapping [14–16]. Many users of these new instruments and techniques attempt traceability or comparison to the stylus instrument measurements [10–12].

The field of surface profiling has been expanded to the atomic scale in recent years by the development of the scanning tunneling microscope (STM) [17–21] and its offshoot, the atomic force microscope (AFM) [21–24]. Long trace STM has also been developed now for the smooth engineering surface measurement [25].

All of these instruments and techniques now reaching the market or still in the development stage are complicating the situation further. When these instruments, all of them scientifically based, well designed, and equipped with modern digital and computer techniques, get together to measure the same surface, a considerable disagreement, say 50% or more, [26] could be obtained in the measured parameters.

This divergence of methods should not be a surprise to the instrument manufacturer, since surface texture measurement results not only provide an assessment of the surface under examination, but also an assessment of the instrument used in that examination. In many cases, differences between measurement results from different instruments are not due to the fact that one surface measuring instrument is right and the other wrong. A difference is more likely due to the different measurement strategies and internal variables that have been adopted for each instrument. In the case of stylus instruments, a few of these considerations include stylus size, instrument bandwidth, computational algorithms [5], and reference datums.

Also, these differences should not be a surprise to a mechanical engineer, since the so called "same surface" is actually not the same everywhere. Even on the surface of calibration specimens, a few percent variation of their roughness average (R_a) [27] value, or even more than 15% on other parameters

has been observed [28,29]. On real engineering surfaces, 50% or more variations of their surface parameter values is commonplace [29]. And it is almost impossible to make an engineering surface with uniformity in surface texture less than 1% [30]. It has also been known to aviation engineers that, during the aircraft engine manufacturing process, the surface roughness of the parts in the continuous grinding process always changed more than ±50% in a working day, but that a variation of 30% in the surface texture usually had only a very small effect on the parts' function [30].

However, such differences are difficult for an inspector to accept. The inspector would base his judgments of "pass" or "reject" on the parameter values of his surface texture measurements. His consideration of the accuracy and agreement of his measurements could be much more rigorous than the requirements of surface function. And differences of opinion between manufacturers and their customers over the functional suitability of components have occurred when parameter values were out of their specified tolerances on the customer's instrument but not on the manufacturer's [5].

One of the considerations in quality control is to assign a "safety margin," like that used in size tolerance quality control. The measured roughness value would have to be smaller than the designed tolerance for maximum roughness because of the uncertainty of the measurement. However, this is not a good solution to our problem. By no means does the smoother surface necessarily produce the better function. Some surface functions, such as lubrication, are degraded when the surface is overfinished. Furthermore, the costs of finishing products increase rapidly as the R_a values of the surface finish decrease, especially on smooth engineering surfaces [2,27,31].

The fractional uncertainty of surface roughness measurement increases rapidly as its R_a value decreases. However, surface quality control is most important on the surfaces of critical parts having small size tolerance and fine surface structure to meet some particular functional [32] requirements. The peak-to-valley roughness (R_t or R_{max}) usually takes up a significant fraction of the size tolerance, for example [33], between 31% and 49% of tolerance values for different diameter ranges and tolerance grades. That fraction increases as the size or tolerance of the workpiece decreases. Due to the developments in the automobile, aviation, space, and microelectronics industries, many critical parts with small size tolerance and fine surface finish have appeared. These smooth engineering surfaces need a high degree of quality control in their manufacturing process for their proper function during use. At present, however, surface metrology cannot meet these needs very well. Therefore, more and more people are aware of the importance of accuracy and agreement in their engineering surface measurements, especially of smooth engineering surfaces. As a representative of one automobile manufacturer said, when he visited our lab about a year ago, "We do pay money for our (surface) measuring error."

1.2 Outlook for Quality Control of Engineering Surfaces

The solution to the problems of engineering surface quality control involves two aspects:

1. To recognize parameters that are functionally important for each application and determine their control values,
2. To maintain accuracy and agreement in engineering surface measurements among various instruments.

The first aspect mentioned above involves controlled experiments. Usually, the parameters and their controlled values, by which the balance between manufacture and function can be well maintained, come from the manufacturer's experience with the product's performance and the consumer's satisfaction. If the balance is not maintained very well, or if either the manufacturing process or the surface function is changed, a controlled experiment should be performed to determine the important surface parameters and their control values, or even the validity of the manufacturing process itself. Otherwise, serious failures could occur when the parts are used. One example comes from an automobile factory line for making camshafts [1]. The manufacturing process there was changed from grinding and lapping to turning and burnishing, but no change was made in the specification of surface texture. Because the texture produced by turning is much rougher than that produced by grinding, the burnishers had to produce an enormous amount of plastic deformation in order to get the cam roughness down to the specified value. This deformation ruined the surface's ability to withstand stress, which caused the cams to fail in large numbers of cars.

During the controlled experiments, previously specified parameters may be replaced by new parameters, which relate better to specific functional properties of the surface rather than merely

statistical parameters characterizing geometrical features. One example is the family of R_k parameters [34, 35]. The reduced peak height R_{pk} serves to describe profile characteristics relating to the running-in behavior of surfaces, the core roughness depth R_k relates to the long-term running behavior, and the reduced valley depth R_{vk} relates to lubricant behavior of the surfaces in static contact.

It is important to measure a large number of surfaces during the controlled experiments in order to ensure that the findings are significant. It is also important to maintain a high degree of repeatability and accuracy in these measurements, which could be very different from that of surface roughness specimen calibrations.

This brings us to the second aspect of engineering surface quality control, the accuracy of engineering surface measurements. Maintaining accuracy involves considerations for: 1) measuring instruments; 2) measured surfaces, including both calibration specimens and the real engineering surfaces; and 3) calibration and measurement procedures. We discuss each consideration in turn below.

1) Measuring Instruments

In a measuring instrument, information about the surface may pass through several stages. In stylus instruments, for example, as the information flows from the stylus tip to the indication on the recorder, the surface information is subject to distortion. Some of this distortion is intentional, such as that produced by profile filters. Other distortion is unavoidable due to practical considerations such as the finite size of the stylus [5]. Therefore, any measurement of a single surface could yield, to a greater or lesser extent, an arbitrary result if there are no limitations on the reference conditions of these measurements.

Different instruments may be manufactured according to different national and international standards whose interrelationships are obscure [6]. Furthermore, many aspects of digital instruments are not covered by national and international standards, and this has led to different instrument philosophies being implemented by the manufacturers. This is most evident when measuring standards or reference specimens. Unless a surface measuring instrument has the same specification as the instrument used during the original calibration and certification of a specimen, the measured parameter value may differ from the marked value. At present the specifications for some aspects of the calibrating instrument are arbitrary [5].

For surface texture results to have any "absolute" interpretation, the standard reference measuring conditions of the measuring instrument must be precisely defined. Thus, the concept of the "reference surface metrology instrument" (RSMI) was suggested [5]. Without a reference instrument conforming to standard reference measuring conditions, direct comparison of parameter values between different instruments is often not satisfactory. And if these disagreements happen for the surfaces of calibration specimens, they can also happen for real engineering surfaces where the measurement problems may be harder to diagnose.

Different calibration laboratories do use different surface measuring instruments. The specification and subsequent adoption of the standard reference measuring conditions by calibration laboratories would ensure compatibility of parameter values of reference specimens, as well as real engineering surfaces.

2) Measured Surfaces—Calibration Specimens and Engineering Surfaces

One of the most interesting parts of instrument development has been that of proving performance. For example, the measurement of small stylus displacements combined with low operating forces and high magnifications required the development of more sensitive proving techniques than have been previously available. Generally the instruments could be adapted to prove themselves [3]. That is, a certified standard could be measured by the instrument to be proved, and the difference between the measured result and the certified value of the standard would show the performance of the instrument. Therefore, the calibration of the existing range of stylus instruments with their wide range of performance calls for different types of calibration specimens. Each type of calibration specimen, with its specially designed surface profile, high uniformity over its surface, and certified parameter values, should be useful for calibrating certain aspects of instrument performance and capable of avoiding a bad calibration due to effects caused by other performance characteristics of the instrument. For example, by using a sinusoidal specimen [36] with small slope and with wavelength long as compared with the stylus size but short when compared with the instrument cut-off length [37], we can calibrate both the vertical and horizontal magnification, or check the R_a readings of stylus instruments. These calibrations would be insensitive to the effects of stylus tip size or filter cut-off of the calibrated instrument.

Since real engineering surfaces are different from those of most existing calibration specimens, a check specimen with engineering surface features is suggested for use in engineering surface measurement. Its measuring area should also have high uniformity, and this is aided by having a unidirectional profile. The surface waveform, wavelength and roughness parameter values should be comparable to those of the measured engineering surface. The check specimen could be measured under standardized reference measuring conditions, and receive its certified parameter value, or even a set of certified values corresponding to various measurement conditions, such as stylus tip size and filter cutoff length. By comparing the measured result with the certified value, the check specimen thus provides an overall check of proper instrument usage in engineering surface measurement.

3) Calibration and Measurement Procedures

The calibration and measurement procedure involves the use of various calibration and check standards ("specimens"). Each calibrated specimen may have a limited range of application according to its own characteristics and those of the instrument to be calibrated. The validity of the calibration of an instrument would depend on the correct association of these individual characteristics [37].

It should be noted that certain calibration or measurement procedures correspond to the measurement of certain kinds of measured surfaces. At national laboratories, such as ours, the routine surface measurements are made mainly on the calibration specimens. Sometimes, when we measure real engineering surfaces, mostly for research purposes, a larger uncertainty is obtained in our measurements because the positional fluctuation of the engineering surface is appreciably greater than the instrumental uncertainty of an individual measurement [3]. This is an important factor inhibiting researchers from investigating their measuring errors in real engineering surface measurements. However, we believe that the use of a greater number of measurements (larger statistical samples) on these types of surfaces together with improved calibration procedures would significantly enhance the accuracy and usefulness of the surface measurements.

At industrial metrology laboratories, routine surface measurements are performed mostly on the real engineering surfaces. Some of these labs maintain their traceability of surface measurement to NIST by sending their calibration specimens to our lab for periodic calibration and by using the same procedure to measure their engineering surfaces as we did on the specimens. The validity of this traceability depends both on the difference between the calibrated specimens and the measured engineering surfaces, and on the difference between their instrument and ours.

All of these aspects mentioned above affect each other. For example, when a controlled experiment shows that new surface parameters, such as the R_k family, should be selected for some kind of functional surface quality control, instruments should be enhanced for the measurement of R_k parameters. In addition, new calibrated specimens, having highly uniform R_{pk}, R_k, and R_{vk} values close to those of the controlled engineering surfaces, should also be developed as intermediate check specimens relating the reference instrument and the instrument used for engineering surface quality control.

At the NIST surface calibration laboratory, we calibrate scores of reference specimens each year. We also issue the sinusoidal calibration specimens: SRMs (Standard Reference Materials) 2071–2075. In addition, specimens developed elsewhere are often sent here for testing their properties. In this paper, we describe the characteristics, utilizations, and limitations of different classes of precision roughness calibration specimens. We also discuss the effects of various reference conditions on specimen calibration and propose a measurement procedure, which involves the use of a check specimen, for engineering surface quality control. For one class of specimens the surface texture design is intended to meet the functional requirements in engineering surface quality control.

2. Calibration and Check Specimens

The reference conditions which should be defined, calibrated, or checked in a stylus instrument measurement are:

a) magnification, both in the vertical and horizontal direction;
b) the stylus tip;
c) the stylus loading;
d) the type of skid or reference datum;
e) the type of filter, reference line, and cut-off length;
f) profile digitization;
g) the algorithms for calculating parameters;
h) the number and distribution of profiles on the surface.

There are four types of calibration specimens according to the ISO 5436 standard [37], each of them are specially designed for the purpose of calibrating *certain* of the above characteristics of the stylus instrument. It is important to prevent these calibrations from errors caused by the *other* characteristics of the calibrated instrument or by the reference conditions during these calibrations. The four types of calibration specimens are discussed below.

2.1 Specimens for the Calibration of Vertical Magnification

At the NIST surface calibration laboratory, a set of step height standards with calibrated step heights ranging from 0.0291 to 22.90 μm is used for the calibration of our stylus instruments. The calibration profiles are taken under unfiltered conditions, and the width of the step or groove on the specimen is wide enough so that the step height measurements are not sensitive to the stylus size. Therefore, our consideration of the reference conditions is focussed on the datum of the step and the algorithm of the step height.

Most stylus instruments have mechanical reference datums provided by their slides. The straightness of the traverse mechanism along the datum is an important factor. It could be more than 0.1 μm for a 2 mm traversing length. At high vertical magnifications, say 50,000× or more, such a straightness error can deform the measured step profile seriously (see fig. 1a). The use of the skid can reduce this error to a large extent (see fig. 1b). However, the skid can only be used when the surface has a smooth area which can serve as a datum for the skid to move on. This approach works well as long as the skid is offset sufficiently from the stylus along the direction of travel so that the skid never crosses the step itself but rides only on the outer flat section of the surface. Otherwise, the step could be damaged seriously when the skid touches it.

Flexure pivots are used in some high resolution stylus instruments instead of slides. The resulting path of the stylus is slightly curved in the horizontal plane. If the measured surface is not leveled perpendicular to the general direction of travel, the measured profile of the step could be deformed to be either convex or concave (see fig. 2b, c).

We recently tested the calibration of step height for errors due to curvature in the measured profile. There were two algorithms in the step height calculation. The first was adopted by ISO 5436 [37].

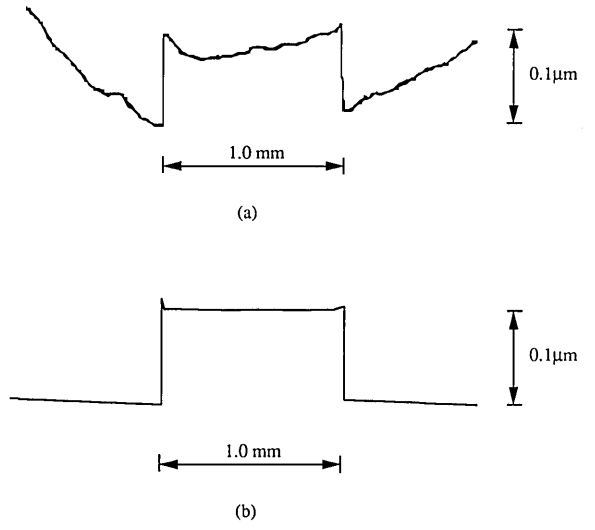

Figure 1. The profile of a step height specimen distorted by the straightness of the traverse mechanism of the stylus instrument.
a) skidless;
b) with skid.

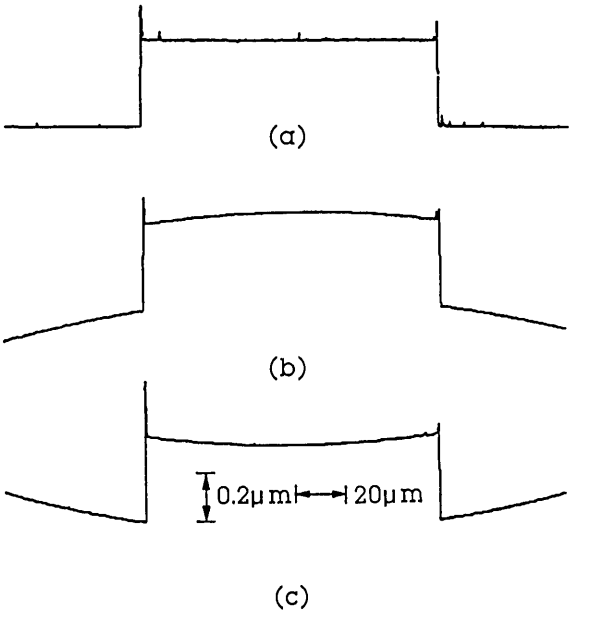

Figure 2. The profile of the step height specimen distorted by the unleveled surface for a stylus instrument with a flexure pivot stage.
a) leveled surface;
b) and c) unleveled surface.

It calculates the step height as the distance between two least square lines as shown in the Appendix. The second one is used at NIST. It calculates the step height as the average of the left and right step heights. Both of these are calculated by one-side step-height algorithms [38]. We now consider the effects on the calculated step heights when the step profile is distorted during the calibration of a stylus instrument or during a step height measurement.

We measured a 0.303 μm two-sided step in leveled and extremely unleveled positions with our high resolution stylus instrument (see fig. 2). In spite of the distortion, we got very repeatable results of 0.3024, 0.3029 and 0.3032 μm corresponding to figure 2a, b, and c, when the step height was determined by the average of the left and right. However, a tremendous error could be obtained corresponding to figure 2b and c if the step height was determined by the two least square lines. This error Δ could be calculated analytically using considerations discussed in the Appendix.

Such extreme profile curvature is unrealistic for a step height measurement. However, a difference of a few percent between results of the two algorithms can arise both in instrument calibration and step height measurement. The step profile can be deformed by the straightness error of the traverse mechanism of the instrument, by an unleveled step surface (for flexure pivot profilers) or by curvature in the specimen itself. These curvature problems become more significant at high magnification. The variations could be remarkably decreased if least squares circular arcs were used instead of least squares lines to determine the step height. That way the algorithm automatically anticipates the surface curvature.

The straightness of the traverse mechanism could be checked with the stylus traversing on an optical flat glass surface. For one of our stylus instruments, such a profile segment is shown in figure 3a. The use of a skid could decrease this error to a large extent (see fig. 3b).

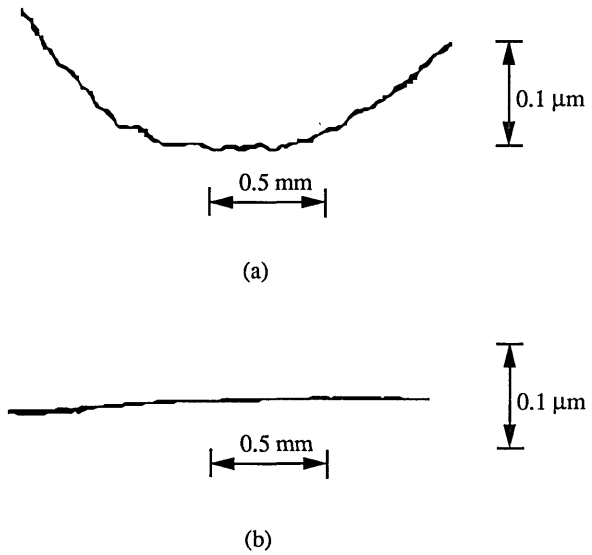

Figure 3. The straightness of the traverse mechanism checked by traversing on an optical flat (lithium niobate) surface.
a) skidless;
b) with skid.

2.2 Specimens for Checking Stylus Tip Condition

The stylus radius and apex angle have been adopted by most national and international standards [27,39]. However, we have gradually found that the stylus radius is a difficult quantity to define, especially for some chisel-shaped styli [40, 41]. We now prefer stylus width w, which is defined to be the distance between the two points of contact when the stylus profile is inscribed in a 150° angle. If the stylus tip has a perfectly round shape, then $w = 0.52r$ [40].

Although several methods, such as SEM, optical microscopy and the razor blade trace, have been developed to measure the stylus size under laboratory conditions [41], the most difficult part of the roughness instrument to check in the workshop is still the stylus [3].

One method involves traversing a calibrated roughness specimen having a series of fine V-shaped grooves, the profiling of which is sensitive to the stylus size. By comparing the roughness value obtained with the true roughness value of the specimen, the stylus size may be determined. These specimens have been widely used for stylus tips over 10 μm. However, finer grooves, useable for checking 2 μm tips, are still difficult to make [3,37].

One set of specimens is fabricated by oxidation and etching of silicon wafers [42]. Each specimen is composed of square-profile arrays of four distinct pitches, with wavelength 6, 20, 60, and 200 μm. The pitches with 6 and 20 μm wavelength are nearly equal to certain standard stylus sizes and could also be used for testing stylus tips. One of the experiments on such a wafer specimen with nominal R_a value of 0.0425 μm is shown in figure 4. Roughness average (R_a) values were measured on the four pitches with styli of different nominal radius. The measured R_a value, which decreases as the stylus

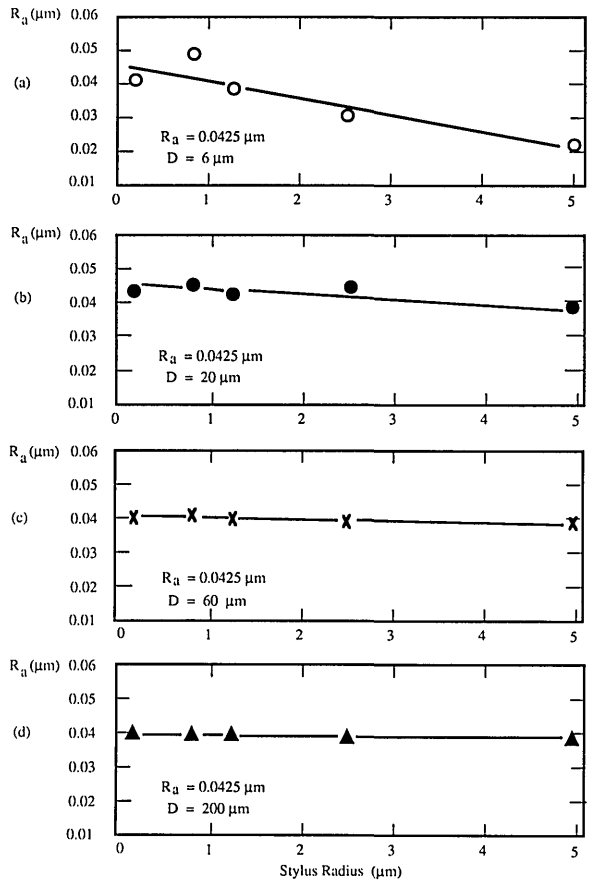

Figure 4. The measured roughness values on a rectangular profile specimen with nominal $R_a = 0.0425$ μm and four different pitches [42].
a) $D = 6$ μm;
b) $D = 20$ μm;
c) $D = 60$ μm;
d) $D = 200$ μm.

size increases, should be less than the nominal value, with the difference depending on the stylus size as well as profile wavelength. Sometimes, when the finer stylus was tested, the measured R_a value could be larger than the nominal value, because of "stylus flight" (see fig. 5) [43]. The stylus tip was losing mechanical contact with the surface under conditions of very low stylus force. Stylus flight could even happen at a slow traversing speed of 0.122 mm/s on our high resolution stylus instrument.

2.3 Periodic Profile Specimens

Various periodic specimens, with rectangular, triangular, arcuate and sinusoidal profiles, are used for the calibration of R_a readings of stylus instruments. They may also be used for the calibration of horizontal magnification, and the sinusoidal specimens may be used for checking the signal filter characteristics of stylus instruments.

The wavelength of these periodic specimens should be long enough, say 30 μm or more, that the measured roughness is insensitive to the stylus size; and short enough, say less than 1/10 of the cut-off length, [37] that the measured roughness is insensitive to the filter of the instrument. Figure 6 shows the effect of the stylus size on R_a results for different periodic specimens. Results for the sinusoidal specimens with 100 μm wavelength are insensitive (fig. 6a), while those for the triangular specimens with 15 μm wavelength are very sensitive to the stylus size (fig. 6b). As an exception, the R_a value increases with the stylus size on the specimen with cusped peaks (see fig. 6c). That is because the

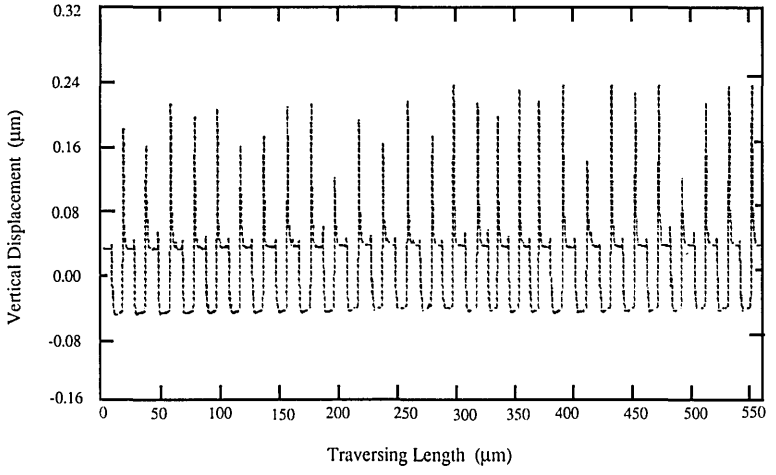

Figure 5. Stylus flight on the surface of a rectangular profile specimen.

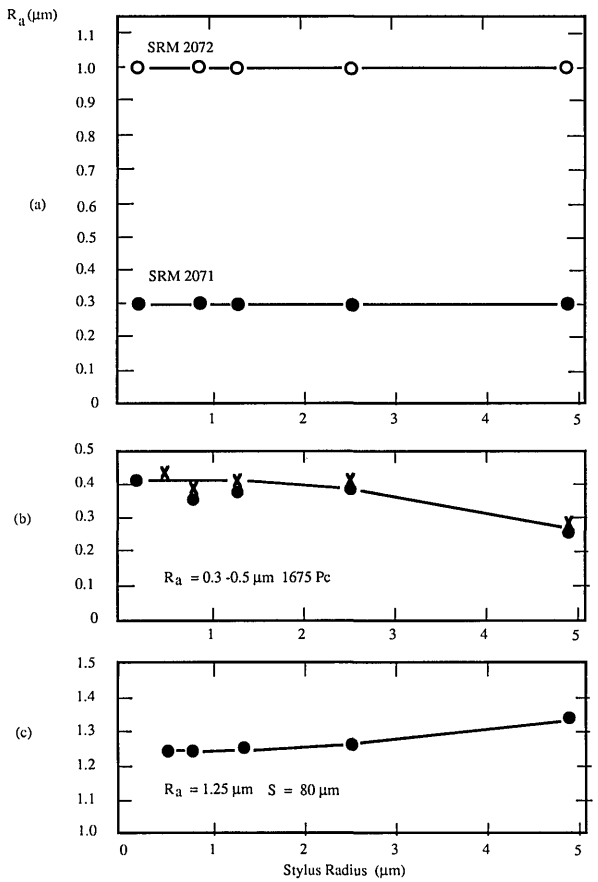

Figure 6. The effects of stylus size on roughness measurements of different periodic profile specimens. The curves are eyeballed fits.
a) sinusoidal profile specimens;
b) specimens for checking stylus tips; × $-M_v$=5,000×; • $-M_v$=20,000×
c) profile specimen with cusped peaks.

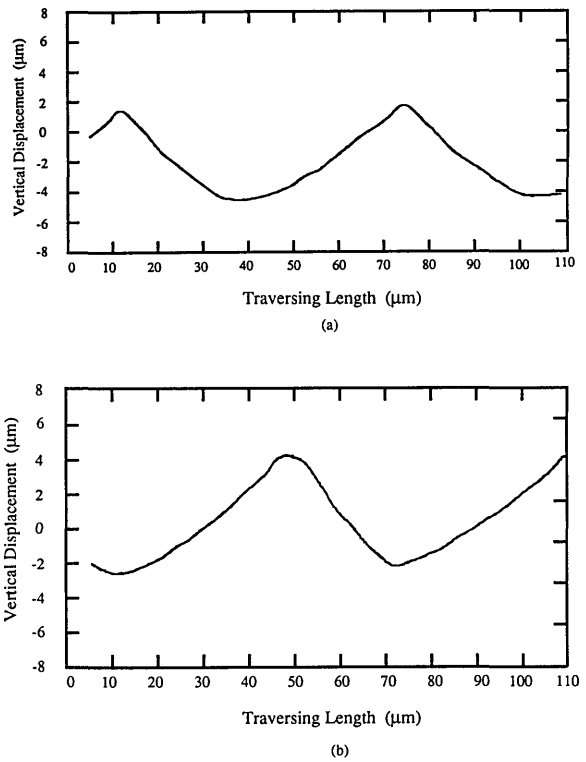

Figure 7. The peaks of the arcuate profile specimen are enlarged by increasing the stylus size.
a) stylus width w = 0.4 μm, R_a = 1.26 μm;
b) stylus width w = 5 μm, R_a = 1.33 μm.

widths of the peaks were enlarged by the bigger stylus size. Figure 7a shows a profile of the latter specimen measured with a stylus of width 0.4 μm, yielding an accurate profile of the cusped peaks. Figure 7b, on the other hand, was measured with a 5 μm width stylus and therefore shows broadening of the peaks, leading to the increased R_a value. This effect could even happen on real engineering surfaces with cusped profiles formed by the cutting tool radius during turning, planing, and side milling processes [44].

The NIST sinusoidal specimens, SRM 2071–2075, with R_a values ranging from 0.3 to 3 μm and wavelengths ranging from 40 to 800 μm, were manufactured by the numerical controlled diamond turning process, and calibrated at NIST's computerized surface calibration system consisting of a stylus instrument integrated with a laser interferometer. The R_a value was measured by a stylus instrument calibrated by one of our step height calibration specimens. Meanwhile, the surface wavelength was calibrated by using a laser interferometer to measure the lateral displacement of the moving stylus. These sinusoidal specimens can yield constant R_a values despite different stylus sizes (see fig. 6a). The surface profiles of some of these were machined with a fine-scale harmonic on the sinusoidal profile as shown in figure 8. The spacing of the fine-scale structure was approximately 4.2 μm. It can be useful for monitoring the quality of the fine stylus tips. The upper profile in figure 8 was obtained with a stylus of width 4 μm, while the lower profile was obtained with a stylus of width 0.5 μm. The fine-scale structure can be also useful for checking different zero-crossing and peak-counting algorithms in the surface profile analysis.

A set of sinusoidal specimens with various amplitudes and wavelengths (λ), λ being equal to 1/3, 1, and 3× cut-off length, were also suggested in ISO 5436 for the calibration of filter characteristics of

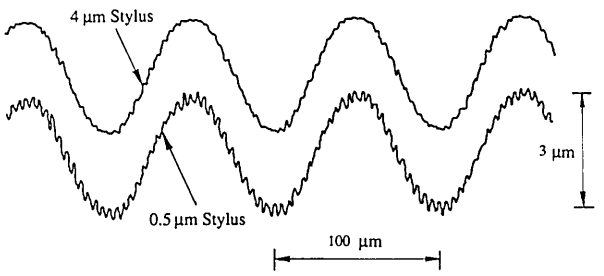

Figure 8. The fine-scale harmonic on the profile of sinusoidal specimen SRM 2072.
a) stylus width $w = 4$ μm;
b) stylus width $w = 0.5$ μm.

stylus instruments [37]. Since our currently existing sinusoidal specimens SRM 2071–2075 cannot cover as wide a range in amplitude and wavelength as discussed in ISO 5436, we can test the filtering characteristics of the instruments at only a few frequencies. We measured the R_a values of the sinusoidal specimen for various settings of the filter cutoff length, and compared these values with theoretical tolerance values calculated from the B46.1-1985 standard [27] or ISO 3274 [39] corresponding to different cut-off lengths. By this method, we calibrated the filter characteristic of our stylus-computerized surface measuring system as shown in figure 9. All of these measured R_a values on SRMs 2072 and 2075 for various cutoff lengths of the system are within the tolerance range.

2.4 Specimens with Random Profile

Many engineering surfaces have random profiles, such as those manufactured by grinding, lapping, polishing, or honing. Others have profiles with random content superimposed on a periodic profile. These include surfaces manufactured by turning, planing and milling processes. Random profiles have wide amplitude distributions (usually Gaussian) and wide spatial frequency distributions. Measurements of these surfaces with different instruments could produce varying results if the spectral responses of the instruments differ significantly. In 1965, Häsing at PTB [45] developed random profile roughness specimens to represent these types of surfaces. These specimens are made with a measuring area having a unidirectional random profile manufactured by the grinding process. The random profile is repeated every 4 mm, a period exactly equal to the evaluation length of the measurement. Therefore, the measured random profile is always essentially constant within

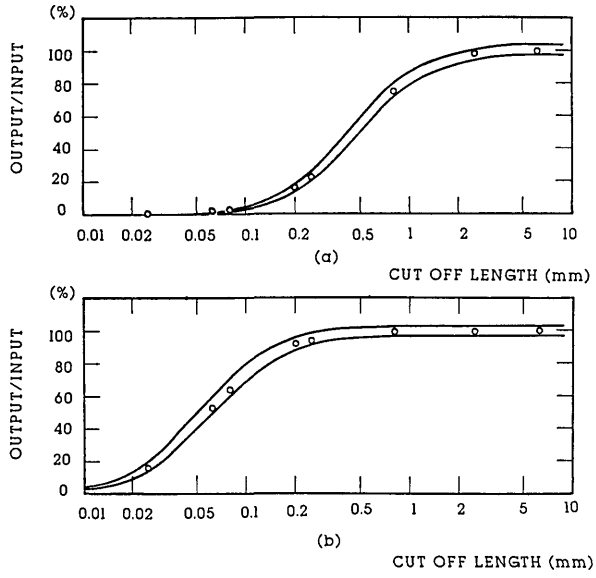

Figure 9. The filter characteristics of the NIST stylus-computerized surface calibration system calibrated by a single sinusoidal specimen.
a) by SRM 2075, $R_a = 1$ μm, $D = 800$ μm;
b) by SRM 2072, $R_a = 1$ μm, $D = 100$ μm.

the measuring area, irrespective of the measuring positions on the specimen. These specimens consist of a set of three pieces, with R_a values of 1.5, 0.5, and 0.15 (or 0.2) μm. With random profile specimens, stylus instruments can be subjected to a summary test covering all stages in the instrument from the stylus tip to the indication of the measured value. These roughness specimens play an important role in getting agreement for engineering surface measurement among various instruments. Hillmann [46] recently reported that a couple of years ago, differences in measured surface parameters of 40% and more were ascertained in national and international comparison measurements. Now, the differences between measurements carried out within the framework of the German Calibration Service amount to only a few percent; in an audit of the European Communities less than 15% was attained [26,46]. Hillmann's recommendations included the use of both well defined measurement conditions and well characterized roughness specimens.

The PTB roughness standards are limited to R_a values ≥ 0.15 μm. However, smooth engineering surfaces with $R_a \leq 0.1$ μm play an increasingly important role both in industry and research. Most smooth engineering surfaces, such as those made by fine grinding, lapping, polishing, honing, and electro-polishing, have random profiles. Since the

importance of smooth engineering surfaces, as well as their production costs are extremely high, the surface quality control becomes increasingly significant. Stylus and other surface measuring instruments have been designed for this purpose, but their measurement results can be divergent. One of the important reasons is that most smooth engineering surfaces have average spatial wavelengths falling in the range of the stylus size itself (~ 10 μm or less). Therefore, measured results are very sensitive to the stylus size. In addition, their surface texture may vary widely from place to place.

In 1986, random profile precision roughness calibration specimens were developed by Song at the ChangCheng Institute of Metrology and Measurement (CIMM) Beijing, China, to aid in identifying various effects in smooth engineering surface measurement [47]. These specimens are similar to PTB random profile roughness specimens but have smoother values of roughness. The measuring areas are composed of several (4–8) identical unidirectional random profile surfaces side by side, and their profile repetitions are equal to the recommended evaluation lengths for measuring them (see fig. 10). The specimens are a set of four pieces having R_a values of 0.1, 0.05, and 0.025 μm (these three with profile repetition of 1.25 mm) and 0.012 μm (with profile repetition of 0.4 mm).

There are also two smooth reference surfaces at opposite ends of the measuring area (see fig. 10). The smooth reference surfaces were made to be situated on the mean lines of the random roughness profiles and have $R_a \leq 0.005$ μm and flatness ≤ 0.01 μm. The reference surfaces provide a mechanical measuring datum for a skid to move on. By this means, the mechanical noise of the stylus instrument can be reduced to a minimum [47] because the smooth reference datum, rather than being a slide located in the drive box, is located very close to the stylus. This consideration is similar to minimizing the Abbe offset in dimensional measurement. Lines of intersection between the smooth reference surfaces and the measuring area also provide a datum for the start and end positions of the random profiles. It makes possible the comparison between the profile graphs obtained with various instruments.

PTB roughness specimens, with their relatively longer peak spacing do not yield R_a results that are sensitive to the stylus size [48] (also see fig. 11a). However, CIMM specimens, with mean peak spacing less than 10 μm are very sensitive to the stylus size (see fig. 11b), and could be used for checking stylus condition. To obtain agreement between results on CIMM specimens measured with various instruments, the reference conditions, especially the stylus size, should be carefully adhered to. Like the PTB specimens, the CIMM specimens simulate the situation of measuring real, smooth engineering surfaces and provide a means of exposing the problems or obtaining agreement when stylus and other instruments are used for measuring smooth engineering surfaces.

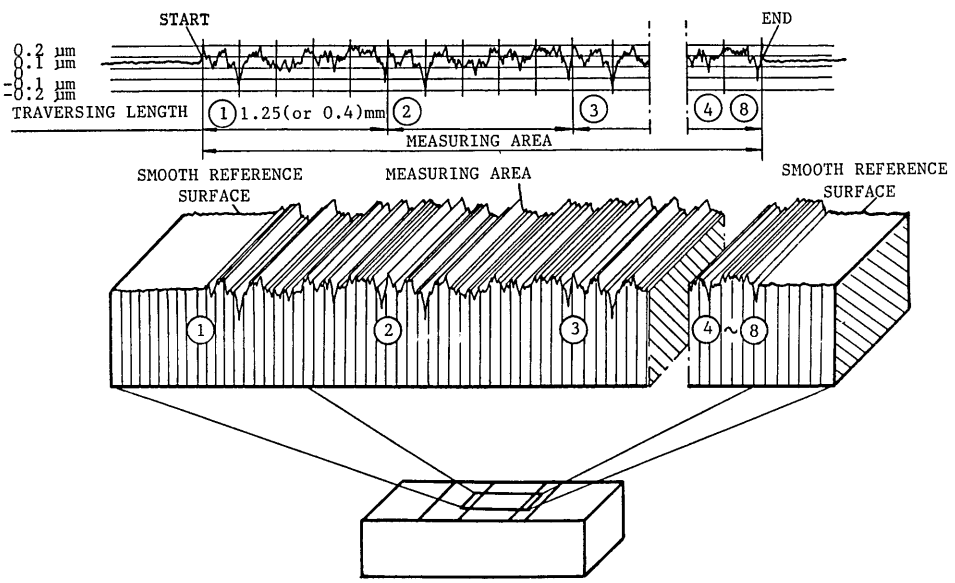

Figure 10. Random profile precision roughness calibration specimen.

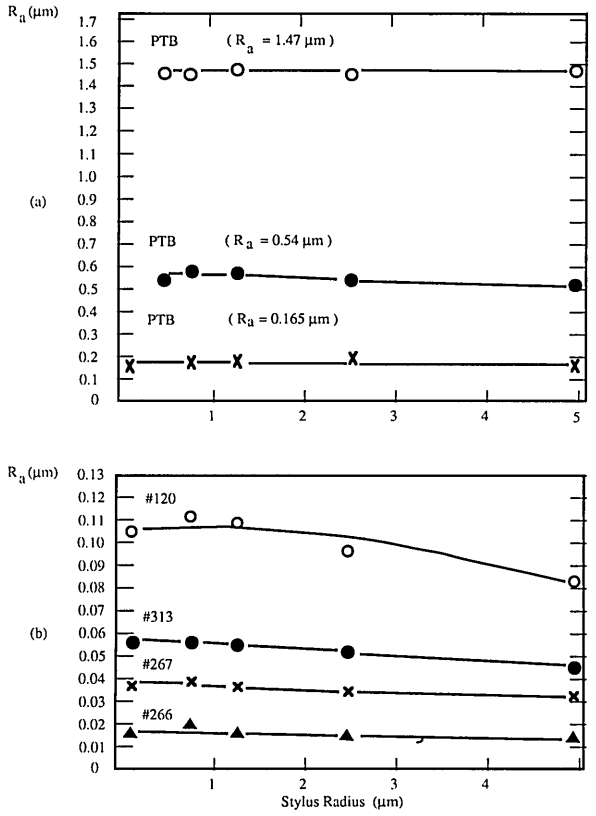

Figure 11. The effects of stylus size of various random profile specimens.
a) PTB specimens;
b) CIMM specimens.

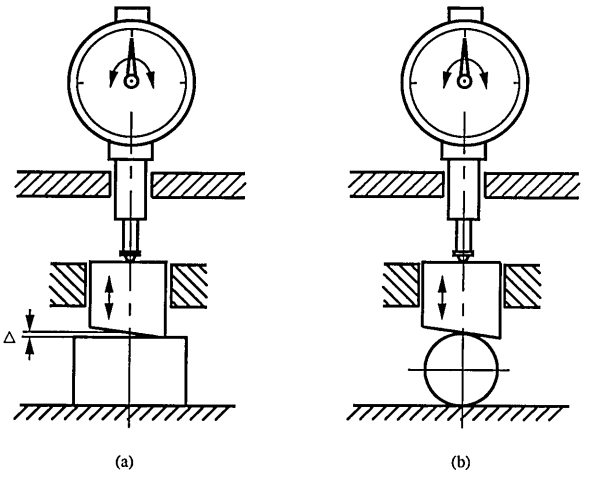

Figure 12. The reference conditions used in the calibration should be as close as possible to those used in the measurement.
a) calibration;
b) measurement.

3. Calibration and Measurement Procedures

3.1 Coordinating Calibration and Measurement Conditions

One of the important principles in metrology is that the reference conditions in the calibration of the instrument should be as close as possible to those used in the measurement. One example comes from length measurement. Having been calibrated by a calibration block, a height comparator could be used for the measurement of the diameter of a ball (see fig. 12). However, the parallelism between two measuring surfaces has a significant effect on the measuring error. If instead, a certified standard ball with the same size and material as those of the measured balls was used for calibrating the comparator, the error from the parallelism between the two measuring surfaces could be minimized, as well as any error due to the elastic deformation.

Another example comes from surface profile measurements. It is well known that a stylus with large width or radius distorts the measured profile seriously and can not be used for measuring the fine structure of smooth surfaces. But certain damaged styli with jagged tips could measure the fine structure of smooth surfaces very well. However, when a rough surface is measured, this type of stylus could cause serious profile distortion. Such a damaged stylus is shown in figure 13a. Using it to measure a wafer specimen [42] with $R_a = 0.0425$ μm and $D = 20$ μm, we obtained a perfect profile graph (fig. 13b), since only the sharp tip (segment B, fig. 13a) contacted the surface. When we measured a rectangular profile roughness specimen with $R_a =$ 0.91 μm and $D = 80$ μm, we obtained the distorted profile shown in figure 13c because segment C (fig. 13a) of the damaged stylus was involved. The profile was also seriously distorted when measuring a 3.042 μm step height specimen because the damaged stylus profile contacted the surface at several different points in succession. This is shown by segment *D*, figure 13a and d. If only the smoothest rectangular profile had been used to check the stylus tip, we would not have detected the problem of the stylus condition for measuring rougher surfaces. This demonstrates once again the importance of having calibration and checking conditions corresponding to the measurement conditions. Therefore, the calibration or check standards, should be as similar as possible to the measured surface.

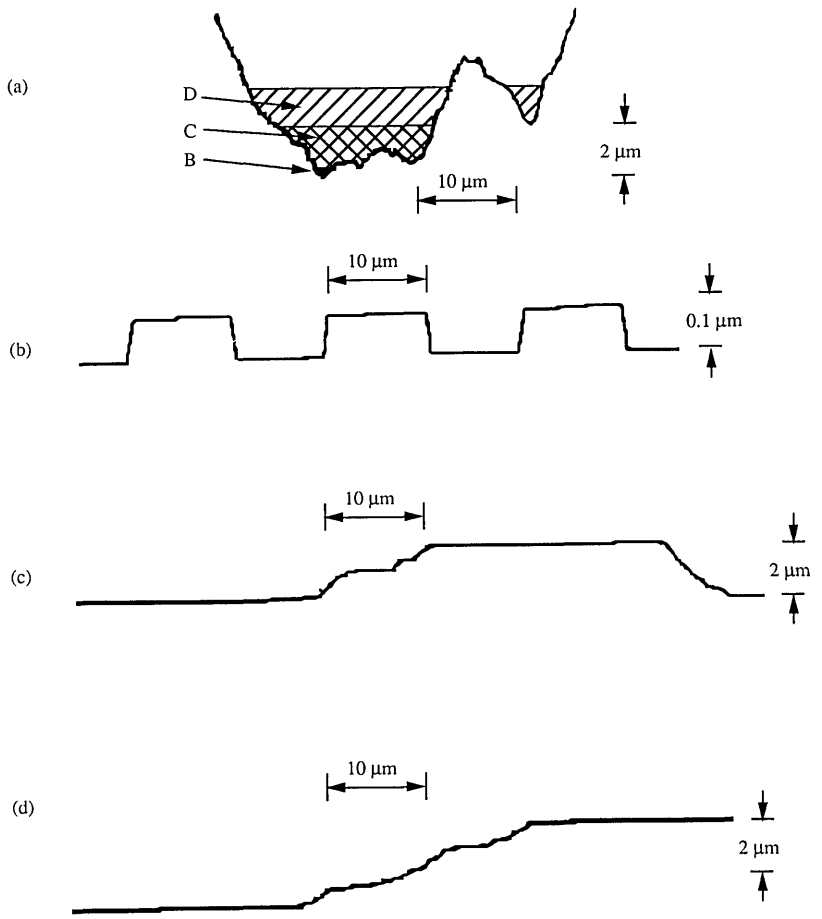

Figure 13. Profiles measured with a defective stylus having a jagged tip profile.
a) the profile of the damaged stylus;
b) the measured profile of a rectangular roughness specimen, $R_a = 0.0425$ μm; $D = 20$ μm;
c) the measured profile of a rectangular roughness specimen, $R_a = 0.91$μm; $D = 80$ μm;
d) the measured profile of a step height specimen, $H = 3.042$ μm.

3.2 Calibration Procedures at NIST

The above principle is being applied in the NIST surface calibration lab. In order to ensure that the calibration or measurement is correct, we measure a check specimen with profiles as similar as possible to the measured surfaces. For example, when calibrating our sinusoidal specimens SRM 2071–2075, we use the following procedure:

a) Calibrate the instrument with an interferometerically measured step height specimen, whose height about the same as the amplitude of the specimens to be measured;
b) Check the instrument calibration by measuring a certified step height or a certified sinusoidal specimen;
c) Measure the specimens under test;
d) Check the measurement again by measuring a check specimen as similar as possible to the measured specimens.

Using the check specimen and getting the results within a given tolerance helps to verify the entire set of measurements taken during a run. We also use roughness specimen calibration control charts, in which the measurement results and uncertainty of each calibrated specimen, as well as the calibration constant before and after these calibrations are recorded day by day. The control charts help us to maintain quality control in specimen calibrations and make these calibrations traceable to measurements taken several years ago.

By using the procedure above, we have calibrated sinusoidal specimens, SRM 207–2073. The resulting uncertainties on the R_a and wavelength values are, for example, (SRM 2072/Serial No. 1021), $R_a = 1.004 \pm 0.027$ μm ($\pm 2.7\%$) and $D = 101.64 \pm 0.24$ μm ($\pm 0.24\%$, both $\pm 3\sigma$). Other periodic profile specimens also have waveforms and sufficiently long wavelengths as to be insensitive to the stylus size. For these, we used the same procedure as for the calibration of our sinusoidal specimens, and a sinusoidal specimen with similar R_a and wavelength was used as a check specimen.

However, our calibrations of some rectangular profile specimens show variations of R_a with stylus tip size of several percent, even for specimens with a wavelength as long as 80 μm. In these specimen calibrations, we made a detailed report of the reference measuring conditions, especially the stylus size, on the calibration report. If the stylus size used in other labs differ from ours, a few percent difference in the R_a value could result.

3.3 Proposed Procedure for Engineering Surface Measurements

If the same procedure were used for engineering surface measurements, one of the important considerations is whether or not the wavelength domain of the measured profile falls in the flat portion of the transmission characteristics of the instrument, as at position X in figure 14 [37], where there is substantially no attenuation caused either by the stylus width at the high end of the spatial frequency spectrum, or by the electrical filter cut-off length λ_c at the low end. Otherwise, a big measuring error could occur. As a demonstration, by using the same procedure as that used for our sinusoidal specimens calibration, we measured the same smooth engineering surface with different stylus widths. The measured surface was a CIMM random profile specimen, with the mean spacing of profile irregularities [49] S_m equal to 6.6 μm. Our results were $R_a = 0.120 \pm 0.013$ μm and 0.165 ± 0.016 μm corresponding to stylus widths of 5 and 0.4 μm, respectively. The two contradictory measuring results should be equally valid, since both of them passed the checking procedure with a sinusoidal specimen as a check specimen which was not sensitive to the stylus size at all. Therefore, if a CIMM random profile specimen were used as a check specimen instead of the sinusoidal specimen, we could spot a change in the stylus condition when measurements of the check specimen are compared with previous ones as on a control chart.

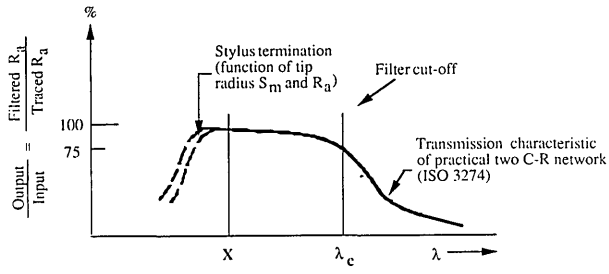

Figure 14. Schematic of the transmission characteristics of the stylus instrument.

The suggested procedure to be used for engineering surface measurement, as well as in the calibration of reference specimens with a stylus instrument is as follows:

1) Calibrate the vertical magnification of the instrument using a calibrated step height specimen whose height covers the range of the amplitudes of the measured profile;
2) Verify that the calibration was correct by measuring a certified step height, or an R_a value of a calibrated roughness specimen, such as a sinusoidal specimen;
3) Measure the engineering surface or specimen to be calibrated;
4) Check the measurement by measuring a check specimen with the same or similar waveform to that of the measured surface. The R_a value of the check specimen should have been measured under standardized reference measuring conditions.

In addition the reference conditions for the stylus instrument measurement, such as filter setting, stylus loading, and straightness of the mechanical motion, should also be carefully checked periodically.

Existing roughness calibration specimens could be used as the check specimens for a wide range of engineering surface measurements. For example, when the measured engineering surfaces have mainly periodic profiles, such as those obtained by turning, planing, or side milling processes, the periodic roughness specimens with triangular, cusped-peak, and sinusoidal profiles could be used as check standards. When the measured engineering surfaces have random profiles, as is obtained by grinding, lapping, polishing and honing processes,

the PTB and CIMM random roughness specimens could be used. They would cover the range of R_a values from 1.5 to 0.012 μm. If the checking measurement shows that the difference between the measured result for the check specimen and its certified value under reference conditions was within a given tolerance, the measurement of the engineering surface is considered to be under good quality control.

At present, different stylus size standards are adopted in different countries, for example, a 10 μm radius in the U.S. [27] and 2 μm radius in ISO and Europe [39,2]. However, there is not a uniformly defined method for checking the effective stylus size in the workshop. Therefore, it is difficult to get agreement in measurements of smooth engineering surfaces. However, the CIMM specimens calibrated by a reference instrument under standardized reference measurement conditions with a fine stylus size could be used as check standards. They make it possible to validate the results of smooth engineering surface measurements and obtain agreement by various instruments.

4. Conclusions and Recommendations

1) Engineering surface quality control, which involves the cycle of manufacture, texture measurement, and functional usage of the engineering surface, consists in choosing the right parameters to measure and obtaining knowledge of the right control values, as well as in maintaining the accuracy of the engineering surface measurement.
2) By controlled experiments, the right surface parameters and their controlled values could be determined for each application. If the manufacturing process or the surface function changed, a new controlled experiment should be performed so that a clear relationship between surface manufacture and function could be maintained. During these controlled experiments, new functional parameters, such as the R_k family, could be adopted.
3) Accuracy of engineering surface measurement would be aided by the establishment of a Reference Surface Metrology Instrument (RSMI), which would verify the standardized reference measuring conditions, and by the use of various calibration and check specimens with their certified parameters measured with the RSMI. A correct measurement procedure would include the choice and use of the check specimen for engineering surface measurement.
4) The check specimen's waveform, and parameter values should be similar to those of the measured engineering surfaces. The parameter value certified by the RSMI (or under standardized reference measuring conditions), provides an overall checking mechanism for all processes in the instrument including stylus tip, filter, and gain. The measuring results are accepted only when the check specimen is measured with the same instrument and a result within a given tolerance of its certified value is obtained.
5) Some precision roughness calibration specimens could be used as check specimens for a wide range of engineering surface measurement, from periodic to random profile, from several μm to 10 nm R_a. But there are still some new specimens that should be developed. These include specimens for testing stylus conditions, and specimens for testing scanning tunneling microscopes and atomic force microscopes to be used in super-smooth surface measurements.
6) With the development of the functional parameters, such as the Frech R & W [50,51] and German R_k family [34,35], the function of engineering surfaces could be assessed more quantitatively. Therefore, it may be possible to optimize functional performance of engineering surfaces by designing their surface texture, material, manufacturing processes, and quality control procedures. We call this combination the "surface texture design." For example, during the finishing of engine cylinder bores by the plateau honing process following the honing process [52], an optimal cylinder bore surface texture could be obtained. This surface has a negatively skewed profile height distribtution, (negative R_{sk}) [49] which results in good performances for running-in, long-term running and lubrication of the cylinder bores. The R_k family of parameters including algorithms and software have been developed for the measurements of such engineering surfaces. Meanwhile, some new specimens should also be developed to simulate these designed surface textures and used as check specimens in engineering surface quality control. These specimens could be manufactured by combining the manufacturing techniques of the PTB and CIMM specimens.
7) Replication techniques make it possible that the "same" engineering surface could be used as a check specimen on a great number of

instruments. Various electro-formed replica specimens with periodic profile have been successfully used in engineering surface measurement. However, for random profile roughness calibration specimens, especially those of $R_a \leq 0.1$ μm, the agreement between the original and the replica specimens would have to be carefully investigated so that the comparison among various instrument measurements could be based on the "same" smooth engineering surface.

5. Appendix

We discuss here certain difficulties with using the ISO Standard 5436 for step height measurement on curved surfaces. According to ISO 5436, the depth of a groove (fig. 15) or the height a step is determined as follows:

"A continuous straight mean line equal in length to three times the width of the groove (3w) is drawn over the groove to represent the upper level of the surface and another (w/3) to represent the lower level, both lines extending symmetrically about the centre of the groove (see fig. 15).

"To avoid the influence of any rounding of the corners, the upper surface on each side of the groove is to be ignored for a length equal to one-third of the width of the groove. The surface at the bottom of the groove is assessed only over the central third of its width. The portions to be used for assessment purposes are therefore those shown at A, B, and C in figure 15.

"The depth *d* of the groove shall be assessed perpendicularly from the upper mean line to the midpoint of the lower mean line.

"*NOTE*: The depth *d* as described here will be equal to the mean of the portion C below the upper mean line."

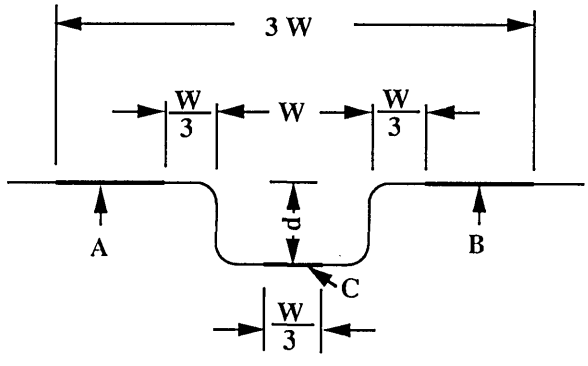

Figure 15. Assessment of step height according to the ISO 5436 Standard.

We now consider the effect when the profile graph is distorted into a curve, perhaps because of the straightness error of the traverse mechanism of the stylus instrument, or because of surface curvature.

As before, the step height *d* is determined by the distance between two parallel least square lines ab and cc, both of them calculated from the curved profile, parts A, B, and C (fig. 16).

Since $R >> H$, where H is the true step height, the error Δ caused by the distorted profile could be determined by the distance between lines c'c' and ab, where c'c' is a least square line calculated from the curve Part C', which is situated on the same curve with parts A and B, and parallel to the profile part C with a distance H (the true step height). Therefore, we have

$$\Delta = d - H = k - k'$$

where k and k' represent the positions of the least square lines ab and c'c' with respect to the origin 0.

The deformed profile, parts A and B, could be represented by

$$Y = R - (R^2 - x^2)^{1/2} \ .$$

The least square line $\bar{y} = k$ could be calculated from

$$f(x) = \int_{5w/6}^{3w/2} (y - \bar{y})^2 \, dx = \min,$$

$$f(x) = \int_{5w/6}^{3w/2} [(R-k)^2 - 2\,(R-k)\,(R^2-x^2)^{1/2}$$

$$+ (R^2 - x^2)]\, dx$$

$$= \left\{ (R-k)^2 x - (R-k) \left[x\left(R^2 - x^2\right)^{1/2} \right.\right.$$

$$\left.\left. + R^2 \sin^{-1} \frac{x}{R} \right] + R^2 x - \frac{1}{3} x^3 \right\}_{5w/6}^{3w/2} .$$

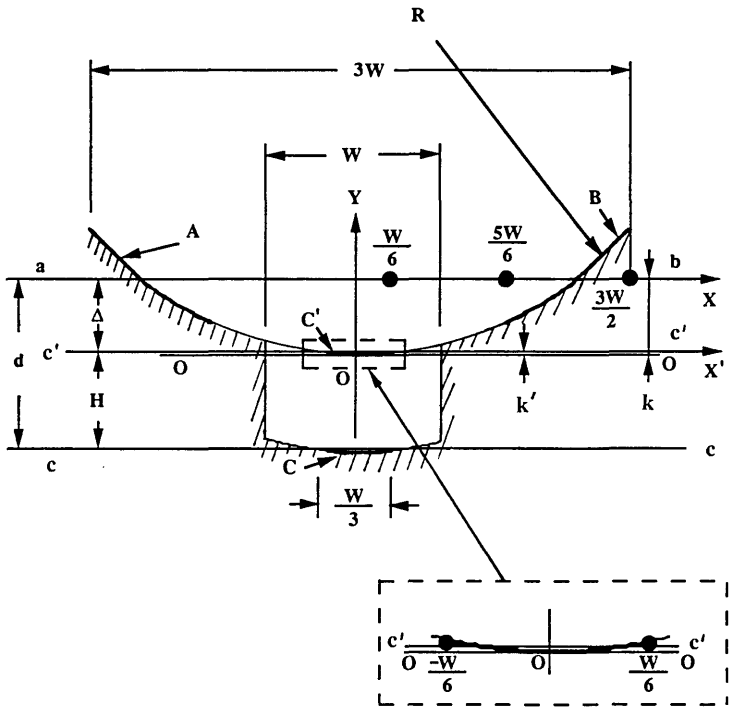

Figure 16. The error of step height when the profile graph is distorted into a curve.

Since $R >> x$,

$$f(x) \approx \left\{(R-k)^2 x - (R-k)\left[Rx(1-\frac{x^2}{2R^2}) + Rx\right] + R^2 x - \frac{1}{3}x^3\right\}_{5w/6}^{3w/2}.$$

Let

$$\frac{\partial f(x)}{\partial (R-k)} = 0 .$$

Then,

$$\left[2(R-k)x - 2Rx(1-\frac{x^2}{4R^2})\right]_{5w/6}^{3w/2} = 0,$$

$$\frac{4}{3}w(R-k) - 3Rw\left(1-\frac{9w^2}{16R^2}\right) + \frac{5}{3}Rw(1-\frac{25w^2}{144R^2}) = 0,$$

$$\frac{4}{3}w(R-k) - \frac{4}{3}Rw + \frac{151}{108}\frac{w^3}{R} = 0,$$

$$R-k = R - \frac{151}{144}\frac{w^2}{R},$$

$$k = \frac{151}{144}\frac{w^2}{R} .$$

The deformed profile Part C′ could also be determined by

$$y' = R - \left(R^2 - x'^2\right)^{1/2}.$$

The second least square line $\bar{y}' = k'$ could be calculated from

$$f'(x')\int_0^{w/6} (y' - \bar{y}')dx = \min.$$

Let

$$\frac{\delta f'(x)}{\delta (R-k')} = 0 .$$

Then

$$\left[2\,(R-k')\,x-2\,Rx\left(1-\frac{x^2}{4R^2}\right)\right]_0^{w/6}=0\,.$$

$$k'=\frac{1}{144}\frac{w^2}{R}\,.$$

Therefore,

$$\Delta=k-k'=\frac{(151-1)}{144}\,\frac{w^2}{R}=\frac{25}{24}\frac{w^2}{R}\approx 1.04\frac{w^2}{R}.$$

Since $R>>x$, the radius R could be calculated by observing the depth or height (y) of the curvature over the entire profile and using the formula below:

$$R\approx\frac{x^2}{2y}\,.$$

From figure 2b, y is calculated by construction to be 0.136 μm. We therefore calculate $R=3676$ mm. As a result

$$\Delta=1.04\frac{w^2}{R}=0.283\ \mu\text{m}\ (+93\%).$$

A similar result is obtained for figure 2c.

6. References

[1] Whitehouse, D. J., Wear **83,** 75 (1982).
[2] BS 1134, Method for the assessment of surface texture, Part 2, London, British Standards Institution (1972).
[3] Reason, R. E., Wear **57,** 1 (1979).
[4] Schneider, E. J., Wear **57,** 17 (1979).
[5] Scott, P. J., Surface Topography **1,** 153 (1988).
[6] Stout, K. J., and Thomas, T. R., eds. Proc. Int. Conf., Lausanne, Elsevier Sequoia (1979) p. vii.
[7] Vorburger, T. V., Ann. CIRP **36,** 2 (1987).
[8] Bennett, J. M., and Mattsson, L., Opt. Soc. Am. (1989).
[9] Bennett, H. E., and Porteus, J. O., J. Opt. Soc. Am. **51,** 123 (1961).
[10] Vorburger, T. V., Teague, E. C., Scire, F. E., McLay, M. J., and Gilsinn, D. E., J. Res. **89,** 3 (1984).
[11] Lieberman, A. G., Vorburger, T. V., Giauque, C. H. W., Risko, D. G., and Rathbun, K. R., Surface Topography **1,** 115 (1988).
[12] Blessing, G. V., and Eitzen, D. G., Surface Topography **1,** 143 (1988).
[13] Vorburger, T. V., and Teague, E. C., Prec. Eng. **3,** 61 (1981).
[14] Lange, S. R., and Bhushan, B., Surface Topography **1,** 205 (1988).
[15] Biegen, J. F., and Smythe, R. A., Surface Topography **1,** 287 (1988).
[16] Bristow, T. C., and Arackellian, K., Proc. SPIE **749,** 114 (1987).
[17] Young, R., Ward, J., and Scire, F., Rev. Sci. Instrum. **43,** 999 (1972).
[18] Binnig, G., Rohrer, H., Gerber, Ch., and Weibel, E., Phys. Rev. Lett. **49,** 57 (1982).
[19] Quate, C. F., Physics Today **39,** 26 (1986).
[20] Golovchenko, J. A., Science **232,** 48 (1986).
[21] Hansma, P. K., Elings, V. B., Marti, O., and Bracker, C. E., Science **242,** 209 (1988).
[22] Binnig, G., Quate, C. F., and Gerber, Ch., Phys. Rev. Lett. **56,** 930 (1986).
[23] Meyer, G., Appl. Phys. Lett. **53,** 1045 (1988).
[24] Erlandson, R., McClelland, G. M., Mate, C. M., and Chiang, S., J. Vac. Sci. Technol. **A6,** 266 (1988).
[25] Prater, C. B., Drake, B., Hansma, H. G., Hansma, P. K., Masse, J., Elings, V. B., Dixon-Northern, B. L., and Peterson, C. M., Bull. Am. Phys. Soc. **35,** 760 (1990).
[26] Division of precision engineering, Braunschweig: PTB, p. 7.
[27] ANSI/ASME B46.1-1985, Surface texture-surface roughness, waviness, and lay, New York: Am. Soc. Mech. Eng. (1985).
[28] Young, R. D., and Scire, F. E., J. Res. Natl. Bur. Stand. (U.S.) **76C,** 21 (1972).
[29] Thomas, T. R., and Charlton, G., Prec. Eng. **3,** 91 (1981).
[30] Reason, R. E., CIRP/Annalen 12 (1962).
[31] MIL-STD-10A, Surface roughness waviness and lay, Washington, DC, Dept. of Defense (1955).
[32] DeVries, M. F., Field, M., and Kahles, J. F., CIRP **25,** 569 (1976).
[33] Osanna, P. H., Wear **57,** 227 (1979).
[34] DIN/4776-1985, Berlin, Deutches Institut fur Normung (1985).
[35] Bodschwinna, Ing., Rauheitskennwerte aus der Abbottkurve zur funktions-bezogenen beschreibung der oberflachengestalt, in *VII Internationales Oberflachenkolloquium*, ed. H. Trumpold, Karl-Marx-Stadt, Technisches Universität (1988), p. 146.
[36] Teague, E. C., Scire, F. E., and Vorburger, T. V., Wear **83,** 61 (1982).
[37] ISO 5436-1985, Calibration specimens—stylus instruments—types, calibration and use of specimens, Geneva, International Organization for Standardization (1985).
[38] Teague, E. C., Metrologia **14,** 39 (1978).
[39] ISO 3274-1975, Instruments for the measurement of surface roughness by the profile method—contact (stylus) instruments of consecutive profile transformation—contact profile meters, system M, Geneva, International Organization for Standardization (1975).
[40] Song, J. F., and Vorburger, T. V., Applied Optics **30,** 42 (1991).
[41] Vorburger, T. V., Teague, E. C., Scire, F. E., and Rosberry, F. W., Wear **57,** 39 (1979).
[42] Berger, J., Surface Topography **1,** 371 (1988).
[43] Song, J. F., and Vorburger, T. V., Stylus flight in surface profiling, (to be published).
[44] For a theoretical treatment on this subject, see Church, E. L., and Takacs, P. Z., Effects of the non-vanishing tip size in mechanical profile measurements, Proc. SPIE 1332 (to be published).
[45] Häsing, J., Werkstattstechnic **55,** 380 (19)65.

[46] Hillmann, W., Comparison of roughness measurements in the European community, Commission of the European Communities, Community Bureau of Reference, BCR. Report, EUR 12 180 EN (1989).
[47] Song, J., Surface Topography **1**, 303 (1988).
[48] Hillmann, W., Technisches Messen **47**, 209 (1980).
[49] ISO 4287-Part 1, Surface roughness—terminology—surface and its parameters, Geneva, International Organization for Standardization.
[50] Bielle, J., Prec. Eng. **7**, 31 (1985).
[51] Fahl, C. F., Wear **83**, 165 (19)82.
[52] Davis, E. J., Sullivan, P. J., and Stout, K. J., Surface Topography **1**, 63 (1988).

About the authors: *J. F. Song was educated at the Harbin University of Technology, China, and holds a masters degree in Precision Engineering. He headed the surface metrology effort at the ChangCheng Institute of Metrology and Measurement before coming to NIST as a Guest Researcher in the Precision Engineering Division of the Manufacturing Engineering Laboratory. T. V. Vorburger holds a Ph D in Physics from Yale University. He has been working in surface metrology at NIST for 15 years and is group leader of the Surface and Particle Metrology group in the Precision Engineering Division.*

Volume 96, Number 3, May-June 1991
Journal of Research of the National Institute of Standards and Technology

[J. Res. Natl. Inst. Stand. Technol. 96, 291 (1991)]

An Automated Reverse-Bias Second-Breakdown Transistor Tester

Volume 96 | Number 3 | May-June 1991

David Berning

National Institute of Standards and Technology,
Gaithersburg, MD 20899

An automated instrument is described for generating curves for the reverse-bias, safe-operating area of transistors nondestructively. A new technique for detecting second breakdown that makes automation possible is highlighted. Methods to reduce stress to the device under test are discussed, as are several other innovations that enhance automation. Measurements using the tester are described, and limitations on nondestructive testability are discussed.

Key words: automated testing; clamp overshoot; crowbar; fast switching; nondestructive; protection circuit; reverse-bias; safe-operating area; second breakdown; transistor.

Accepted: March 5, 1991

1. Introduction

High-voltage, power-switching transistors are used in a wide variety of applications including such diverse fields as power conversion, motion control, and electronic ignition. A critical element for specifying the performance and reliability of these transistors is their turn-off capability. Most applications using high-voltage switching transistors require the transistor to turn off from a state of high-current conduction at low voltage with a load circuit that may be somewhat inductive. Often during turnoff, the voltage rises to a high value before the current begins to fall, and there is a period of time when the transistor experiences a very high level of peak power dissipation. If a particular combination of current and voltage exceeds the switching capability of the transistor, it may enter second breakdown and be destroyed. It is desirable to have test apparatus that can test the transistors by simulating conditions that are typical of actual usage in circuits, and it is especially desirable that the testing be nondestructive so that one transistor can be used to determine a safe operating area curve.

Equipment for testing the turn-off switching capability of high-voltage switching transistors has been previously described [1–3], and various data obtained from such testers have been discussed [4, 5]. These previous testers are difficult to use because each test requires multiple set-up steps and once a breakdown is observed, subjective interpretation is required. This paper describes a new tester that was developed to automate the process of making measurements of reverse-bias, safe-operating area. The tester was designed as a stand-alone instrument that can be used manually, or with a computer that has an IEEE-488 interface controller. This paper concentrates on the special techniques required to automate the measurement of reverse-bias safe operating area. The circuit details, including a complete set of schematic drawings, are published elsewhere [6].

2. The Automated Tester

Figure 1 is a block diagram of the tester. This tester, like others, works by connecting the device under test (DUT) in a common emitter (source) configuration with a voltage supply and load inductor in the collector (drain) circuit. The DUT is turned on for a period of time sufficient to charge the inductor to the desired test current. The device is then turned off. The collapsing field in the inductor causes the collector voltage to rise to a level where the device may break down. As the voltage rises across the DUT, either the voltage will be limited by an external clamp imposed by the tester, allowing the device to safely turn off, or the device may begin to avalanche with or without entering normally destructive second breakdown. Second breakdown is characterized by a sudden collapse of voltage. If the DUT does experience second breakdown, it will be destroyed unless the current and voltage are removed very quickly. The success of making breakdown measurements is critically dependent on the speed of diverting current away from the DUT after the onset of second breakdown. The automated tester incorporates a fast breakdown detector and shunting "crowbar" circuit that diverts up to 100 A of test current from the DUT within 65 ns of device voltage collapse. The time to divert the current includes both a circuit propagation delay and current fall time and decreases to about 30 ns for test currents under 40 A. The voltage slew rate of the crowbar is 200 V/ns. The maximum test voltage is 2000 V, and the clamp voltage can be set at any level up to this maximum voltage.

The DUT base (gate) drive circuits are both constant current sources that can source and sink up to 25.5 A each for device turn-on and turn-off, respectively. A settable drive clamping circuit allows voltage limits to be imposed to prevent base-emitter breakdown and provides voltage drive when testing MOS gated devices.

Although the specific details and performance capabilities of the various earlier testers differ from those that characterize this tester, most of the testers have the same basic building blocks as those described up to this point. To automate the reverse-bias test, some additions and refinements over manually operated testers are needed. The most important improvements that must be made involve the protection circuit that detects second breakdown and diverts current from the DUT once breakdown occurs.

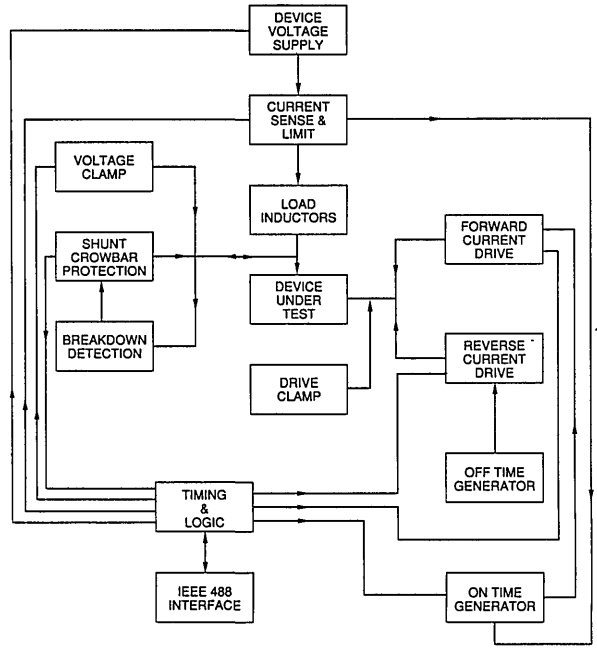

Figure 1. Block diagram of the tester.

2.1 Breakdown Detector

Presently, two methods are commonly used for nondestructive reverse-bias testing. One method uses a dV/dt detector that senses voltage collapse with a small capacitor which is coupled to an amplifier, which in turn drives a crowbar switch that shunts the current around the DUT. The other method is a pre-trigger scheme which always fires a crowbar during the test for breakdown. This scheme requires multiple tests whereby the trigger delay is adjusted in small increments until breakdown is observed. The first method is faster, in that multiple tests are not required to determine the presence of breakdown, and the second method is less demanding of the speed of the crowbar circuit. Figures 2 and 3 are oscillographs of device voltage and current waveforms taken by using the new tester that demonstrate problems with the above methods when automation is considered.

Figure 2a shows voltage and current waveforms when a 500-V power MOSFET is turned off very fast. A severe voltage overshoot reaching a peak of 480 V can be seen even though the clamp voltage was set to 180 V. However, no breakdown has occurred. The overshoot is caused by the parasitic inductance and diode turn-on delay in the clamping circuit in the presence of high dI/dt. Figure 2b shows voltage and current waveforms for actual

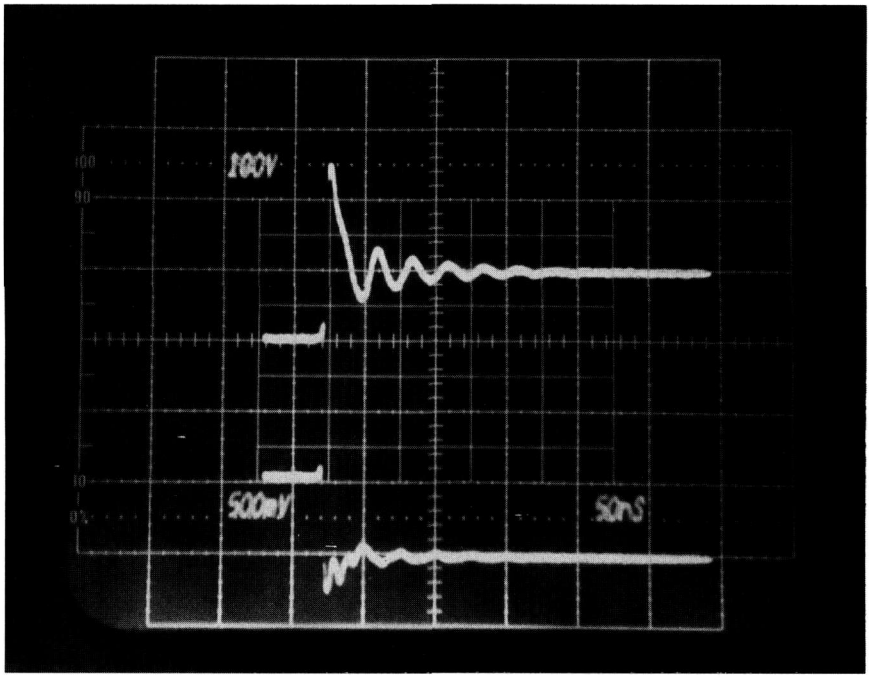

Figure 2a. Voltage and current waveforms showing voltage overshoot for fast MOSFET turn-off. Top trace: 100 V per small div.; bottom trace: 5 A per small div; time: 50 ns per small div.

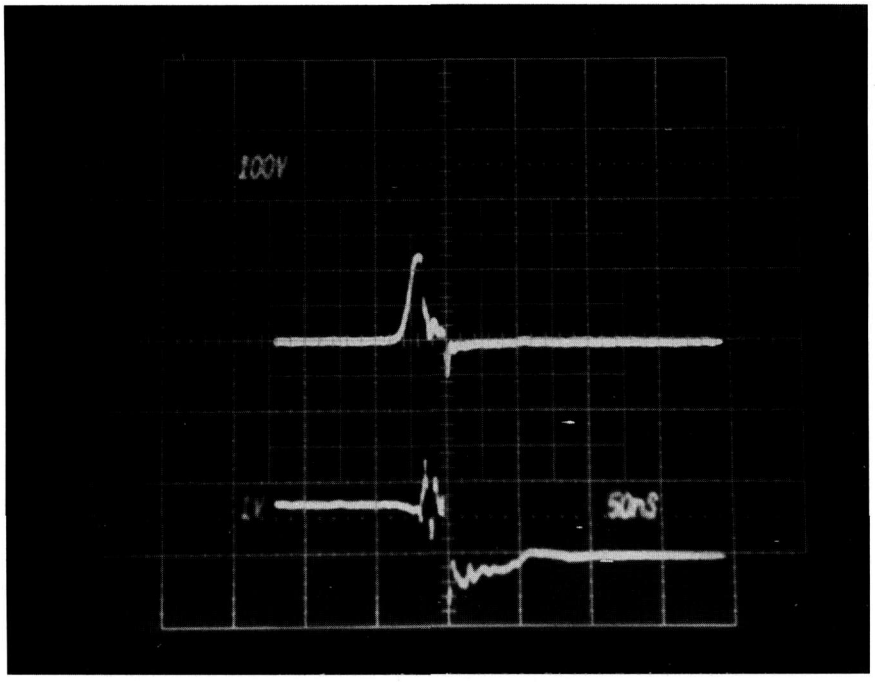

Figure 2b. Voltage and current waveforms for second breakdown in 200V MOSFET. Top trace: 100 V per small div.; bottom trace: 10 A per small div.; time: 50 ns per small div.

second breakdown in a 200 V MOSFET. A dV/dt detector that is adequately sensitive to detect the collapse of voltage as second breakdown in figure 2b will be triggered falsely by the overshoot in figure 2a where there was no breakdown. False indications of breakdown are clearly unacceptable when automating the test for reverse-bias safe-operating area.

Figure 3 shows voltage and current waveforms for a power MOSFET that sustains in avalanche for 220 ns before entering second breakdown. Automating a tester based on a pre-trigger scheme is very difficult for delayed breakdowns because subjective interpretation is needed to determine if the voltage collapse is a result of second breakdown or of the crowbar firing. A further problem is that the length of time that a device stays in the avalanche mode is often subject to time jitter from test to test.

A unique breakdown detector circuit was therefore developed that uses both voltage and current to determine the presence of second breakdown. Figure 4 is a simplified schematic of the breakdown detection and protection crowbar. Half of a dual triode is used as a diode detector, and the other half is used as a comparator. The (+) input of the comparator (cathode) is pulled negative upon DUT voltage collapse, thus pulling the output plate negative and firing the crowbar unless the (−) input (grid) is driven negative by output from a dI/dt sense transformer that determines the presense of increasing clamp current. Increasing clamp current thus blocks the firing. Vacuum tubes are used because of their inherent high-voltage capability and low interelectrode capacitance. Tubes are free of recovery problems, do not need overvoltage protection, and are unsurpassed in speed.

2.2 Crowbar Performance Enhancement Techniques

Figure 4 also shows some additional important features. There are two sets of clamp diodes with the crowbar placed between them. During testing, the diodes nearest the DUT are reverse-biased to the maximum extent possible to keep parasitic capacitance at the test fixture low. A large negative voltage is applied to the crowbar cathode to increase the speed of current diversion from the DUT. A reverse blocking diode and saturable inductor work together to reduce current reversal in the DUT when the crowbar fires. The actual crowbar circuit utilizes 16 vacuum tubes connected in parallel that conduct the current for several hundred nanoseconds, after which SCRs (not shown) take over the crowbar function. The DUT voltage

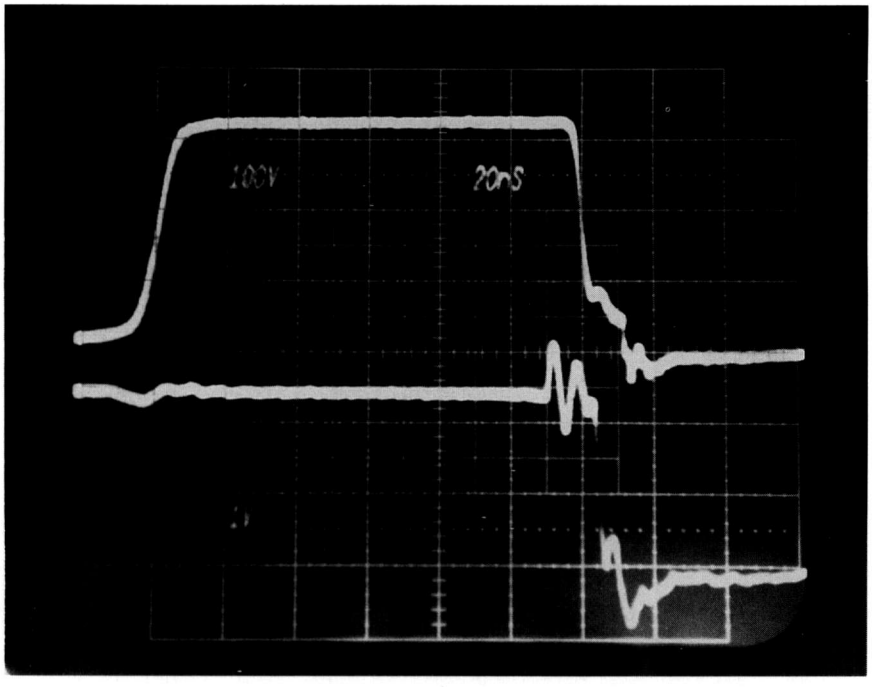

Figure 3. Voltage and current waveforms for sustained avalanche with second breakdown in MOSFET. Top trace: 100 V per small div.; bottom trace: 10 A per small div.; time: 20 ns per small div.

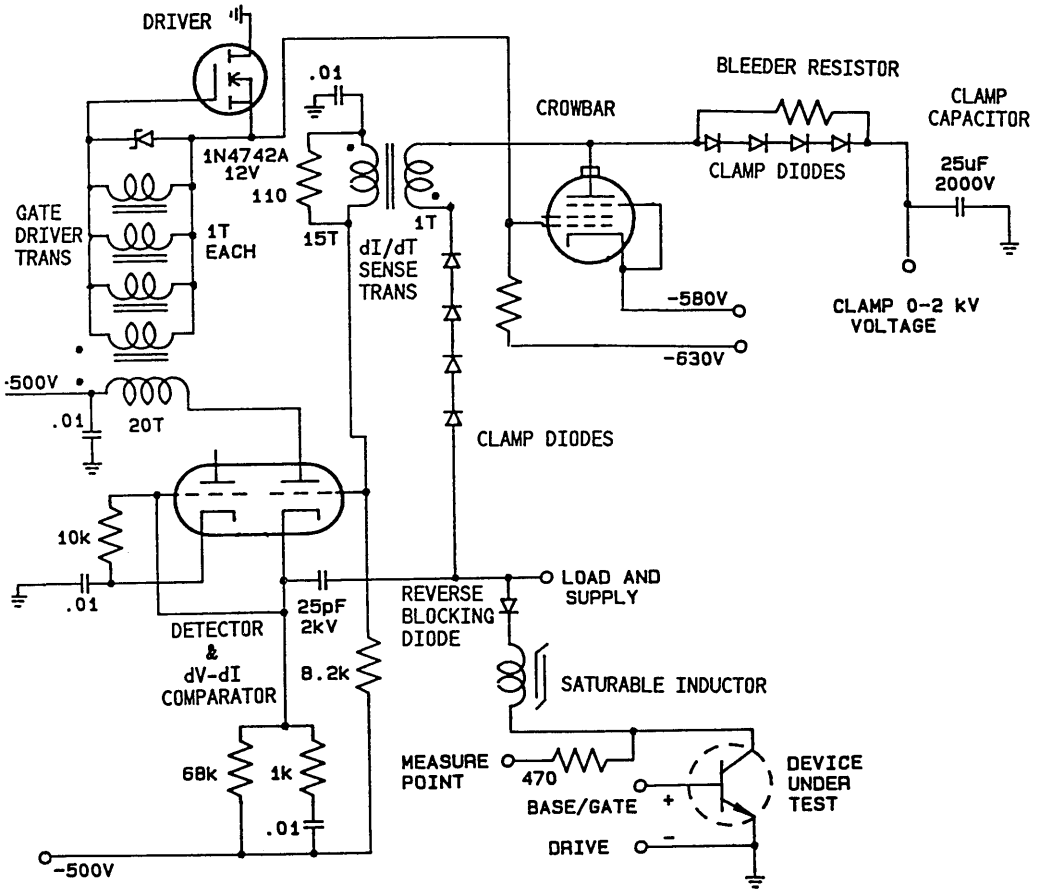

Figure 4. Simplified schematic showing critical elements of breakdown detector and crowbar circuit.

is measured through a 470 Ω resistor to reduce parasitic capacitive loading. This resistor causes a minor degradation to the bandwidth of the voltage measuring system.

2.3 Device Test Current and Test Load

In manually operated testers, a test current is usually set by running a series of tests with the clamp voltage set sufficiently low so as to prevent breakdown. The current is observed on an oscilloscope and increased or decreased to the desired level by adjusting the on-time or supply voltage. A desirable feature for automating the reverse-bias, safe-operating-area test is to have the ability to use the DUT current as an independent variable. Test current is determined by a number of factors, including the duration of the time that the DUT is turned on, the effective on-resistance of the device, resistances in the load circuit, and supply voltage. This tester incorporates a current limit detector which is coupled to the on-time generator. During either manual or automated tests, the on-time can be set to its maximum value, and as the DUT current ramps up and reaches a desired current set point, the on-time is terminated, the test is executed, and the new value of on-time is stored for subsequent tests.

While in principle the load for the DUT is simply an inductor, the use of several inductors with different values and saturation characteristics placed in series permits automated measurements over a wider range of currents. Figure 5 shows the device load used in the tester. The 100 μH inductor is linear up to the full 100 A test-current capability of the tester. The 300 μH inductor saturates at 15 A, and the 1 mH inductor saturates at slightly less than 1 A.

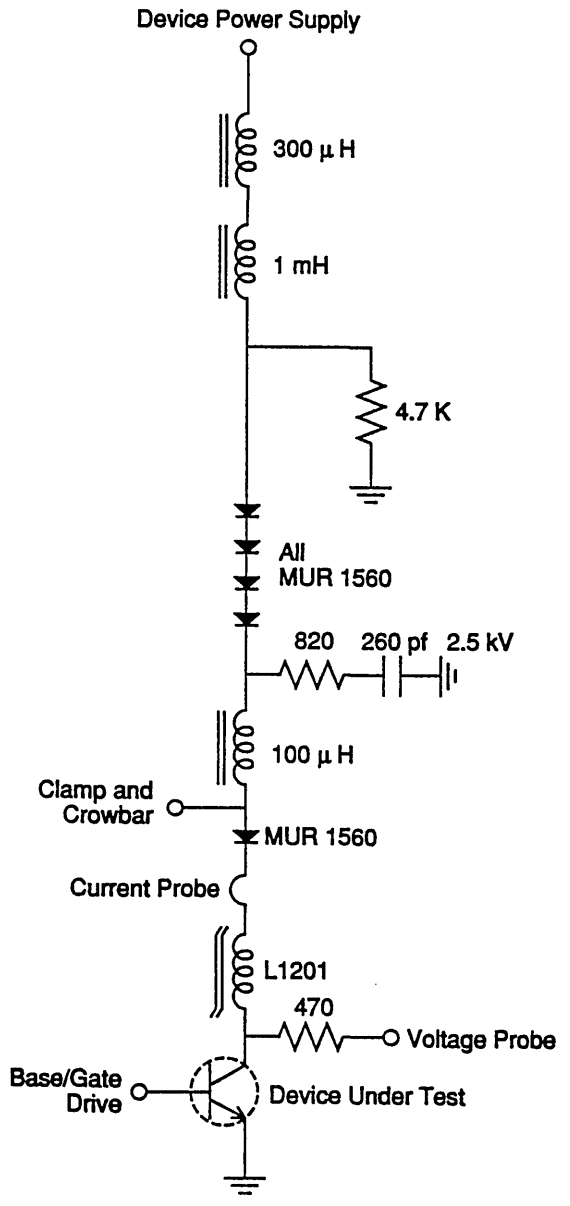

Figure 5. Load circuit for the Device Under Test.

The 1 mH inductor works in conjunction with the four diodes and the 260 pF capacitor to prevent the voltage on the DUT from rapidly snapping to zero when the current in the 100 μH inductor goes to zero. Such a rapid voltage transition would otherwise be detected as a device breakdown. The resistors associated with this L-C-diode network are used for damping. The 1 mH inductor needs to store only enough energy to assure that the 260 pF capacitor is left in a charged state when the current goes to zero. The effect of the 1 mH inductor on the breakdown test itself is of no consequence, as it is in saturation for all test currents of interest, and it only adds a delay in the ramping up of the current when the device turns on.

The effective load inductance for test currents is 400 μH for currents up to 15 A, and about 100 μH for currents between 15 and 100 A. The dual-inductor system enhances the accuracy of the current-limit circuit at the lower currents because dI/dt is reduced, and additional time resolution is provided to establish the proper on-time pulse width needed to achieve the desired set current.

2.4 Voltage Clamp Power Supply

A large clamp capacitor is needed to effectively clamp large currents while maintaining a nearly constant voltage. During the course of making automated measurements, it is desirable to change the clamp voltage as fast as possible, and a power supply is needed that can both source and sink a significant amount of power. During repetitive testing, the clamp current charges the clamp capacitor, and this charge must be removed. A two-quadrant switching amplifier that sources and sinks up to 2000 V was developed to meet these requirements.

The switching amplifier can deliver a power output of up to ±60 W, or ±30 mA at up to 2000 V. Negative power represents power absorbed by the amplifier (negative current, positive voltage), and most of this power is not dissipated as heat, but converted back to the rectified power mains. A simplified schematic of the amplifier is given in figure 6a. The amplifier is configured as a rectified voltage source in series with a constant current sink, with the output taken between these and applied to a low-inductance 25 μF oil-filled capacitor. The capacitor is located as close to the DUT as possible.

The voltage-source portion of the amplifier is a quasi-resonant converter, a topology chosen to offer both the advantage of the wide control range of pulse-width modulation (PWM), and the inherent current limiting of the parallel resonant converter. The PWM circuit is driven by a feedback signal that includes the output of the amplifier and the output of a controlling DAC. The current-sinking portion of the amplifier is a parallel resonant converter operating at resonance. The ac feedback keeps the circuit self-oscillating at resonance over the wide range of supply voltage at the clamp capacitor. The interesting property of this circuit is that on a dc basis it behaves like a constant current sink over a wide range of supply voltage. The output current of this converter is fed back to the rectified power mains that runs the entire tester through a suitable

transformer and rectifier, and is voltage-clamped by this rectified mains voltage. The output current returned to the rectified power mains is proportional to the clamp voltage. The operating frequency is about 110 kHz.

Figure 6b shows the output of the clamp supply when the tester makes a test at 2000 V. Once a test command is given, the clamp supply is gated on by changing the data fed to the DAC from zero to the desired test value. A time delay allows the voltage on the clamp capacitor to reach the set point before the breakdown test. Once the test is executed, the voltage returns to zero.

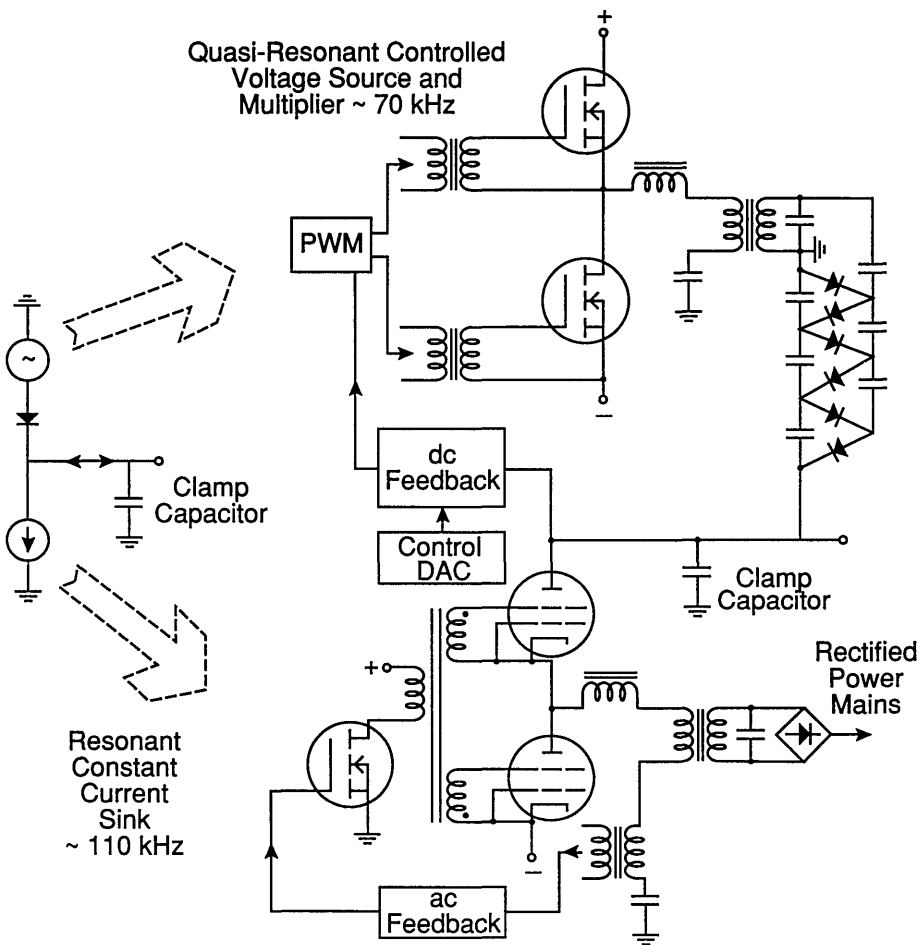

Figure 6a. Simplified schematic of two-quadrant switching amplifier for clamp supply. A conceptual circuit appears to the left, and the actual implementation is shown on the right.

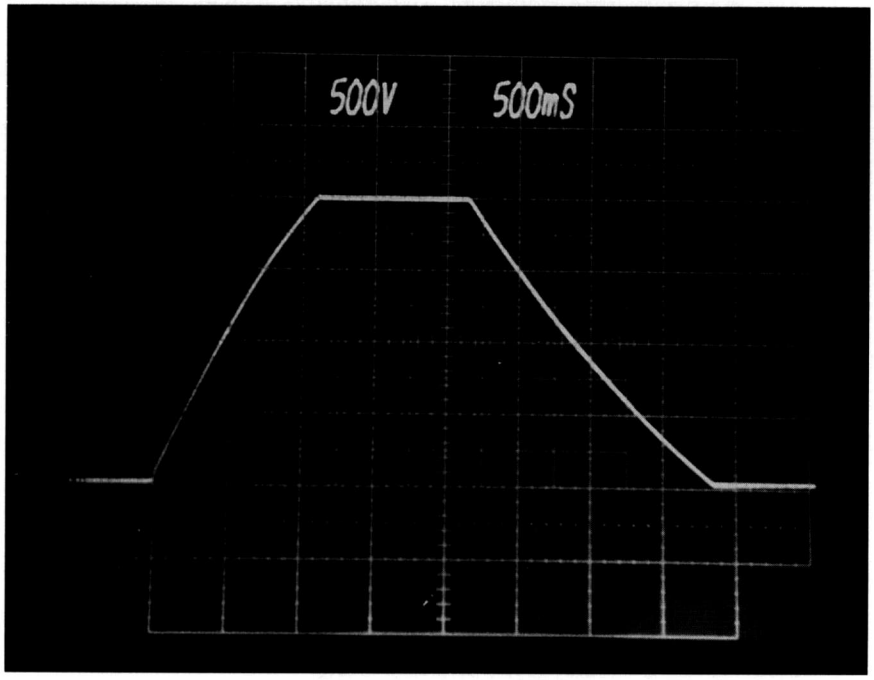

Figure 6b. Clamp voltage for a test at 2000 V. Once a test command is given, voltage on the clamp capacitor ramps up to the desired clamp voltage, the test is executed, and the voltage returns to zero. Scale: 500 V per div.; time: 500 ms per div.

2.5 Tester Architecture

The functions in the tester that are programmable include test current, on-time, turn-on drive current, turn-off drive current, clamp voltage, and test-start. The tester can return a "device failed" message. The first five parameters above are represented as 8 bit binary numbers in the tester, and can be set either remotely through an IEEE 488 interface, or on the tester with rotary optical encoders. An encoder and 7-segment-type display are provided for each parameter. The interface uses the Fairchild 96LS488 chip[1] which can be used in non-microprocessor, asynchronous systems such as this tester.

[1] Certain commercial equipment, instruments, or materials are identified in this paper to specify adequately the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

3. Breakdown Measurements

The measurement of the second breakdown voltage for a transistor can be done in two different ways, which can give two different numbers. One method is the unclamped measurement, whereby the clamp voltage is set well above the expected breakdown voltage. When the test is executed, the peak voltage at the point of voltage collapse is measured with a storage oscilloscope or a fast digitizer. The other method is the clamped measurement, whereby testing is begun by setting the clamp voltage well below the expected breakdown voltage, and incrementally raising the clamp voltage until second breakdown occurs. The clamp voltage is then equal to the breakdown voltage. The unclamped method generally gives a higher breakdown voltage indication than the clamped method. The unclamped method can give an artificially high number because a transistor can often withstand a higher voltage for a very short period of time before the voltage actually collapses. The clamped method can give an artificially low number because clamp overshoot can trigger second breakdown.

Figures 7a and 7b demonstrate the differences in the two measurement methods. In figure 7a, three different clamp settings cause three different responses for a bipolar transistor. For this figure, the transistor was turned off very hard, with a turn-off reverse base current of 4.8 A for a collector current of 6 A. For one trace, the clamp voltage was set well above the peak voltage recorded, which was about 640 V. Another trace was generated with the clamp set at 410 V. The voltage reached a peak value of about 510 V and the transistor voltage collapsed, but rather slowly compared to the unclamped case. A third trace was generated with the clamp set to 400 V, and the transistor turned off successfully, with the voltage reaching a peak of about 500 V. In figure 7b, the same transistor was tested with the same test conditions as in figure 7a, except that the turn-off current was reduced to a much more appropriate value of 1.0 A. Again, the first test was unclamped, with a peak recorded voltage of 550 V. When the clamp was set to 500 V, the transistor broke down with a peak voltage of about 510 V. With the clamp set to 490 V, the transistor did not break down and the voltage reached approximately 500 V.

In view of the above measurements, it is clear that some care must be taken when determining the second-breakdown, safe-operating area (SOA). Clearly, clamped measurements give a more conservative (lower voltage) SOA than unclamped ones, but clamped measurements can be overly conservative if the device is turned off too fast. The most accurate SOA is determined when the two methods are combined, by automating this tester with a programmable fast digitizer. To combine the methods, the clamp is raised incrementally until second breakdown occurs, and at the same time, the voltage waveform is digitized and the peak voltage reached is recorded for the breakdown.

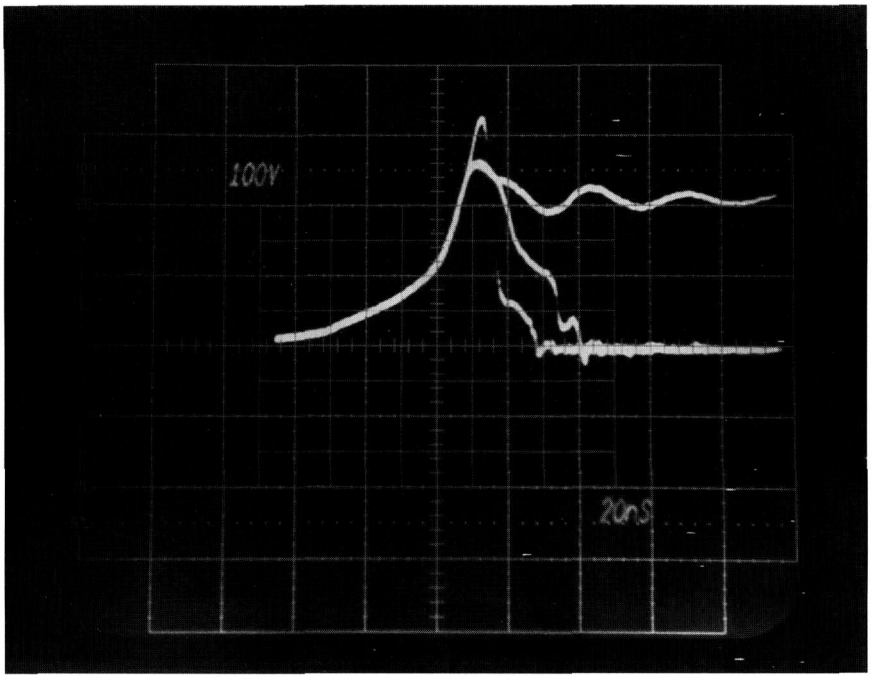

Figure 7a. Device collector voltage for three different clamp settings for high turn-off base current. The transistor breaks down for two of the clamp settings. Scale: 100 V per small div.; time: 20 ns per small div.

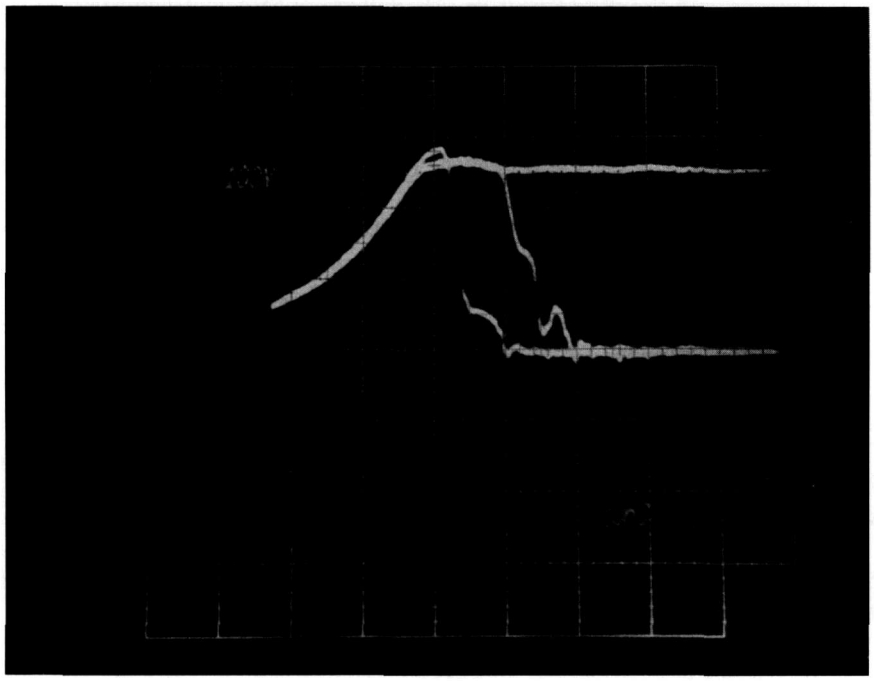

Figure 7b. Same device and conditions as above; however, the turn-off is reduced, and a different set of clamp voltages is used. The voltage difference between unclamped and clamped measurements is reduced for reduced turn-off drive.

Figure 8 gives some typical SOA curves for a bipolar transistor as measured by the tester under computer control. The measurements were made using the clamped technique without a digitizer, and the turn-off currents used were sufficiently low so as to avoid significant overshoots. One set of data was generated when the tester was programmed to make a series of tests with collector currents from 1 to 20 A, and the turn-on and turn-off base currents were adjusted for each collector current so that they were 1/5 the value of the collector current for turn-on and turn-off gains of 5. The other set of data was generated when the tester was programmed to run the tests for the same collector currents as previously used, but the turn-on and turn-off currents were maintained at the fixed values of 2.0 and 0.5 A, respectively. The data of figure 8 follow a general trend observed for bipolar transistors in that second breakdown occurs at lower voltages for higher collector currents, and also at lower voltages when higher turn-off currents are used.

Second breakdown data are given in figure 9 for a MOSFET. These data are representative of the SOA of MOSFETs, showing a nearly constant second breakdown voltage with drain current. For low currents, it is common for a device to avalanche without entering second breakdown as this device does. The data points represented as circles on the graph indicate that the device sustained avalanche without entering second breakdown when the clamp was set to 400 V (this voltage is not an indication of the actual voltage of the avalanche).

4. Limits On Nondestructive Testability

Some transistors avalanche for a relatively long period of time before entering second breakdown, and thus absorb much more energy than transistors that break down without much delay. When transistors sustain avalanches for microseconds, they are often degraded or destroyed by the reverse-bias SOA test. Figure 10 shows typical voltage (top) and current waveforms that are generated by such a transistor during the reverse-bias second breakdown test. The voltage across the device rises when it begins to turn off, but begins to level off as the device avalanches. No clamp is acting. The avalanche voltage rises as the device heats internally. The current ramps down during the avalanche period because the load inductor maintains a higher voltage on the node nearest the device. At some point the device enters second breakdown, and the current is brought to zero by the protection circuit.

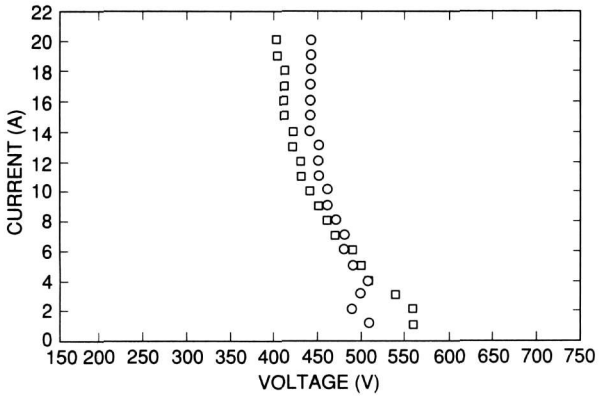

Figure 8. SOA curves measured by the tester under computer control for a bipolar transistor. The squares represent the SOA limit when the turn-on currents and turn-off currents are set to 1/5 the value of the collector current. The circles represent the SOA limit when the same transistor is tested with a fixed turn-on current of 2.0 A and a fixed turn-off current of 0.5 A.

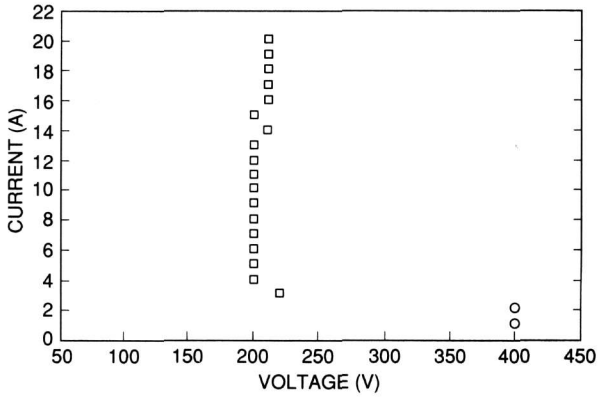

Figure 9. The squares represent the SOA curve for a power MOSFET. The circles at the lowest test currents indicate that the transistor did not enter second breakdown when the clamp voltage was set to 400 V, but rather sustained avalanche at a lower voltage.

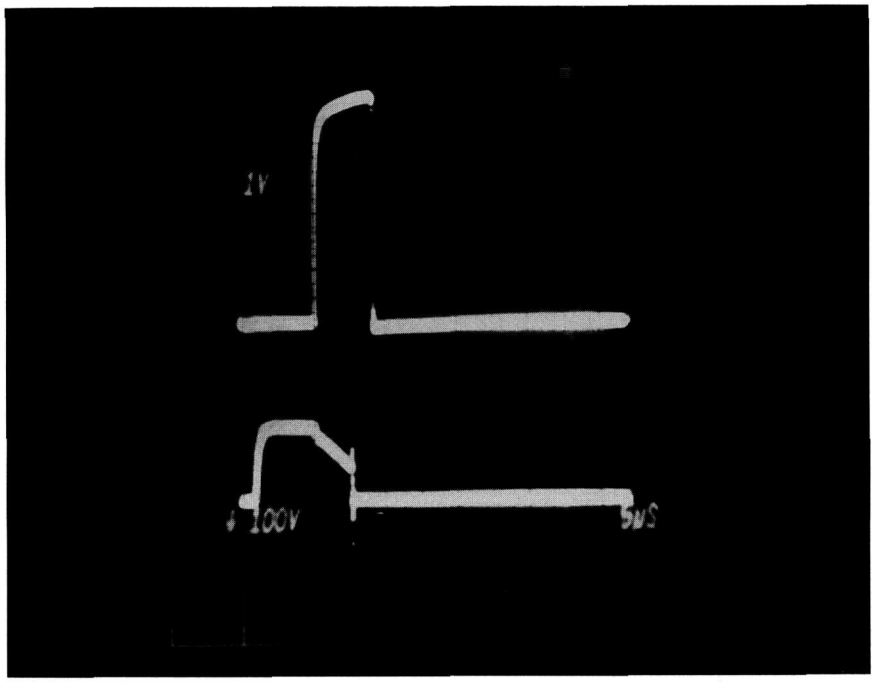

Figure 10. Voltage and current waveforms for a transistor that sustains avalanche for a relatively long time before entering second breakdown. Such behavior is often destructive, even when a protection circuit removes power very quickly. Top trace: 100 V per small div.; bottom trace: 10 A per small div.; time: 5 μs per small div.

Bipolar transistors can usually survive a long sustaining period, but MOS-type devices are generally degraded or destroyed. Past experience has always indicated that the success with which devices are saved is strongly dependent on the speed of the protection crowbar circuit. Although this automated tester has an extremely fast crowbar, it is not able to save many of the devices that have long sustain times. The question arises as to whether these devices could be saved if the protection were even faster.

There are also other possible stresses on the device when it enters second breakdown. The localized portion of the device where second breakdown is instigated absorbs additional energy during voltage collapse because the parasitic capacitance of both the test fixture and the device itself is discharged through the breakdown site. The total parasitic capacitance of the test fixture in the automated tester is 83 pF, and internal device capacitances can be substantially larger. Once the crowbar circuit acts, there is some reversal of current in the device, although this automated tester keeps it to a minimum with the saturable reactor and reverse-blocking diode mentioned previously. It is possible that these additional stresses are responsible in part for the destruction of these devices.

A further attempt was made to nondestructively test these devices by constructing another tester. Reliability and ease of use concerns were set aside to make the fastest device protection possible. In this special tester, the load inductor and crowbar current shunt are eliminated and replaced with a controlled current source that can be turned off quickly. The controlled current source is driven in such a way as to simulate the load inductor in the traditional reverse-bias SOA test. The waveforms shown in figure 10 were made by using this special tester. Additional goals achieved with this special tester were reduced test-fixture capacitance and no current reversal in the device when the protection acts. The parasitic test-fixture capacitance is 62 pF, and there is no clamp provision. This special tester can test at voltage and current levels of up to about 1600 V and 25 A, respectively.

Figure 11 is a simplified schematic of the critical portion of this tester. Representative waveforms that govern the circuit's operation are also shown. A single beam-type pentode acts as a current source and replaces the inductor in the traditional test. The current source is in series with the DUT. However, the source is located on the negative voltage side of the DUT, such as the source terminal of an n-channel FET as shown in the figure. Location of the current source in the drain circuit would lead to higher parasitic capacitance at the test fixture and would complicate the driver circuits for the pentode. During the test, the gate of the DUT is turned on first. The current-source pentode is then turned on with a positive-going step function with an amplitude that determines the desired test current. This step function, which can have an amplitude of up to 1500 V at 2 A (with respect to the cathode) for the highest test currents, is applied to grid 2. Grid 1 is maintained at 0 V during this phase of the test. These conditions are maintained for about 10 μs to give the charge distribution and carriers in the DUT time to reach equilibrium. The DUT is then turned off, and the voltage across the device rises. At the same time, the step function applied to grid 2 begins a linear ramp down with a slope that can be adjusted to simulate inductors of different sizes. If the simulated inductor is set to a sufficiently small value, or the test current is sufficiently low, the DUT may simply avalanche until the current goes to zero. If the DUT experiences second breakdown, the voltage collapses, and the collapse is sensed as a fast dV/dt signal and is used to drive grid 1 of the pentode negative to turn it off. The current source is thus open-circuited, and there cannot be any current reversal in the DUT as there is with the crowbar that depends on the recovery of a diode.

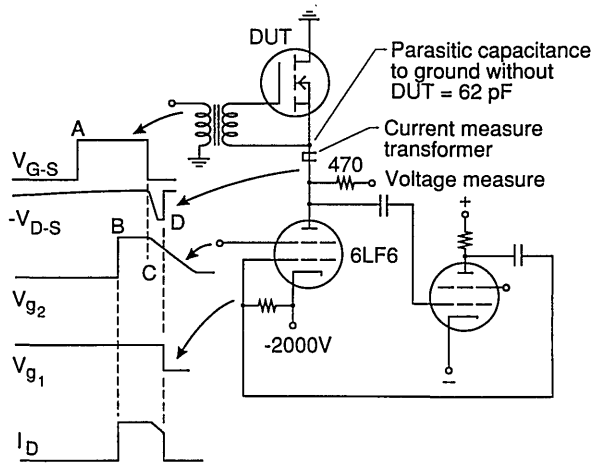

Figure 11. Simplified schematic of special reverse-bias SOA tester that was developed for extremely short device protection times. Timing waveforms for operation are indicated and are as follows:
A—Device gate is turned on
B—Constant current source is turned on
C—Device is turned off and current source ramp-down begins
D—Device breaks down, and current source is turned off fast.

The protection in this tester is extremely fast, as demonstrated by the waveforms in figure 12. Shown are voltage (top) and current waveforms for a MOSFET second breakdown. The test current was 20 A and the breakdown voltage was 550 V. The time scale was 10 ns, and the current was brought to zero within 10 ns of voltage collapse. Voltage collapse corresponds to the upturn in current just before it was brought down. The upturn is caused by discharge of the test-fixture parasitic capacitance. The voltage waveform does not give a good time reference for breakdown because of substantial time jitter from test to test on this fast time scale, and two tests are needed to capture the two waveforms. The voltage transistion was actually faster than it appears in the figure because of bandwidth limiting caused by the 470 Ω isolation resistor. The current was measured with a Pearson 411 current-to-voltage transformer, and the speed of the measurement system may be a limiting factor in determining the actual speed of the protection.

As fast as this special reverse-bias SOA tester is, it is not able to consistently save the devices that sustain long avalanches which cannot be saved by the automated tester. It is perhaps less destructive because some devices that can survive only one or two tests on the automated tester can survive three or four on this special tester before being degraded. It is possible that devices that do sustain for long periods of time before entering second breakdown actually are degraded by localized high temperatures before actual voltage collapse.

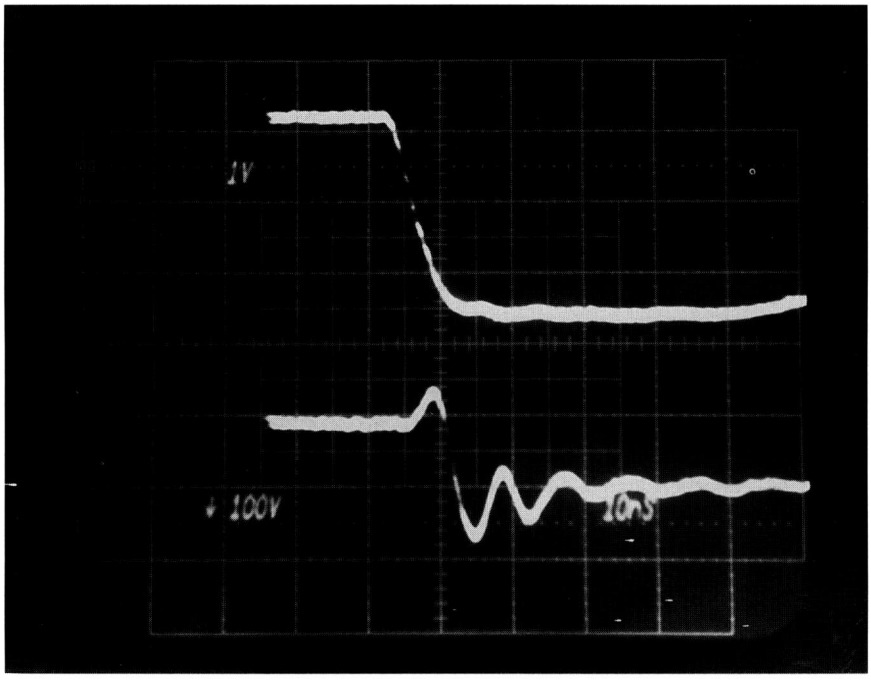

Figure 12. Voltage and current waveforms for second breakdown in a MOSFET as observed with the special tester. The current is removed in less than 10 ns after breakdown. Top trace: 100 V per small div.; bottom trace: 10 A per small div.; time: 10 ns per small div.

5. Conclusions

An overview for the design of an automated reverse-bias safe operating area tester has been given, with emphasis on special circuits that permit automation. Some measurement examples have been discussed, as well as sources of error in the measurements. The reverse-bias second breakdown measurement is generally nondestructive, provided current and voltage in the device under test are removed very quickly once second breakdown occurs. Some devices that sustain avalanches for relatively long times (microseconds) before entering second breakdown are not saved by even the fastest of protection circuits.

Acknowledgment

The author would like to acknowledge the contribution Allen Hefner of NIST made to this work by writing the software to run the automated tester.

6. References

[1] Berning, D. W., Semiconductor Measurement Technology: A Reverse-Bias Safe Operating Area Transistor Tester, NBS Special Publication 400-54, March 1979.

[2] Jahns, T. M., Investigation of Reverse-Bias Second Breakdown in Power Transistors, Massachusetts Institute of Technology MS Thesis, Dept. of Electrical Engineering, May 1974.

[3] Carpenter, G., Lee, F., and Chen, D., A 1800V, 300A Nondestructive RBSOA Tester For Bipolar Transistors, PESC '88 Record, 19th Annual Power Electronics Specialists Conference, Kyoto, Japan, April 11–14, 1988, pp. 1330–1338.

[4] Blackburn, D. L., and Berning, D. W., An Experimental Study of Reverse-Bias Second Breakdown, 1980 IEDM Tech. Digest, 1980 IEEE Intl. Electron Devices Meeting, December 1980, pp. 297–301.

[5] Blackburn, D. L., Turn-off Failure of Power MOSFETs, PESC '85 Record, Proc. 1985 IEEE Power Electronics Specialists Conference, June 1985, pp. 429-435. (Also reprinted in IEEE Trans. Power Electronics PE-2, 136–142 (April 1987).)

[6] Berning, D. W., Semiconductor Measurement Technology: A Programmable Reverse-Bias Safe Operating Area Transistor Tester, NIST Special Publication 400-87, August 1990.

About the author: *David Berning works as an Electronics Engineer in the field of semiconductor device measurements in the Semiconductor Electronics Division of the Electronics and Electrical Engineering Laboratory at NIST.*

Volume 96, Number 3, May–June 1991
Journal of Research of the National Institute of Standards and Technology

[J. Res. Natl. Inst. Stand. Technol. 96, 305 (1991)]

High Resolution Synchrotron X-Radiation Diffraction Imaging of Crystals Grown in Microgravity and Closely Related Terrestrial Crystals

Volume 96 Number 3 May–June 1991

Bruce Steiner, Ronald C. Dobbyn, David Black, Harold Burdette, Masao Kuriyama, and Richard Spal

National Institute of Standards and Technology,
Gaithersburg, MD 20899

Lodewijk van den Berg

EG&G Energy Measurements,
Goleta, CA 93116

Archibald Fripp and Richard Simchick,[1]

NASA Langley Research Center
Hampton, VA 23665

Ravindra B. Lal and Ashok Batra,

Alabama A&M University,
Huntsville, AL 35762

and

David Matthiesen and Brian Ditchek

GTE Laboratories,
Waltham, MA 02254

Irregularities in three crystals grown in space and in four terrestrial crystals grown under otherwise comparable conditions have been observed in high resolution diffraction imaging. The images provide important new clues to the nature and origins of irregularities in each crystal. For two of the materials, mercuric iodide and lead tin telluride, more than one phase (an array of non diffracting inclusions) was observed in terrestrial samples; but the formation of these multiple phases appears to have been suppressed in directly comparable crystals grown in microgravity. The terrestrial seed crystal of triglycine sulfate displayed an unexpected layered structure, which propagated during directly comparable space growth. Terrestrial Bridgman regrowth of gallium arsenide revealed a mesoscopic structure substantially different from that of the original Czochralski material. A directly comparable crystal is to be grown shortly in space.

Key words: crystal growth; diffraction; gallium arsenide; high resolution; lead tin telluride; mercuric iodide; microgravity; synchrotron; topography; triglycine sulfate.

Accepted: March 9, 1991

1. Introduction

The performance of electro-optic devices varies according to the crystal growth and fabrication procedures used; and increases in the ability to control these procedures now promise substantial improvement in such devices. In particular, reduction of convection in the microgravity found in space now offers control of one very important parameter in crystal growth. Nevertheless, the absence of comprehensive knowledge of the principal structural defects engendered during the various stages of crystal growth and device fabrication and of the roles played in device performance by the various defects has severely restricted device improvement to date.

In order to begin to shed light on the principal irregularities found in various electro-optic detector materials and their influence on device performance, we have observed and compared

[1] Lockheed Engineering and Sciences Co., Hampton, VA.

irregularities found in three crystals grown in space and in four *directly comparable* crystals grown on the ground. These irregularities, which were observed in mercuric iodide, lead tin telluride, triglycine sulfate, and gallium arsenide by high resolution synchrotron x-radiation diffraction imaging provide new clues to the nature and origins of the principal irregularities in these important materials and on their respective influence on detector performance.

Detector materials are of special interest because the performance of detectors made from space-grown mercuric iodide has been reported to be far superior to similar devices made from similar ground-grown material. The charge carrier mobility of x-and gamma-ray detectors made from space-grown crystals was at least six times higher than for similar detectors made from ground-grown crystals [1] [2]. This is expected to lead to increased energy resolution in the radiation detectors made from this material. Determination of the principal irregularities in these materials is of interest: 1) for the scientific insight to which it can lead, 2) for the optimization of expensive space growth, and ultimately 3) for the establishment of desirable growth conditions in far less extreme and expensive environments on the basis of the insight achieved.

Establishment of the specific *nature* of the mesoscopic irregularities in these materials, determination of their *level of incidence*, and observation of their *distribution* in space-grown and comparable terrestrial materials are all important to these goals. Fortunately, recent technical advances in diffraction imaging with highly parallel monochromatic synchrotron x-radiation present the first opportunity for crystal growers to obtain all of these parameters simultaneously [3, 4, 5].

2. Imaging Goals

2.1. General Considerations

X-ray topography alone has long promised to provide this information simultaneously on the *nature*, *prevalence*, and *distribution* of structural irregularities over the macroscopic areas important to integration with crystal growth parameters. However, this information clearly has not been available. A principal impediment to the fulfillment of this promise has been x-ray beam divergence. Individual irregularities and the immediately surrounding matrix in a typical crystal are illuminated by laboratory x-ray sources over an angular range measured in arc minutes. The divergence of this incoming beam (at a given point on the sample) unfortunately supports diffraction simultaneously from many irregularities along with that from the surrounding regular regions. The spread in wavelength in white beam synchrotron radiation also permits diffraction simultaneously from irregular and regular regions. In both instances, contrast that would otherwise be present is severely reduced or eliminated entirely. Even where some contrast remains, the spatial information contained is convoluted by the differing Bragg angles in a way that will not permit unfolding and subsequent detailed analysis.

Lattice deviations influencing diffraction by seconds of arc, which are frequently critical to satisfactory interpretation and understanding of mesoscopic irregularity, are rendered visible in diffraction only by a source of monochromatic radiation parallel within an arc second. In such a beam, spatial fidelity also is preserved at the micrometer level. The recent availability of such a source of x-radiation thus now permits the realization of the long term promise of x-ray topography. Among the irregularities that can now be observed *over areas large enough to interpret in terms of crystal growth* are the following.

2.2 Lattice Orientation and Strain

Of the various types of irregularity in high-quality crystals, perhaps the most pervasive is gradual change in the lattice. The orientation of the lattice or the magnitude of the lattice parameter, or both, may vary. For any one orientation of the crystal with respect to the incident beam, such variation results in diffraction only from a portion of a single grain.

Diffraction from grains whose *lattice orientation or parameter varies monotonically or aperiodically* yields images of restricted regions of a single grain. The part of such a grain that is in diffraction shifts gradually as the crystal is rotated. The moving edge of this image is characteristically soft and relatively indistinct in high resolution diffraction images.

In other systems, such as the Czochralski growth of doped material, *lattice variation* may be *oscillatory*, leading to *striations* in diffraction images of crystals cut obliquely to the local growth direction and oriented slightly off of the Bragg condition. Contrast is inverted on opposite sides of the Bragg diffraction peak. These striations record, like tree rings, not only variation in chemical composition but, taken together, also the *shape of the crystal at various stages in its growth*; and they can be deciphered in a somewhat similar if more complex and sophisticated manner [5, 6].

2.3 Grain and Subgrain Boundaries

Sharp contrast in the image of a crystal can delineate homogeneous grains or subgrains. In contrast to the preceding case, the boundaries of such an image do not move as the crystal is rotated in the Bragg direction. Where the lattice orientation of a pair of such homogeneous grains differs by rotation in the diffraction plane by more than the acceptance angle, only one of these grains (or subgrains) will diffract at a time; and, if it is not strained, it does so in its entirety. Variations in real-time images of such a crystal permit rapid and detailed assessment of the *relative misorientation* (in the Bragg direction) *of the various grains and subgrains* with respect to one another.

Where the lattice orientation of a pair of such contiguous grains differs in a direction orthogonal to the plane of diffraction, both grains may appear in diffraction simultaneously, but the resulting images are displaced with respect to one another in this direction. The pair of these images is either separated or overlapped, depending on the relative inclination of the two lattices.

2.4 Dislocations

Dislocations typically appear in diffraction images taken in Laue geometry (transmission) as linear features that are broader at one end than the other. The broadening of one end of such a feature arises from scattering deep within the crystal, while the sharp end locates the intersection of the dislocation with the x-ray exit surface of the crystal. The *orientation of a dislocation* can be determined with high precision for those cases in which the intersection of the dislocation with both entrance and exit faces is distinct in the diffraction image.

Variation in the visibility of such a line feature in successive diffraction images indicates the *direction of atomic displacement associated with a dislocation*, which is parallel to its Burgers vector. However, since the visibility of such a dislocation varies relatively slowly, that is, as the cosine of the angle between the Burgers vector and the diffraction vector, the determination of this direction is most precise when contrast can be observed to disappear at one unique angle. When the direction of diffraction is oriented normal to the atomic displacement, such a feature vanishes from the diffraction image.

2.5 Phase Domain Boundaries

Twins are distinguished normally by absence of diffraction from regions between sharp parallel boundaries visible in some diffraction directions but not in others. The contrast in the latter when observed by high-resolution beams may be affected by very slight differences in lattice alignment.

With angular collimation of the order of an arc second, the images of other boundaries recorded in Laue geometry may also become visible when the diffraction vector falls along the boundary. Such boundaries are visible even under these restrictive conditions only when they separate atomically coherent regions differing by an atomic phase shift [7, 8]. Those boundaries that have been observed to date to fulfill these conditions appear to separate *antiphase domains*. Radiation with a divergence of the order of a second of arc or less is necessary to image such boundaries.

2.6 Additional Phases

The absence of diffraction from particular regions of a crystal under all diffraction conditions supporting diffraction from the rest of the crystal strongly suggests the presence of a *second phase*, although in principle the non diffracting regions may simply be misoriented with respect to the rest of the crystal. In stoichiometric materials, the boundaries of two phases are sharply delineated. In alloys, this sharpness is vitiated by the gradual changes in composition that may be permitted.

2.7 Surface Scratches

The strain associated with surface scratches is linear, but sometimes gently curved, and typically non crystallographic in orientation. They are typically distinguished also by three other characteristics: 1) uniform width, 2) sharp edges, and 3) contrast reversal either laterally, longitudinally, or both, particularly as the Bragg peak is scanned. The latter two characteristics are evident both in observation in Bragg geometry and when scratches are present on the exit surface in Laue geometry.

3. Current Imaging Capability

Suitable synchrotron radiation sources now offer opportunities to fulfill the long awaited promise of x-ray topography; but the degree of success in their realization depends upon the particular parameters of the storage ring and its beam lines. Since the precise orientation of the x radiation at individual points on the sample is crucial to the analysis, the

vertical source size, together with the distance of the sample from the tangent point on the ring, may limit the utility of the images produced.

The x-ray storage ring of the National Synchrotron Light Source at Brookhaven National Laboratory offers the most suitable combination of characteristics of any existing storage ring, providing an unusually bright beam whose degree of vertical divergence at a point on a sample mounted on beam line X-23A3 is 1.5 arc seconds.

Although this 1.5 arc second beam provides a considerable improvement over other sources, it is not yet sufficient for diffraction imaging with optimum sensitivity to defects. For useful sensitivity to irregularities, which requires further improvement in beam divergence by another order of magnitude, i.e., 0.1 arc second, the optics of the monochromator are crucial. Such a beam is necessary for rendering critical features visible, for preservation of the spatial information in the image within the plane of diffraction, and for displaying essential clues to the strains upon which the success or failure of detailed analysis can depend [6].

With such a dedicated 0.1 arc second monochromatic capability, however, which is available on a routine basis only on Beam Line X-23A3 at the National Synchrotron Light Source, the detection and interpretation of irregularities are limited principally by their density. Irregularities can be recorded photographically with a spatial resolution of 1 μm. Observation with an x-ray vidicon and charge coupled device (ccd) cameras readily provides complementary information with a spatial resolution of 35 μm with intermediate sensitivity in real-time, and 20 μm with shot noise limited sensitivity in quasi real-time.

4. The Crystals

Three crystals of mercuric iodide, two of lead tin telluride, one of triglycine sulfate, and one of gallium arsenide are included in the present study.

The three mercuric iodide crystals were grown by *identical* physical vapor transport *procedures*, one on Spacelab III, a second from *identical source material* under full gravity, and a third also terrestrially from more highly purified material recently available. The resulting crystals were state of the art for each material, as demonstrated by the performance of detectors made from directly comparable material and by the diffraction images.

The two lead tin telluride crystals were grown *by identical Bridgman techniques from identical source material*, one on Space Shuttle STS 61A and the second terrestrially. The images strongly suggest that this material also is state of the art for such a ternary crystal.

The triglycine sulfate crystal consisted of a terrestrial seed and additional growth *by identical techniques from identical solution* on Spacelab III. The images indicate that this material contains relatively few irregularities.

The gallium arsenide crystal consisted of a Czochralski seed and Bridgman regrowth, all terrestrial, but carried out under procedures identical to those due to be employed shortly in a space experiment. The purpose of this particular regrowth experiment is to examine various aspects of the growth rather than to grow immediately the most regular material. The terrestrial material is of interest in its own right for diffraction imaging just now in addition because it is the first Bridgman material to be observed by high resolution diffraction imaging.

5. The Images

5.1 Mercuric Iodide

5.1.1 Terrestrial Crystal Compared with Spacelab III Crystal While the terrestrial mercuric iodide crystal grown from source material identical to that used for a crystal grown on Spacelab III diffracts over a range of one half degree, a large central portion is sufficiently regular to diffract only within a few minutes of arc. Full high-resolution diffraction images of this terrestrial crystal appear in figures 1 and 2, and an enlarged portion of the first in figure 3.

Most of the central portion of the crystal is in diffraction in the (1 1 12) image in figure 1, indicating lattice regularity with respect to rotation around a [1$\bar{1}$0] axis of the order of a few arc seconds. However, the absence of diffraction in a wide [110] (vertical) stripe in the center of this figure indicates that the lattice is deformed by a sharp twist of about 10 minutes of arc around the orthogonal axis defined by this stripe. The extent of this twist is determined from the 100 μm width of the stripe, and the knowledge that the photographic plate was located about 3.5 cm from the crystal. This twist of the crystal lattice is evident also in the (0 1 11) diffraction image, figure 2, for which the crystal was rotated azimuthally 45°. In this orientation, the misalignment of the two parts of the crystal precludes bringing them simultaneously into diffraction. Examination of a number of full images

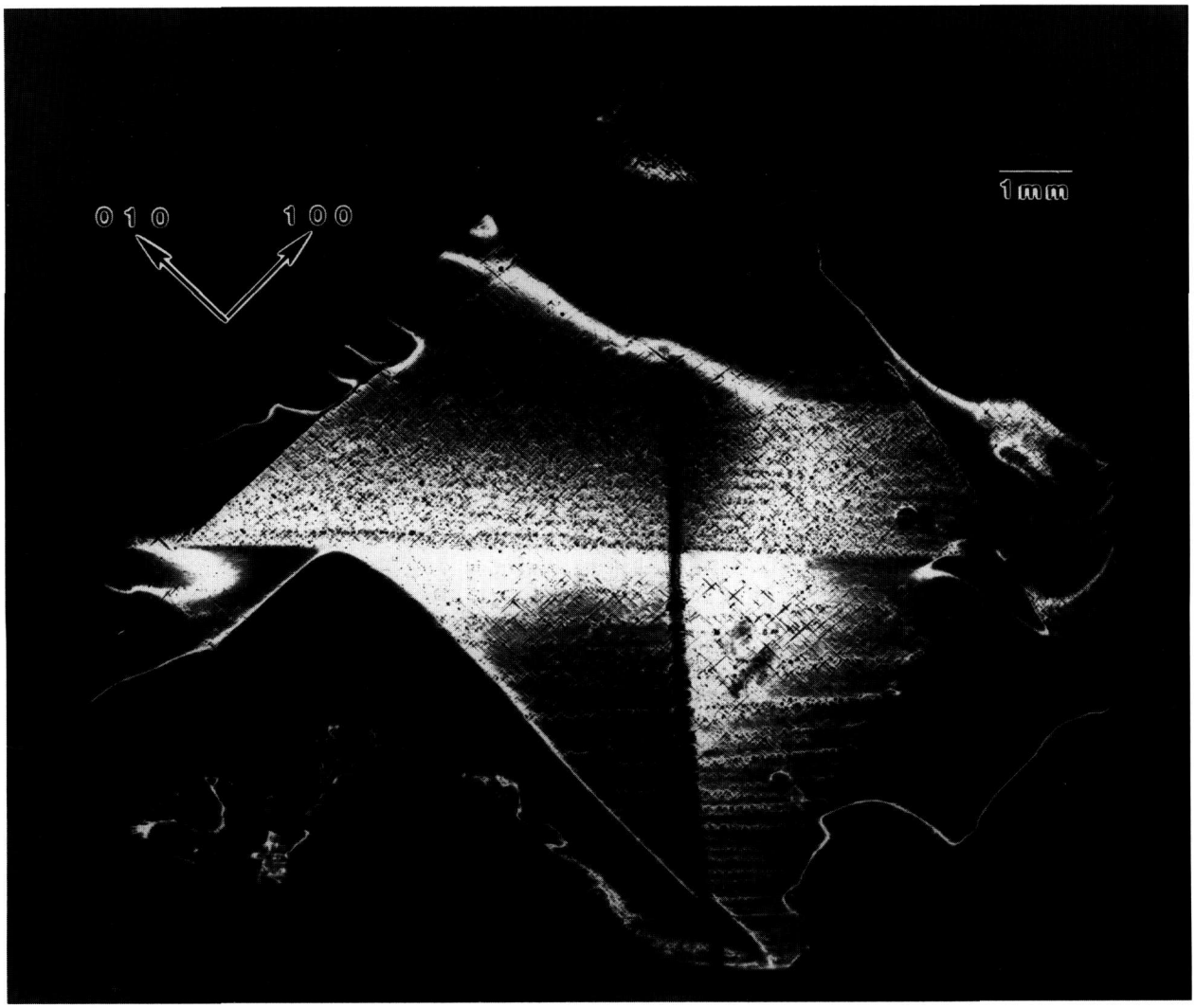

Figure 1. High-resolution (1 1 12) 8 keV diffraction image in Bragg geometry of the (001) surface of terrestrial HgI_2 crystal comparable to that grown on Spacelab III. Lighter areas diffract more strongly.

and of a sequence of real-time images indicates that the principal lattice twist axis itself bends through several minutes of arc, differing slightly in the two subgrains.

The other principal large feature of the full images of this crystal is a set of textural stripes, which are oriented in the $1\bar{1}0$ direction. Enlargements such as that in figure 3 show these stripes to consist of a relatively high density of discrete features, typically out of diffraction in these images and therefore ascribable to one or more additional phases. Some of these features take the form of thin {100}-oriented stripes a few micrometers wide; they are sometimes crossed. The others are more irregular, globular features, 1–60 μm in diameter. These may differ completely from the stripes, but may simply represent similar stripes normal to the (001) image and projected on it. In those regions characterized by a high density of discrete features, diffraction appear to be restricted to small (≈5 μm) cells of the type observed in scanning cathodoluminescence microscopy [9].

The other areas of the crystal contain similar features that are out of diffraction, but with a much lower density. In addition, however, these regions contain thin, curved features marked by varying sections of higher diffraction, lower diffraction, or alternating regions of higher and lower diffraction in in tandem. The inability to observe diffraction in Laue geometry in the current series of experiments, because of the sample thickness, precludes firm identification now of these features as dislocations.

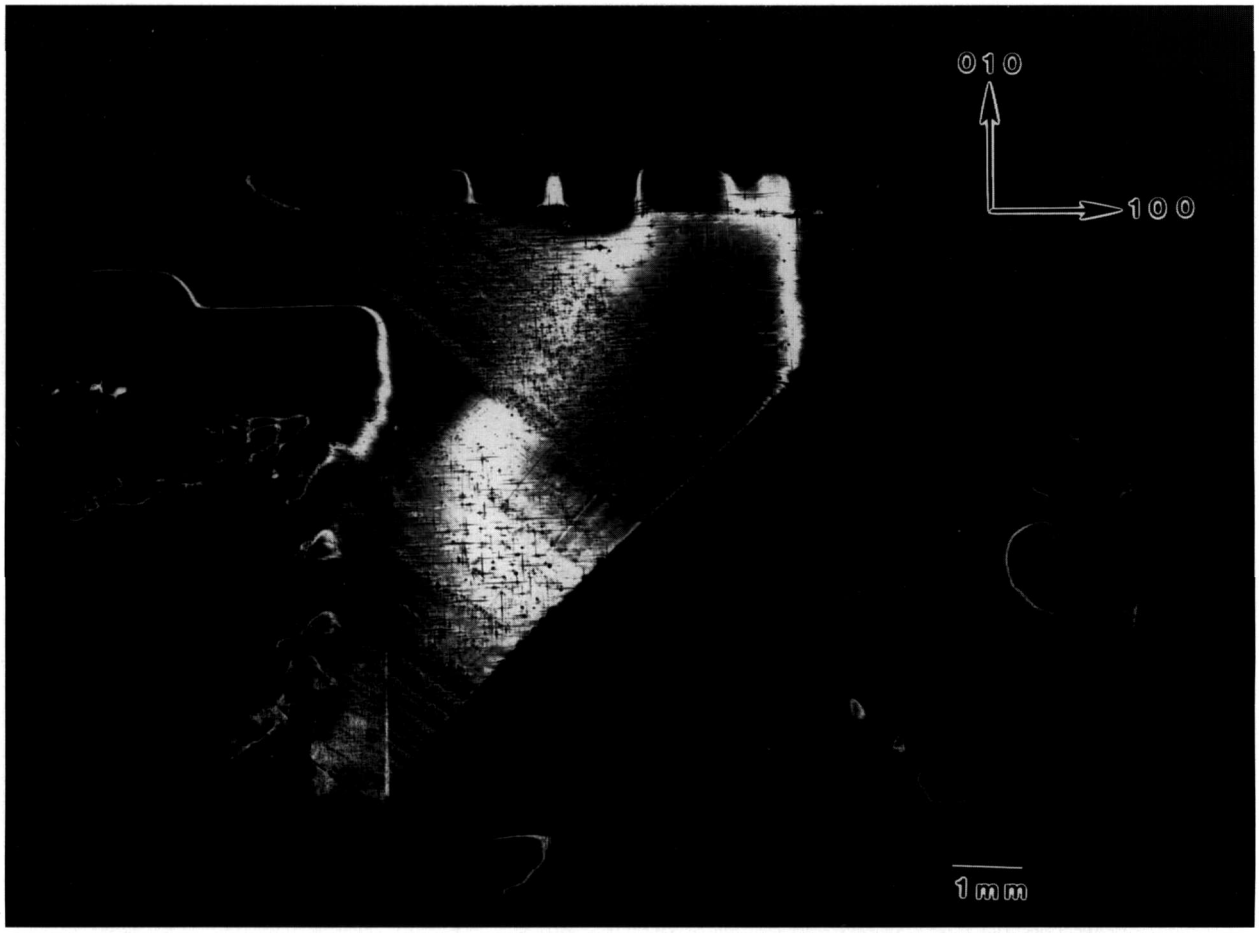

Figure 2. High-resolution (0 1 11) 8 keV diffraction image in Bragg geometry of the (001) surface of terrestrial HgI_2 crystal comparable to that grown on Spacelab III. Lighter areas diffract more strongly.

The nature and arrangement of these various features in this sample make this crystal into a Rosetta stone in understanding the evolution of irregularity in mercuric iodide. One principal question that arises is associated with the origin of the lattice twist. Does it appear during growth or only later during subsequently handling of this very soft crystal? Six distinct observations all indicate that this lattice twist occurred during growth and indeed indicate the growth direction. The first two observations are that the twist axis does not extend across the entire crystal, and that once started, the magnitude of the apparent separation of the two parts of the images does not increase. It is difficult to conceive of such a partial lattice twist, one lying precisely in the (001) plane, developing through inadvertent mishandling. The third observation is that this twist axis is normal to the $[1\bar{1}0]$ layered texture formed by a high density of additional phase features. These layers appear to be broad striations formed during growth and to indicate its direction; the [110] lattice twist axis appears to be aligned with the crystal growth direction. Fourth, the gradually curved nature of some of the linear multiple phase configurations in the vicinity of the lattice rotation is more consistent with growth than with post-growth bending. Fifth, the onset of the lattice twist immediately precedes a major textural change that appears to be growth-related. The final observation is the bending that has been noted in the lattice twist axis, bending that differs in the two resulting subgrains.

Examination of the interfaces between the widest stratum of high-density features and the adjacent low feature density strata confirms the growth orientation and the origin of the lattice twist. The linear additional phase features in the low-density layer immediately adjacent to the high-density region near the center of the crystal appear correlated closely with individual features in the high-density region. Growth thus took place in this part of the crystal from the high-density stratum to

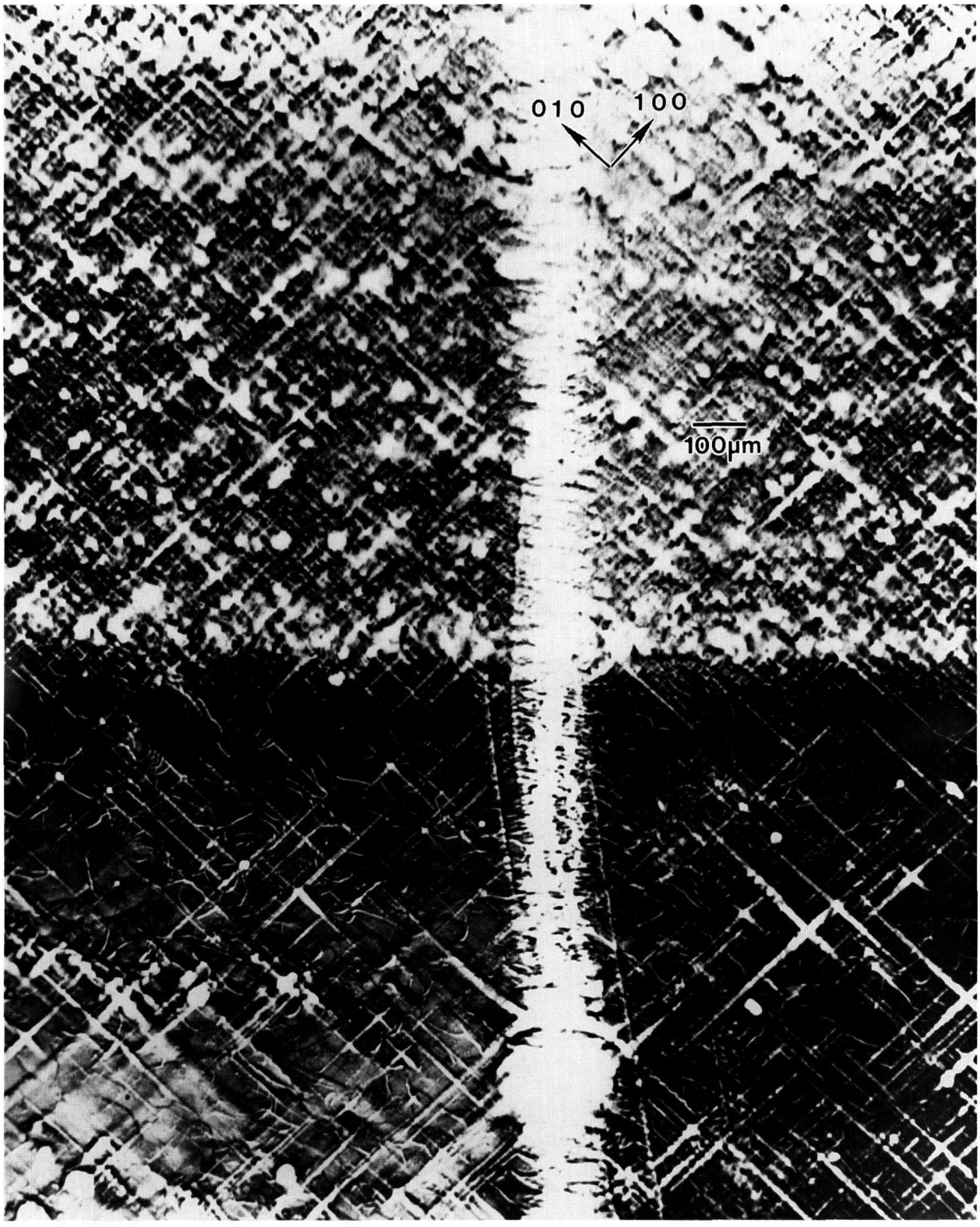

Figure 3. Enlargement of central portion of figure 1, (1 1 12) diffraction. Darker areas diffract more strongly.

the low-density stratum, i.e., in a direction projecting onto the (001) crystal surface in the [$\bar{1}10$] direction. Moreover, as just noted, the sharp lattice twist appears to begin immediately preceding the onset of the broad textural stratum of high density of (precipitate) features, where it joins the preceding low feature density textural stratum.

All of these observations are consistent with a growth model in which growth begins in the extreme [110] corner of this crystal (in fig. 1 this is the top corner) and proceeds relatively uneventfully in the [$\bar{1}10$] direction (downward in fig. 1), or in a direction projected onto this direction in the images, until just before the onset of the wide swath of a high density of additional phase features, one of which initiated the sharp lattice twist. The twist then propagated for the remainder of the growth. During this subsequent growth, briefer periods of relatively high additional phase density alternate with periods of relatively low additional phase density.

The nature of the additional phase material is suggested by evaporation of such crystals. As material is removed, small specks of foreign material similar in size to the additional phase features observed in this study accumulate on the surface, at an irregular rate. Chemical analysis indicates these specks are neither mercury nor iodine precipitates but rather consist of organic and metallic impurities with a 70% carbon content and a wide variety of metals. It is tempting to associate these observed impurity formations with the additional phase features observed in diffraction and therefore to conclude that these impurities reside in such crystals in discrete form.

The morphology of the diffraction images permits us to develop two alternative growth models, which tie together all of these observations. Growth over a region of a few micrometers forms a crystal with a relatively high degree of purity and crystal perfection, creating small regions that diffract strongly. Impurities are rejected from the crystal during this stage of the growth process, in a manner similar to constitutional supercooling, and accumulate near the growth surface.

In one model, the level of impurities after growth of a few micrometers accumulates to such an extent that they precipitate out, marking the local growth surface in {100} directions. At reentrant corners of such {100} growth surfaces a globular precipitate possibly forms. In a second model, the rejection of impurity stimulates dendritic growth, which leaves the linear features observed in {100} directions. In this case, the globular form of the precipitate may form in the reentrant dendritic locations. Alternatively, the features that appear to be globular may simply represent the cross section of dendrites normal to the image surface. None of our observations to date permit us to distinguish absolutely between these two models.

Either model involves modulation of the general impurity level by an as yet unidentified process that produces textural stripes or striations delineated by changes in the density of precipitates. The resulting composite formed in either model resists deformation. It consists of relatively pure and thus relatively strain-free components.

5.1.2 Spacelab III Crystal A crystal grown in Spacelab III from material identical to that used for the terrestrial growth of the crystal shown in the preceding section diffracts over a wider angular range, about one and one half degrees. A full high-resolution diffraction image appears in figure 4 and an enlarged region of this in figure 5. It is clear from the appearance of the full images as well as from the one and one half degree acceptance angle for diffraction that the lattice orientation or parameter of the space crystal in its entirety is less uniform than the comparable terrestrial crystal shown in figures 1-3: that is, less of the space crystal appears in diffraction at a given angle of incidence than does the comparable terrestrial crystal, indicating gradual variation either in lattice parameter or lattice orientation, or both.

Perhaps closely related, but potentially far more important, is substantial reduction in the enlargements of the images of the Spacelab III crystal of arrays of features that are out of diffraction, the textural arrays characteristic of the comparable terrestrial crystal. A few irregular regions of the order of 50 μm across that are out of diffraction are observed, but they are much less pervasive and sharply delineated than are those in the terrestrial crystal. None of the crystallographically oriented regular regions that are typically out of diffraction in the images of the comparable terrestrial crystal are observed in the Spacelab III crystal. The formation of regions of additional phase thus appears to be almost completely suppressed in the crystal grown in microgravity.

The Spacelab III sample differed from the terrestrial sample not only by its growth in microgravity but also by the superposition of graphite electrodes; so that its performance as a neutron and x-ray detector could be measured. Since graphite is relatively transparent to x rays, these electrodes were not expected to interfere with the imaging process itself. While they could in principle have affected the surface strain, we found no evidence for this.

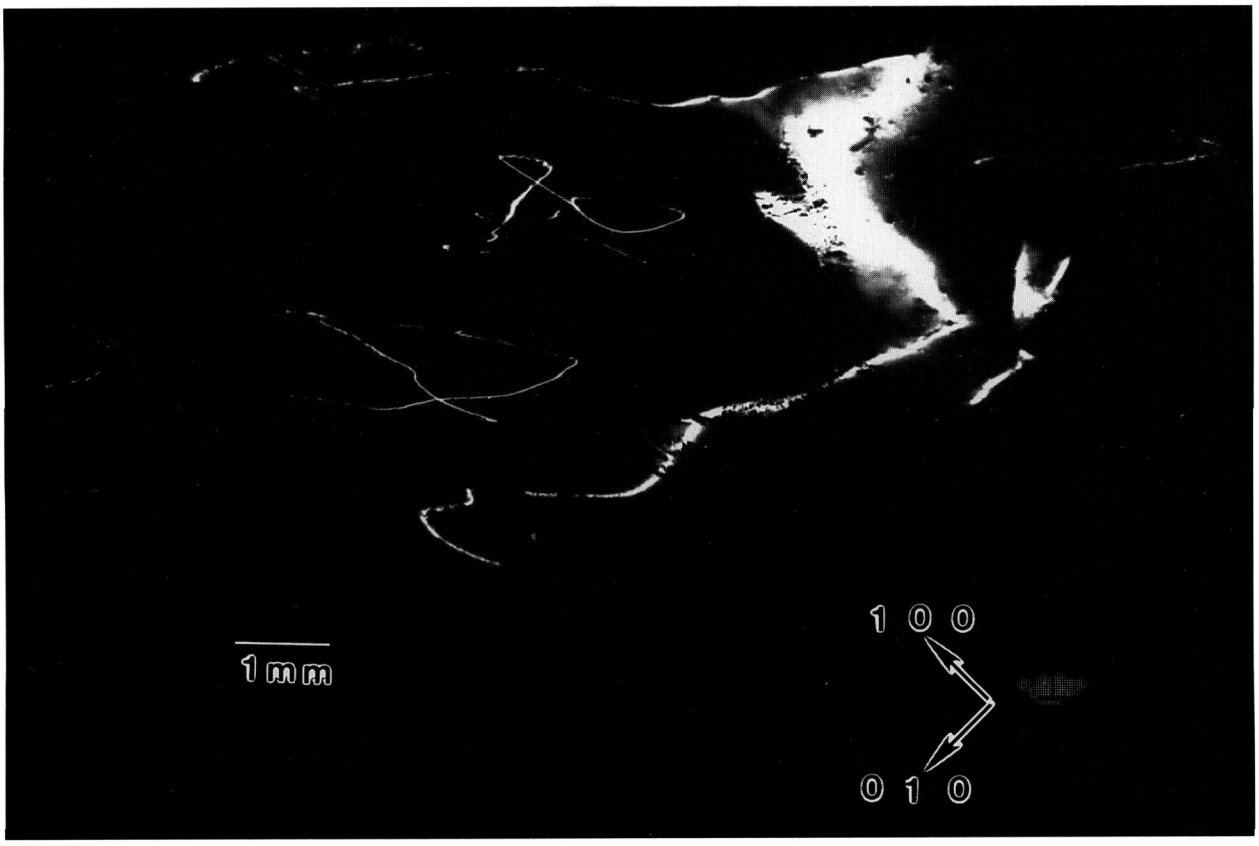

Figure 4. High-resolution (1 1 10) 8 keV diffraction image in Bragg geometry of (001) surface of Spacelab III HgI_2 crystal. Lighter areas diffract more strongly.

However, this crystal was not encapsulated. With the passage of the 5 years since the growth of this crystal, some deterioration in electronic performance of the device made from it actually has been observed, as is characteristic also of unencapsulated devices made in terrestrial environments.

The observed gradual variation in lattice is consistent with varying retention within the lattice of some foreign material. This leads to increased interest in the results of the growth of a mercuric iodide from the much purer material on a future flight.

5.1.3 Terrestrial Crystal to Be Compared to a Future Flight Crystal Models ascribing a central role in the structure and properties of mercuric iodide detectors to impurities are reinforced by observation of a third mercuric iodide crystal, grown terrestrially in an identical manner from higher purity material similar to that to be used on a future flight. It diffracts over a full two degrees. A full high-resolution diffraction image appears in figure 6, with an enlarged region in figure 7. The extent and character of the diffraction in these images, reflecting the general lattice uniformity, resembles much more the diffraction from the Spacelab III crystal than that from its terrestrially-grown counterpart. Moreover, the absence of an array of small features that are out of diffraction also gives these images much more the character of those from the Spacelab III crystal than those from the terrestrial crystal grown about the same time from similar material.

The performance of devices made from the new high purity material approaches the original performance of the device made from the Spacelab III crystal. The improved performance of the Spacelab III crystal is traceable to the higher mobility of its charge carriers. By contrast, however, the purified terrestrial crystal here is characterized by improved carrier lifetime. Although the electronic improvements are quite distinct in these two cases, in neither crystal do we find the additional phase features that we have observed in the first terrestrial crystal. Thus absence of additional phase precipitates appears to be much more important to device performance than the generally higher level of lattice uniformity that we observe in the first terrestrial crystal. The stiffening provided by additional phase precipitates apparently comes at too high a price in terms of charge carrier trapping.

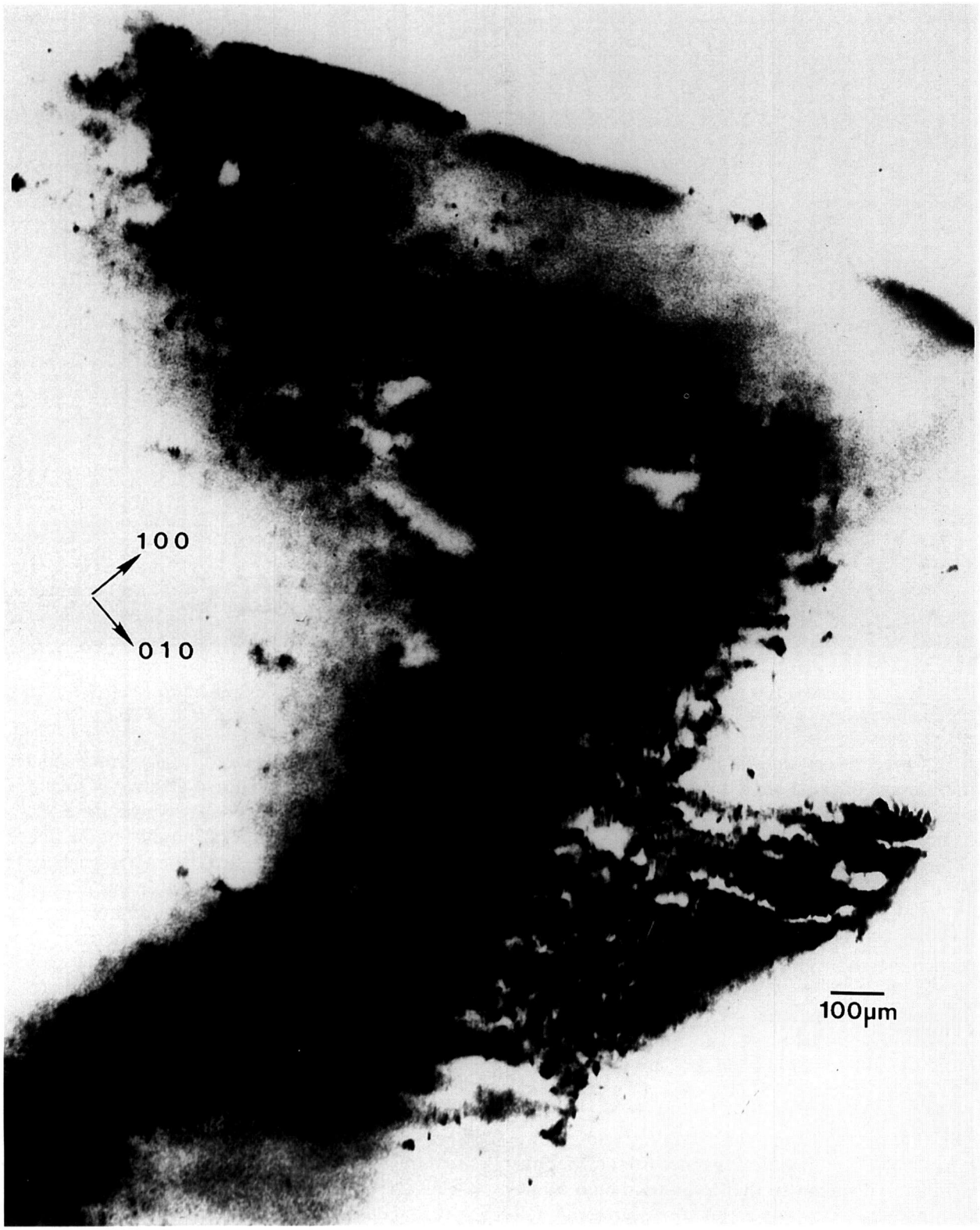

Figure 5. Enlargement of main region of figure 4, (1 1 10) diffraction. Darker areas diffract more strongly.

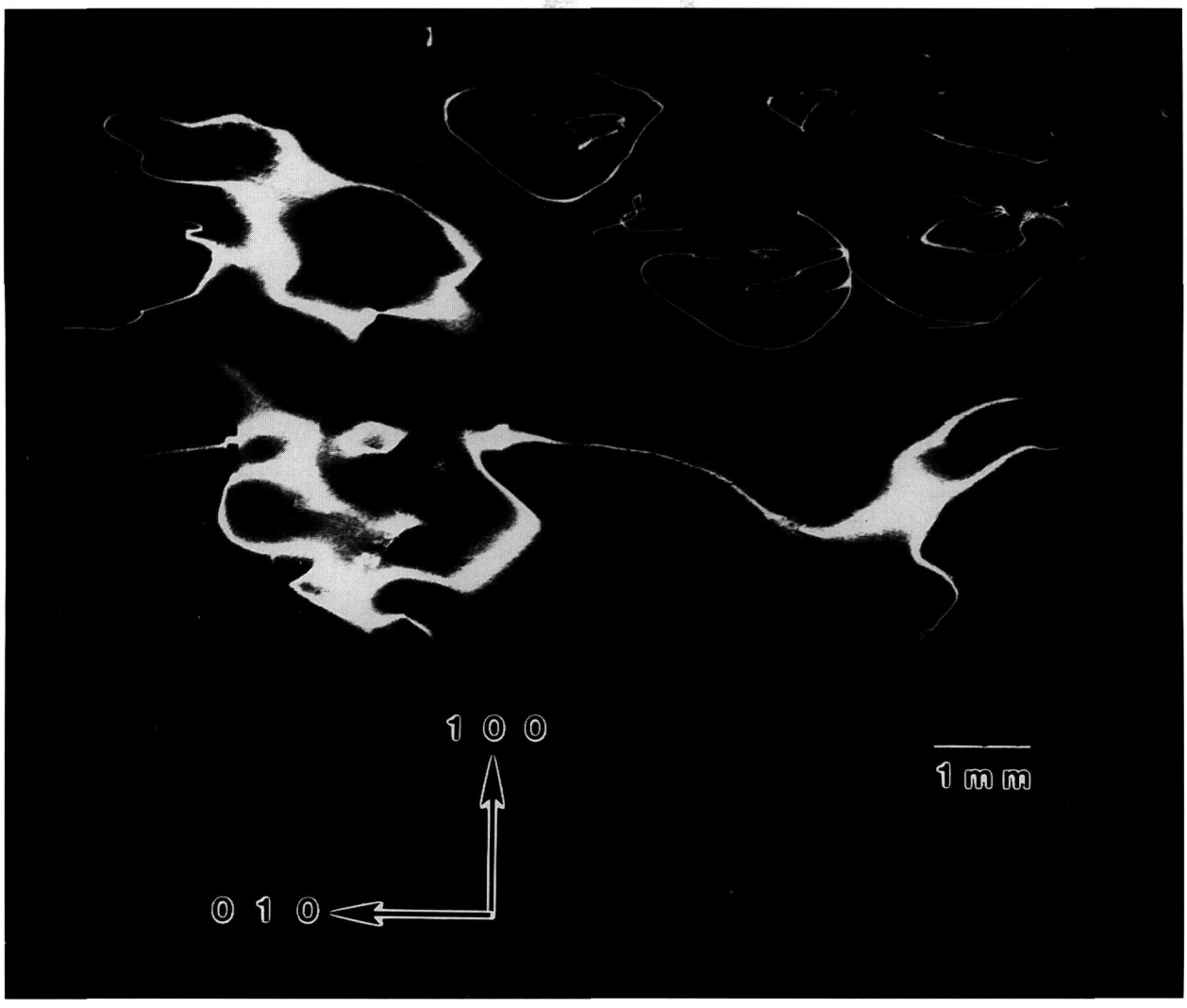

Figure 6. High-resolution (1 0 10) 8 keV diffraction image in Bragg geometry of the (001) surface of terrestrial HgI_2 crystal, comparable to crystal to be grown on future flight. Lighter areas diffract more strongly.

Future space growth of this high purity material now assumes particular interest. Will incorporation of residual impurities in the final crystal even below their currently low level in the charge material be achieved in space growth? And, if so, will these lower impurity levels lead to greater general lattice uniformity? And finally, will this new level of regularity result in still further improvement in device performance, improvements both in carrier mobility and in carrier lifetime?

5.2 Lead Tin Telluride

5.2.1 Terrestrial Crystal Comparable to Space Shuttle STS 61A Crystal Various regions of the terrestrially grown sample of lead tin telluride similar to one grown on Space Shuttle STS 61A diffract as the crystal is rocked over a full two degrees. Full high-resolution diffraction images of two distinct grains are shown in figures 8 and 9. Growth was in the [001] direction, which is oriented "down" in all figures.

The sample was a regular half cylinder. The sharply delineated irregular outlines of the image in figure 8 thus indicate immediately that several grains are present: the curvature of the [1$\bar{1}$0] (right hand) edge indicates that a subsidiary grain started to grow almost simultaneously with the main grain. Then, after 1 cm of growth, a third grain started to grow between the center of the boule and the opposite edge of the main grain. It grew laterally more rapidly than the nucleating grain, however, displacing and, after another 2 cm, completely overtaking the growth of the nucleating grain. The new grain is brought into diffraction in figure 9 simply by rotation of the sample about the boule (growth) axis.

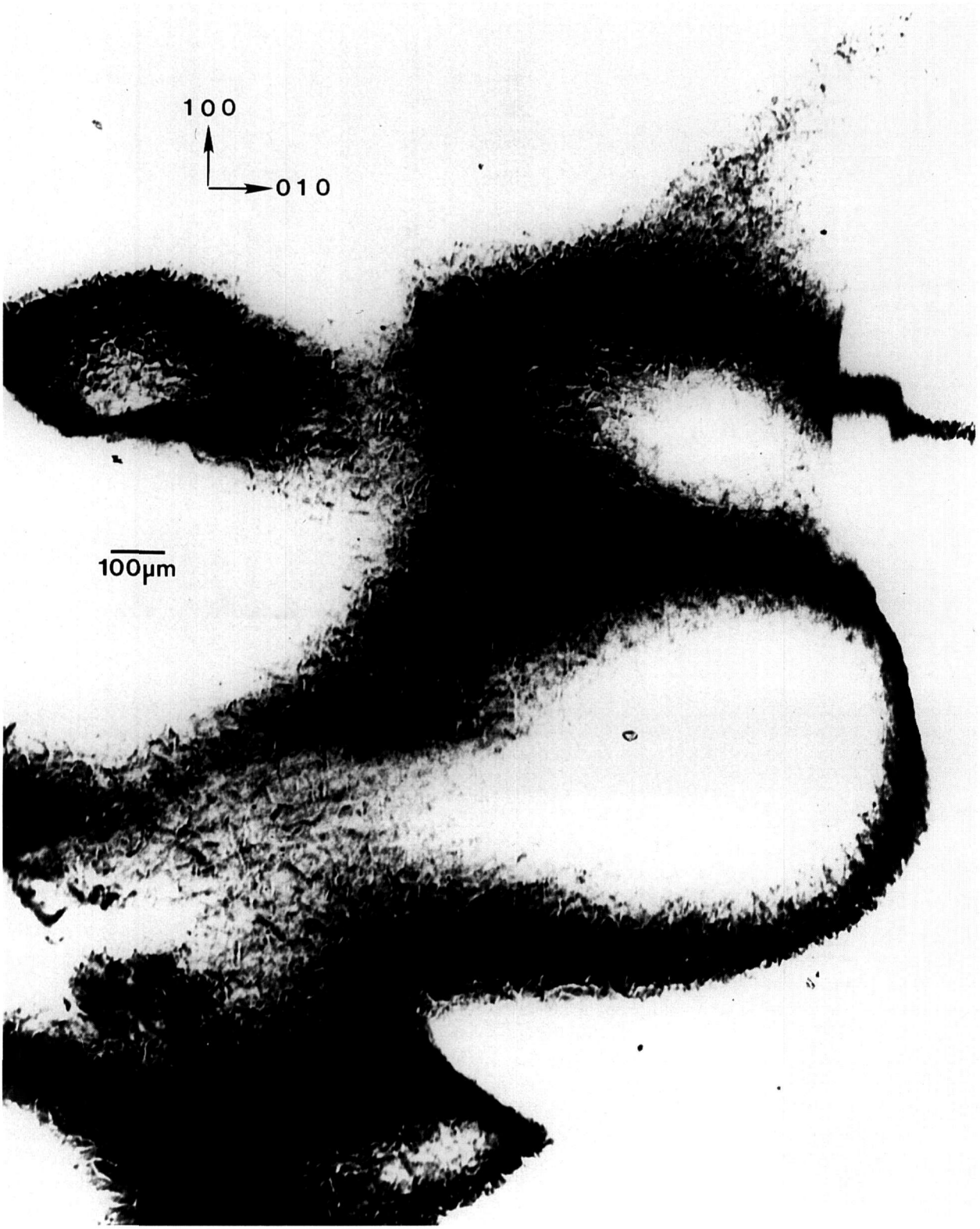

Figure 7. Enlargement of lower left portion on figure 6, (1 0 10) diffraction. Darker areas diffract more strongly.

Figure 8. High-resolution (220) 8 keV diffraction image in Bragg geometry of the approximately (220) surface of terrestrial PbSnTe crystal. The growth direction is [001]. Lighter areas diffract more strongly.

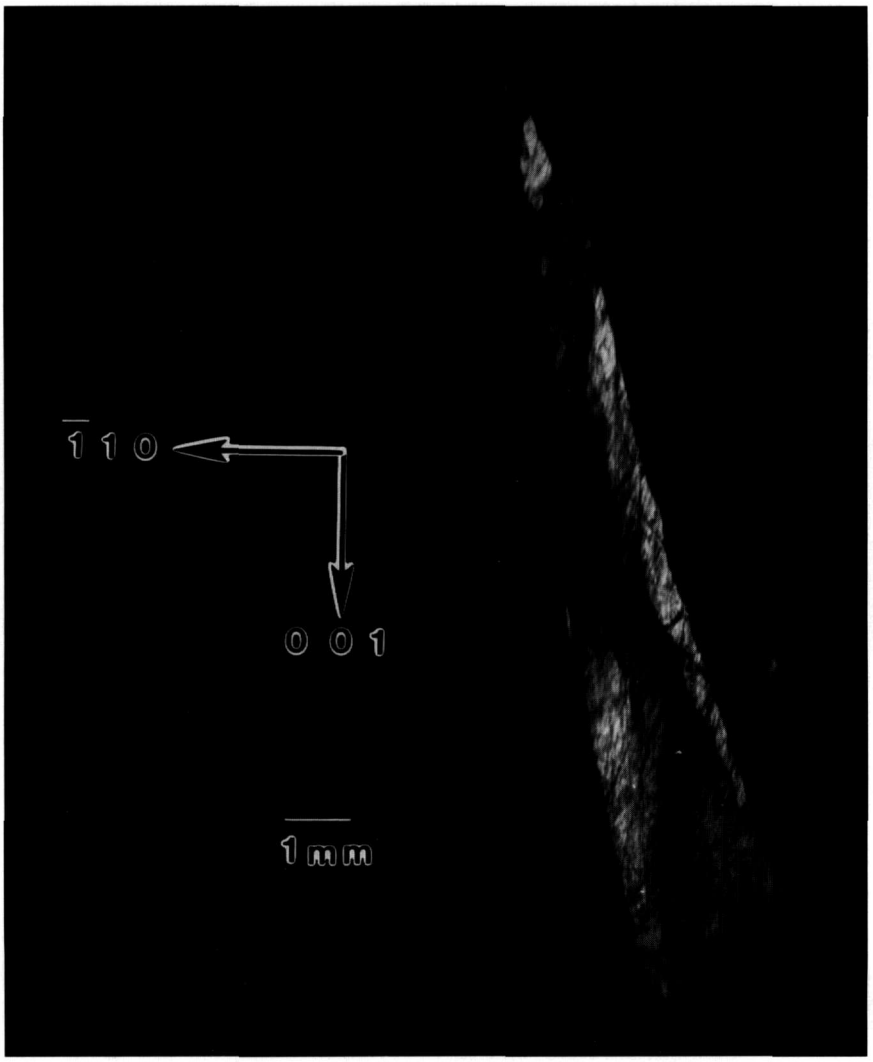

Figure 9. High-resolution (220) 8 keV diffraction image in Bragg geometry of the approximately (220) surface of terrestrial PbSnTe crystal. The angle of incidence is 33 arc minutes larger than that for figure 8; the growth direction is [001]. Lighter areas diffract more strongly.

Subgrains within each of the main grains are clearly visible through terraced variation in contrast and can be studied in real time images as the crystal is rotated. The generally strong diffraction from a 1.5 cm length of each of the two principal grains observed is notable, however, in light of the increase in tin level from 14 to 18% during the first 3 cm of growth visible in these images . The fractional change in lattice constant over the 1.5 cm length of the grains is 4×10^{-4}, which changes the Bragg angle by 90 arc seconds. Nevertheless, diffraction is observed in a single image of one of the grains through broadening by kinematic scattering, which is difficult to quantify, as well as by local compositional variation. Because of the mixture of these two broadening mechanisms, unfortunately we can not use the broadening to evaluate the degree of local compositional variation.

Other aspects of this variation are evident in enlargements such as figure 10. Cellular regions of high diffraction varying in size from ten to several hundred μm are observed. They are separated by lines of reduced diffraction that are 10-50 μm wide.

Many of these lines at first glance appear to be scratches because of their curvature and random orientation. However, three characteristics typical of surface scratches, such as those visible for example in the gallium arsenide images to which we turn

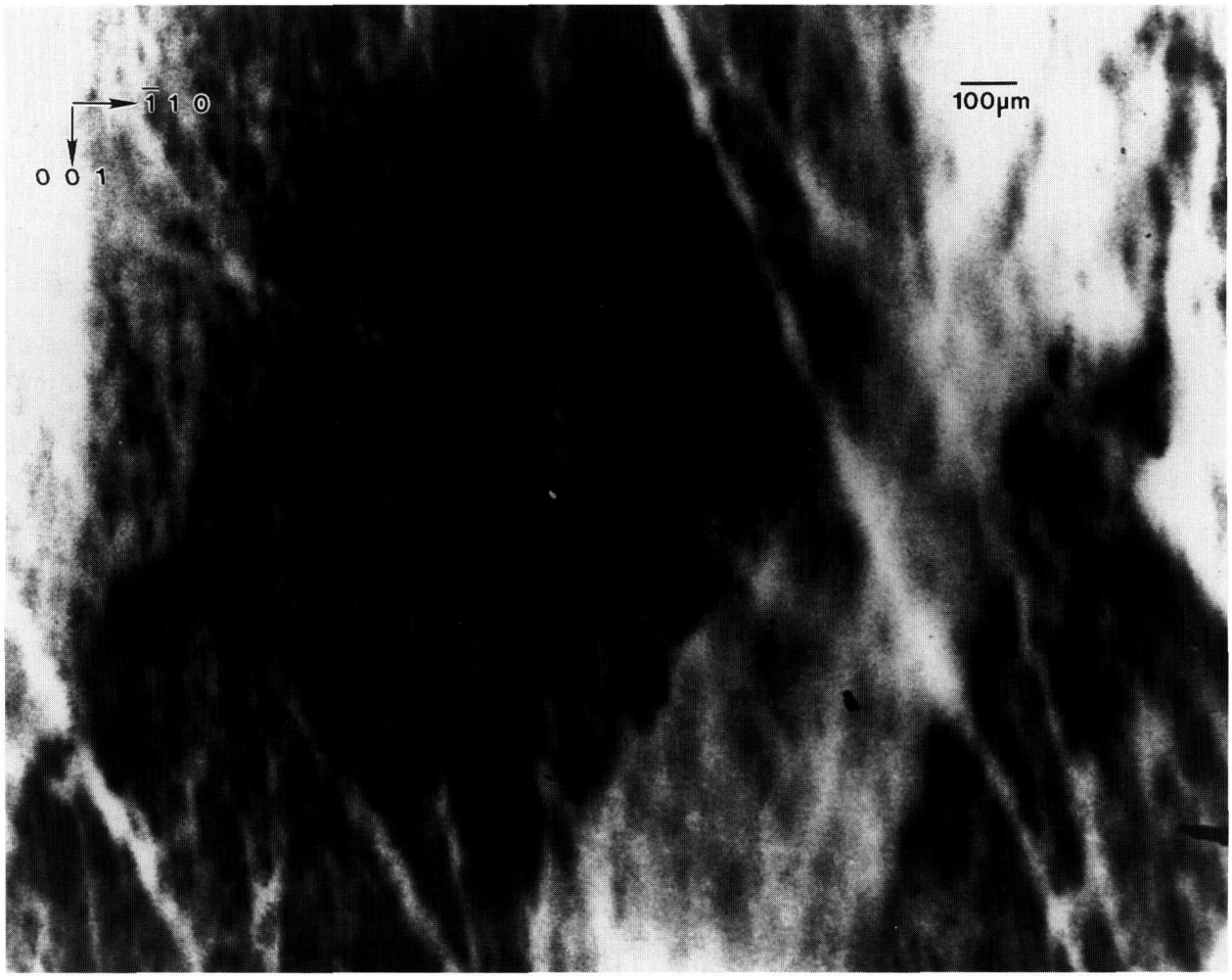

Figure 10. Enlargement of central portion of figure 8, (220) diffraction. Darker areas diffract more strongly.

later, are not observed in these linear features. First, the lines vary in width, both from line to line, and even over the length of a given line. In reality these lines separate cellular regions of high diffraction. Second, the boundaries of the lines are very indistinct. And third, contrast reversal is never observed in them. They are invariably out of diffraction over their entire length, even as the crystal is rotated while it is observed in real time. Thus, while we cannot rule out scratches, the linear features here differ markedly in several respects from those of typical scratches in other materials. Moreover, they are not observed in the image of the space-grown sample, whose images follow.

We are thus left with the postulate that the highly diffracting cells are separated by material of another phase. The indistinctness of the boundaries between these regions of differing phase strongly suggest gradual change in chemical composition on a scale of 1-10 μm or so, in contrast to the sharp delineation between diffracting and non diffracting features in the images of mercuric iodide discussed above.

The pseudobinary phase diagram along the lead-tin axis predicts complete miscibility [11]. However, the observation of similar structure following electrolytic etching led earlier to a series of experiments on the metal/tellurium ratio, which delineated its importance in the growth of this material. This earlier work provides a satisfactory model for the current observations as well [12, 13]. While the metal constituents are widely recognized to be interchangeable, a single phase is preserved only with tellurium concentration in excess of 51%. Below this value, two phases are formed, differing in metal/tellurium ratio. Since the tellurium concentration of the current crystals is 50.1%, two phases are actually to be expected. Constitutional supercooling may also play an important role, depending on the temperature gradients imposed [14].

5.2.2 Space Shuttle STS 61A Crystal A full image of a crystal grown on flight STS 61A appears in figure 11 and an enlargement of the central portion of this in figure 12. The multigrain nature of the STS 61A crystal is generally similar to that of the terrestrial crystal. But, while these images appear qualitatively similar to the full images for the corresponding terrestrial crystal, they differ in important ways.

Each grain is more generally uniform than those of the terrestrial crystal. This uniformity follows a drastic reduction in the incidence of linear features and subgrains. As a result, variation in diffraction on a scale of 10–100 μm is greatly reduced. Thus, while the granular structure resembles that for the terrestrial crystal, variation within individual grains from the intrusion of a distinct second phase appears to be suppressed in microgravity. The absence of thermo-solutal instability for this system in microgravity was noted in the preceding section.

5.3 Triglycine Sulfate

A normal slice from a disc-shaped terrestrial seed crystal with additional growth achieved on Spacelab III diffracts into images, each of which appears over less than half of an arc minute. Full high-resolution diffraction images appear in figures 13-16. The character of the diffraction from this crystal is very different from that of the others. This crystal was thin enough and low enough in atomic number to allow diffraction in Laue geometry. Moreover, superimposed images of this crystal taken as it was rotated about its [100] and [001] axes appear in closely spaced groups, each associated with the diffraction directions expected for diffraction from one set of (h00) or (00l) planes, respectively. The various images have similar, but not identical, shapes. Subsequent work, summarized in table 1, indicates that the members of a given group of images appearing at nearly similar diffraction angles come into diffraction at differing sample orientation.

Table 1. Layers on sample of triglycine sulfate

Layer	Figure No.	(hkl)	Orientation (°)
C2	13	300	+6.6
B3		200	+2.1
B1	14	200	+2.5
A2	15	001	−19.2
A3	16	001	−36.8

The appearance of images in groups indicates that this crystal consists of layered grains whose lattices are similar but rotated with respect to one another by rotation about the [100] and [001] axes. Since each image is ostensibly nearly "complete," the grain boundaries are roughly parallel to the (010) crystal surface. In an optically thick material, transmission through such a layered crystal would be precluded by the misalignment of the successive grains. However, this crystal is optically thin, permitting the observation of symmetrical diffraction from each of the grains in turn. From the occurrence of similar features in pairs of images, which can be ascribed to features shared by adjacent grains at their interface and the degree of clarity, we can assign a tentative order to the various grains as intersected by the x-ray beam. This is the order in which their images are presented in the figures and in table 1. Most of the features thus appear to be associated with irregularities at the granular interfaces, although radiographic effects from each layer are present.

The seed portion of this crystal takes up most of each image. Space growth was in the [00$\bar{1}$] direction along that one edge of the seed. The absence of a clear demarcation between the seed and new growth in this region in most of the images is in contrast to the terrestrial growth of comparable material. The interface between the seed and the new growth is visible in figure 14, but irregularities in this grain do not appear to propagate into the new growth in the central portion of the disc. In the other grains, irregularities from the seed indeed appear to have propagated into the part grown in microgravity. Toward the edge of the disc, the irregularity observed in all of the grains is consistent with rapid growth anticipated from the defects observed near the edge of the seed. Irregularities near the edge in such systems before faceting becomes fully developed later in growth are observed optically as well.

The last of these images, figure 16, differs in shape along the growth edge from the others. It thus appears that new growth did not occur uniformly on all layers. On this one layer, growth appears to have been much slower than on the others, although this may represent initial etching of the seed crystal associated with premature contact with the solution.

Figure 11. High-resolution (220) 8 keV diffraction image in Bragg geometry from the approximately (220) surface of PbSnTe crystal grown on STS 61A. The growth direction is [001]. Lighter areas diffract more strongly.

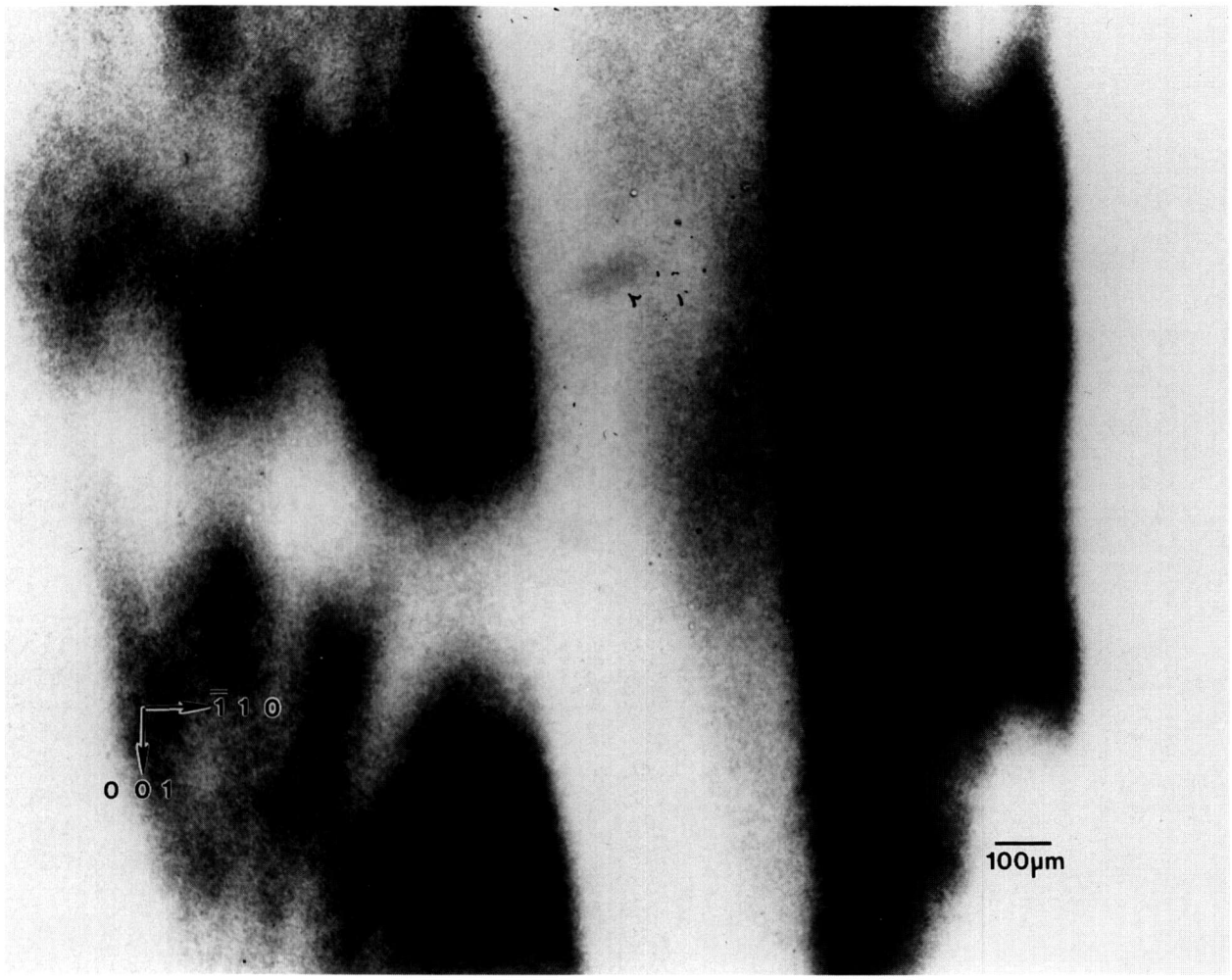

Figure 12. Enlargement of central portion of figure 11, (220) diffraction. Darker areas diffract more strongly.

5.4 Gallium Arsenide

A terrestrial crystal of lightly selenium-doped Bridgman grown gallium arsenide diffracts over several degrees. Growth of identical material by identical techniques is scheduled for an early space flight. A low resolution diffraction image of the terrestrial crystal, achieved by rocking the crystal 4° around a $[11\bar{2}]$ axis during diffraction is shown in figure 17. An infrared image of the same crystal is shown in figure 18. The similarity of these two images is striking. The additional information in a full high-resolution diffraction image of this crystal, figure 19, is evident. An enlargement of a portions of this image is shown in figure 20.

The demarcation of the Czochralski seed from the new Bridgman growth is very clear in those images in which this region is in diffraction. The seed/growth boundary is delineated in two ways. First, toward the periphery of the boule it marks a smooth limit to diffraction, past which the lattice does not diffract under the same conditions. Thus, either the lattice constant, or orientation, or both differs in the new growth. Second, in the one region of the seed interface supporting diffraction from both sides, the mesoscopic structure of the growth is observed to be transformed at the interface. The cellular structure of the seed is characteristic of diffraction images of Czochralski-grown undoped gallium arsenide. In the new Bridgman growth, the formation of cells appears to be completely suppressed. Freedom from other demarcation, however, indicates that, to the extent permitted by the lattice parameter match, the two lattices continue uninterrupted. The lattice parameter mismatch appears to set up a gradual warping of the crystal lattice. Further analysis of the features observed is precluded by the inability to observe diffraction in Laue geometry.

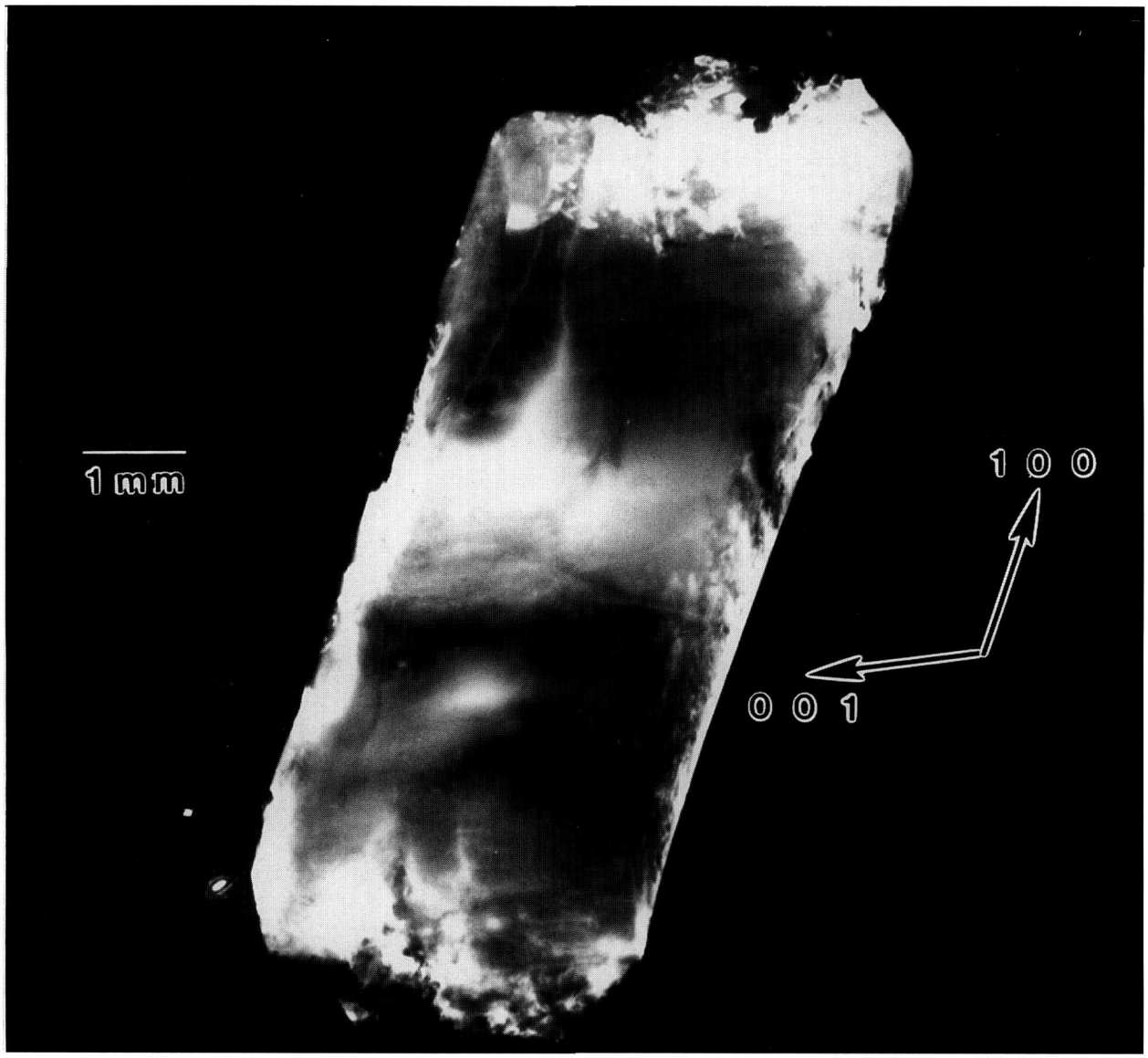

Figure 13. High-resolution (300) 10 keV diffraction image in Laue geometry from the (001) surface of grain C2 of TGS crystal from Spacelab III. Lighter areas diffract more strongly.

6. Summary of Initial Observations

These results are summarized in the following sections. The three crystals of mercuric iodide, two of lead tin telluride, one triglycine sulfate crystal, and one gallium arsenide crystal show remarkable differences in their irregularities.

6.1 General Effects of Microgravity

Seven very different crystals do not provide a sample that is sufficiently large to form definitive conclusions. Nevertheless, our observations provide guidance for further evaluation and crystal growth. The formation of pervasive multiple phases observed in two terrestrial crystals appears to have been greatly suppressed on growth of two comparable crystals in microgravity.

6.2 Mercuric Iodide

A terrestrial specimen comparable in growth procedure and source material composition to one grown on Spacelab III displays more than one phase. The features containing the additional phase appear to be globular in part and partly in the form of thin layers oriented along the {100} crystallo-

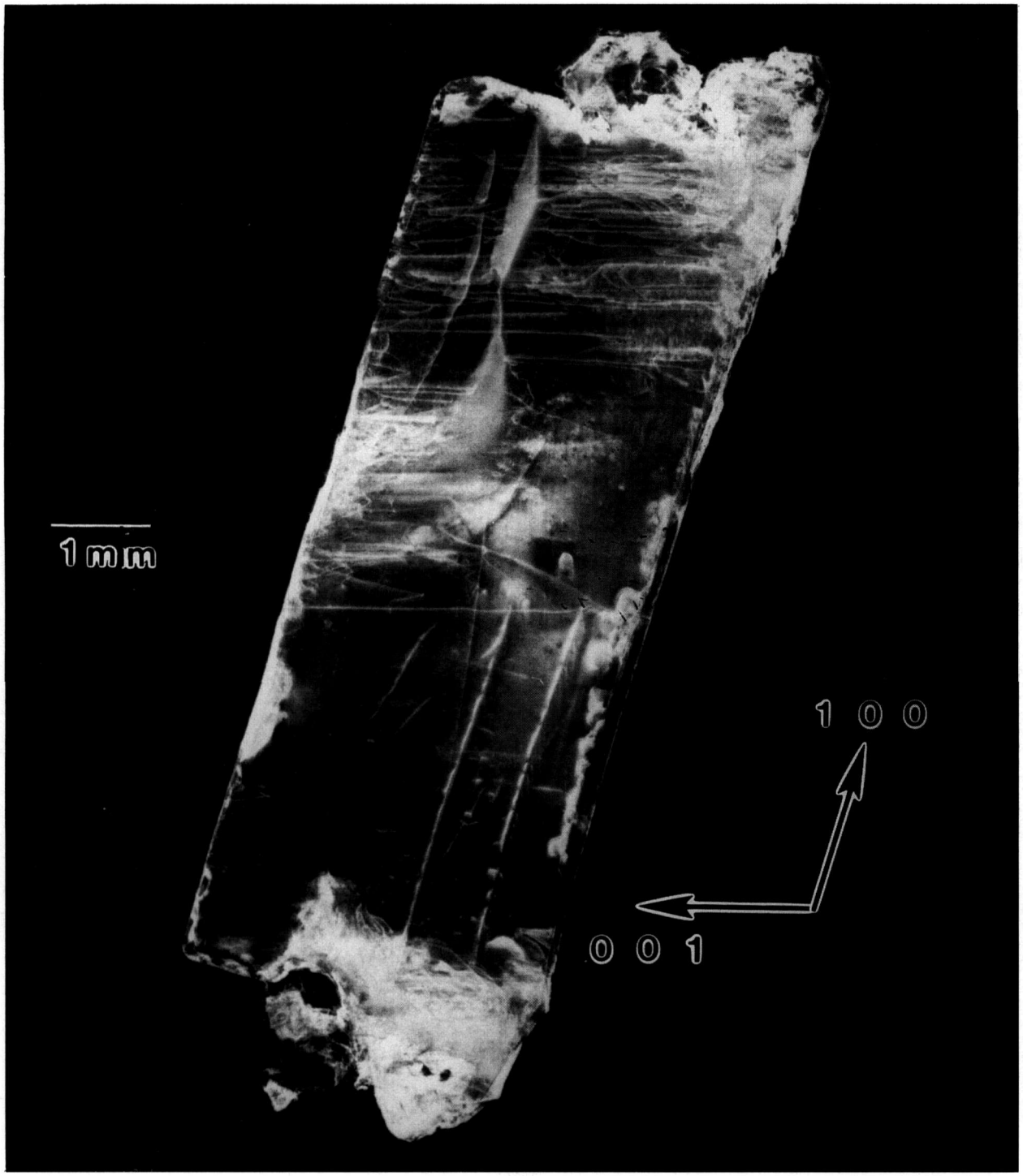

Figure 14. High-resolution (200) 10 keV diffraction image in Laue geometry from grain B1 of TGS crystal from Spacelab III. Lighter areas diffract more strongly.

graphic directions of the matrix. One of these features appears to have initiated a sharp lattice twist by 10 minutes of arc around an axis aligned with the growth direction and to have stiffened the two resulting subgrains. Formation of the additional phase material is suppressed both in a comparable Spacelab III crystal and a recently grown high purity terrestrial crystal. At the same time, the general regularity of the lattice of these latter crystals is lower than in the earlier terrestrial crystal. Superior

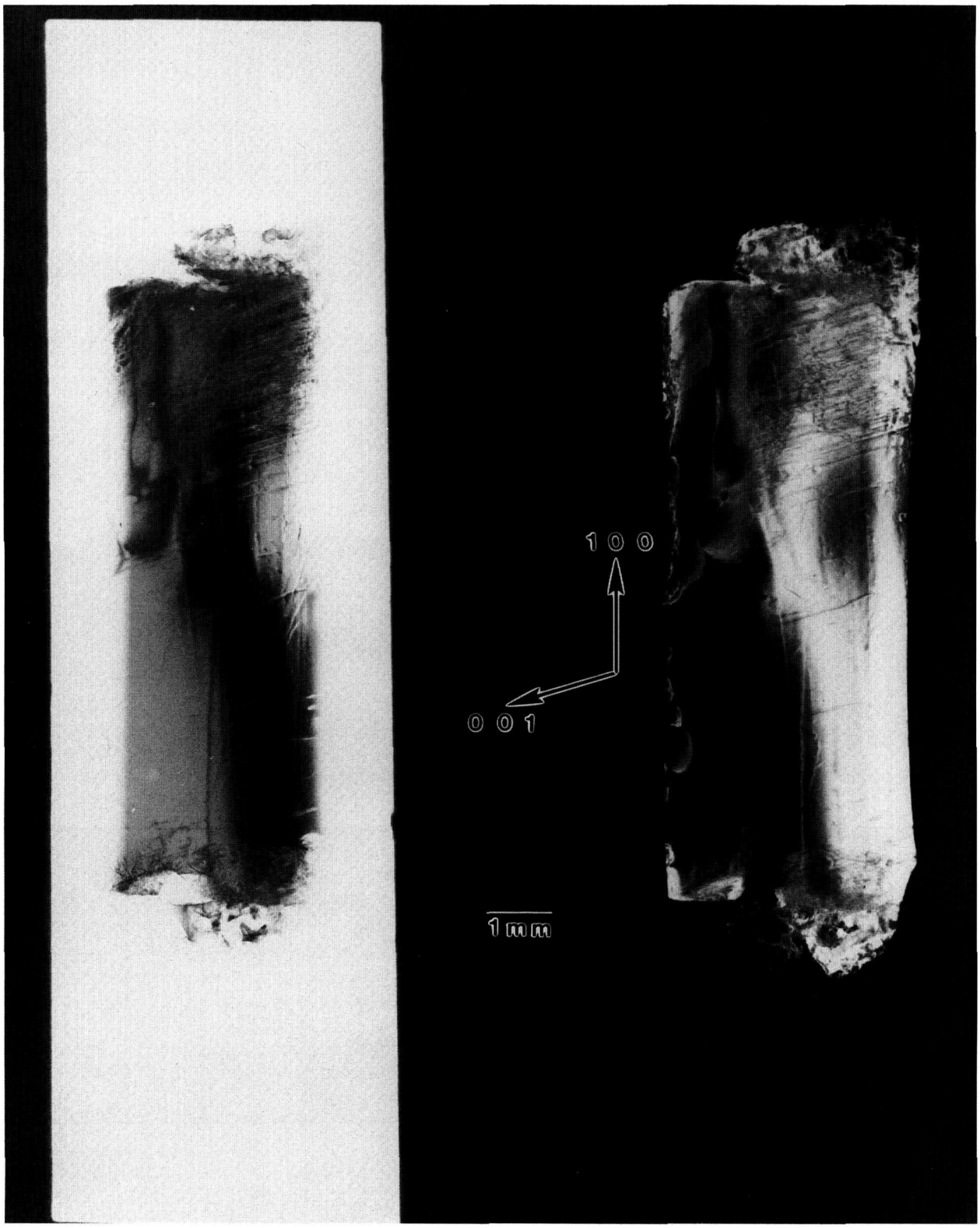

Figure 15. High-resolution (001) 10 keV diffraction image in Laue geometry from grain A2 of TGS crystal from Spacelab III. Lighter areas diffract more strongly.

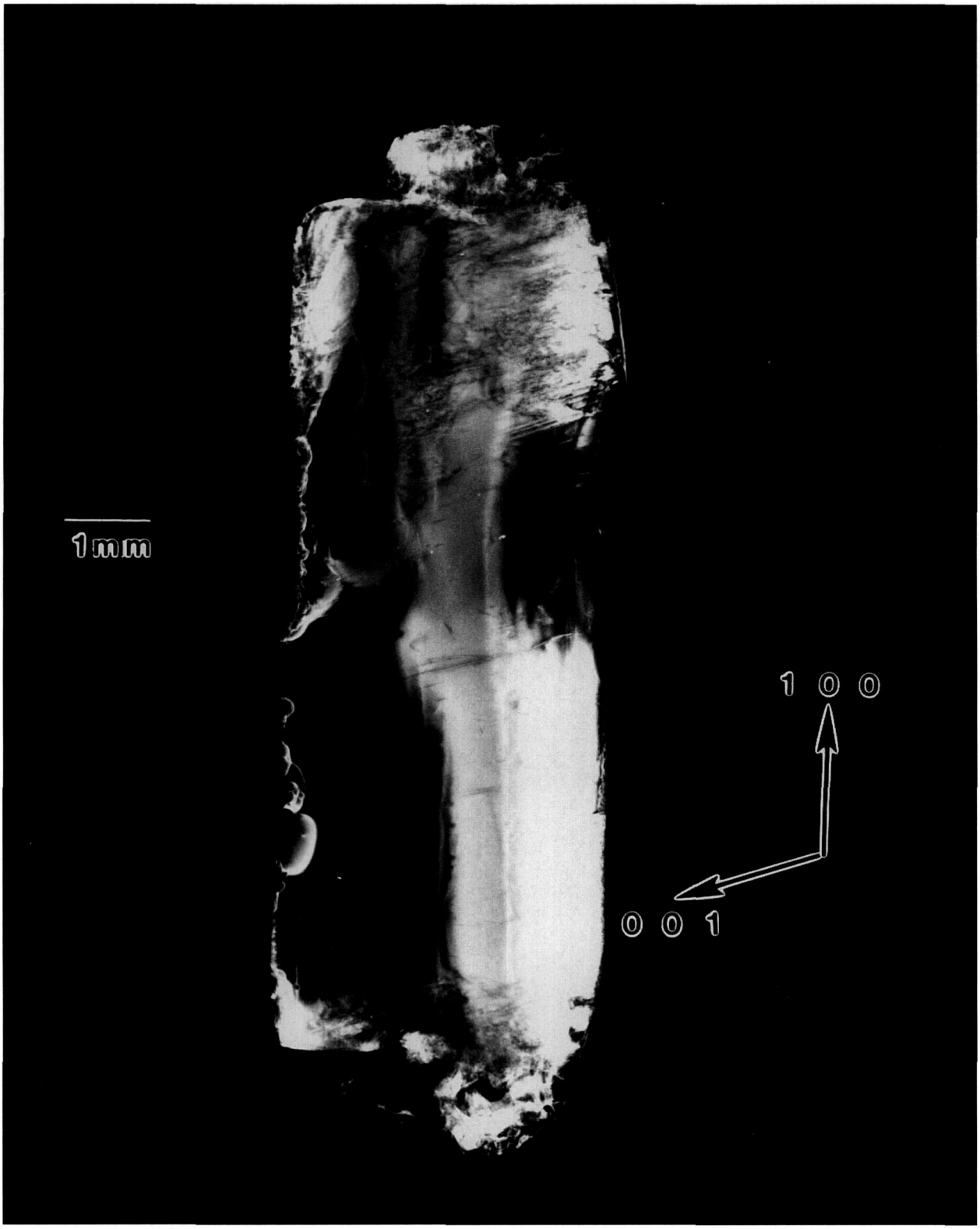

Figure 16. High-resolution (001) diffraction image in Laue geometry from grain A3 of TGS crystal from Spacelab III. Lighter areas diffract more strongly.

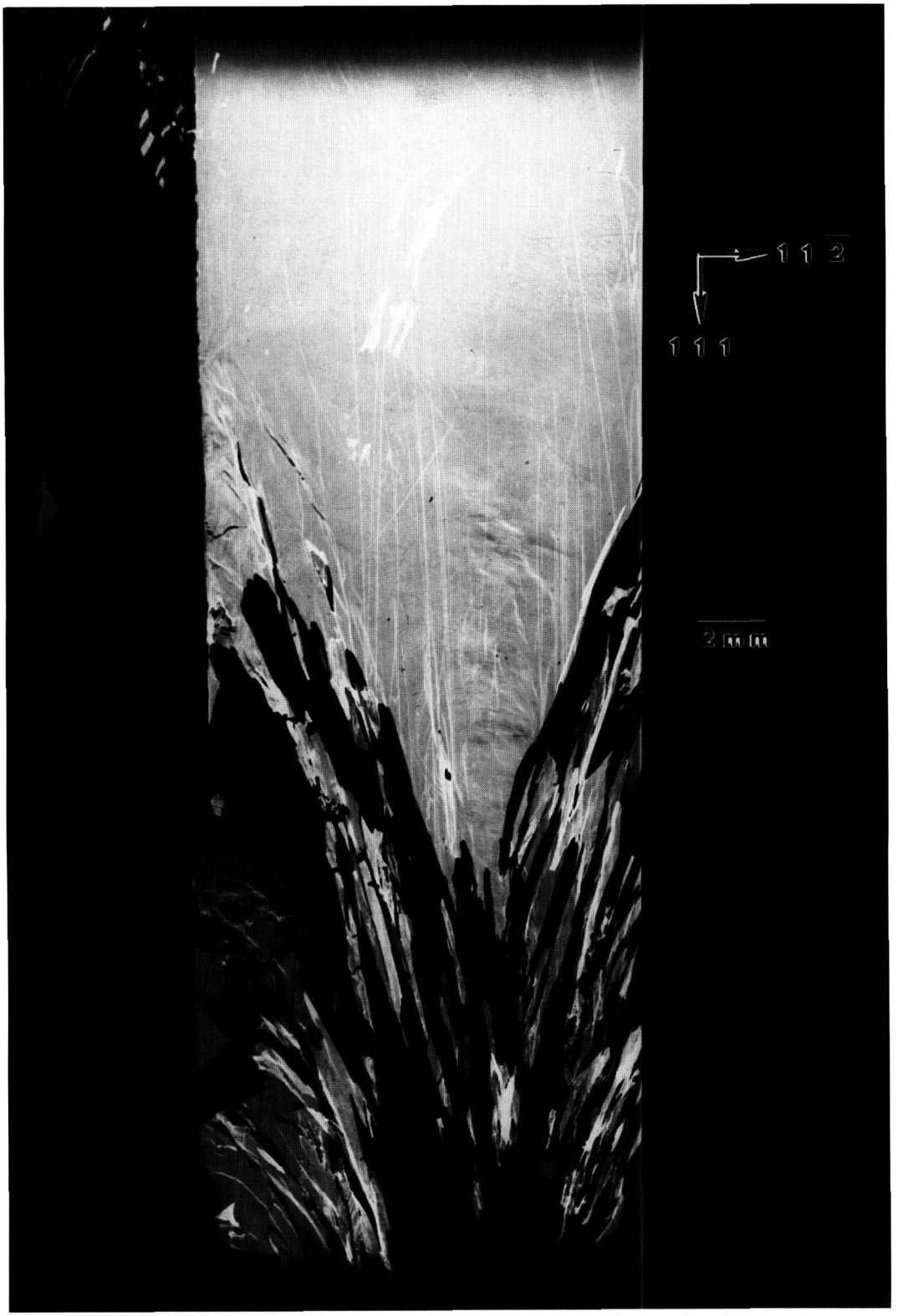

Figure 17. Low resolution (220) 8 keV diffraction image of the approximately (220) surface of terrestrial GaAs crystal, in Bragg geometry. The sample was rocked through an angle of 4° during this exposure; the growth direction is [111]. Lighter areas diffract more strongly.

Figure 18. Infrared image of crystal shown in figure 17.

performance of detectors made from these materials thus appears to be limited far more by sharp discontinuities associated with additional phase(s) than by slow variation in the lattice.

6.3 Lead Tin Telluride

The mesoscopic structure of lead tin telluride appears also to be influenced strongly by the intrusion of additional phase material. But in this instance all available evidence points to identification of the additional phase with the major constituents. Indeed, the presence of two phases has been predicted for systems with tellurium concentration very close to 50%.

Although the STS 61A crystal has grain structure that appears to be similar to that of the comparable terrestrial crystal, the formation of the subgrain variation characteristic of the terrestrial sample is suppressed in microgravity. This suppression is correlated with the predicted thermo-solutal stability in microgravity.

6.4 Triglycine Sulfate

Interpretation of the space-growth of triglycine sulfate has been complicated by the layered structure that we observe in the seed. Defects in one of these layers appear not to have propagated in the central portion of the disc in microgravity, while defects in the other seed layers appear indeed to have propagated into the new growth. In addition, one of the seed layers appears to have grown at a rate slower than the others, but this may simply represent inadvertent contact between the seed and the solution prior to space growth.

In the IML 1 mission scheduled, use of multifaceted natural seeds is planned. They will be characterized not only for various physical properties but also for defects and structural properties prior to flight. In this way, definite information should be obtained on the generation and propagation of defects during growth.

6.5 Gallium Arsenide

Although gallium arsenide has not yet been grown in space in the NASA program, space growth directly comparable to that used for our terrestrial sample is scheduled shortly. Meanwhile, the mesoscopic structure of our terrestrial Bridgman regrowth has been observed to differ from that of the original Czochralski-grown material. The Bridgman-grown lattice also appears to be warped more than that of the Czochralski-grown seed.

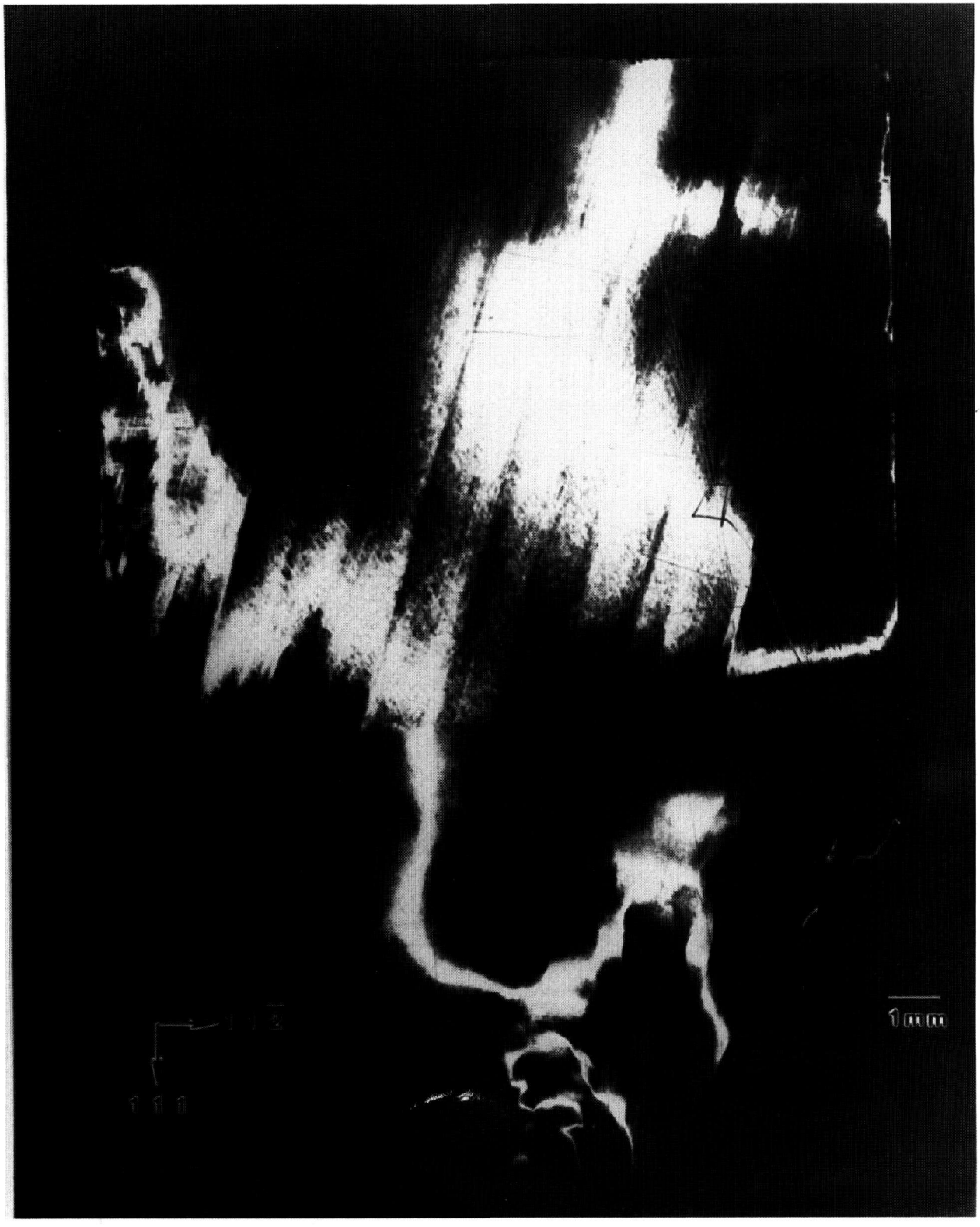

Figure 19. High-resolution (220) stationary diffraction image from the approximately (220) surface of terrestrial GaAs crystal in Bragg geometry. The growth direction is [111]. Lighter areas diffract more strongly.

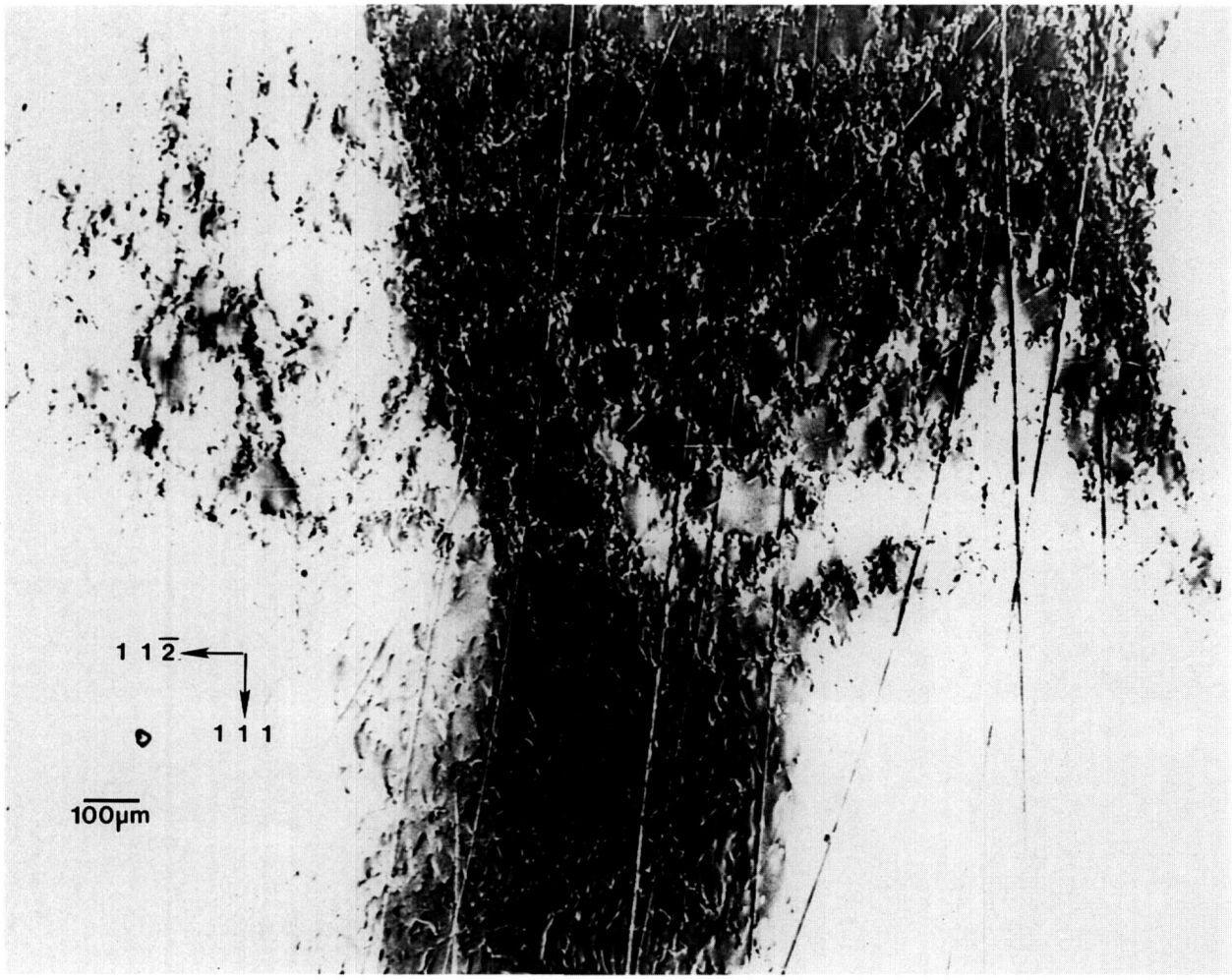

Figure 20. Enlargement of figure 19, (220) diffraction, near the center of the seed/new growth interface. Darker areas diffract more strongly.

Acknowledgment

The results of the collaboration described in this report would have not have been possible without the multifaceted support of the NASA Microgravity Sciences and Applications Division. Thanks to this strong NASA support, strong guidance has been given to the growth of several important crystal systems.

7. References

[1] van den Berg, L., Schnepple, W. F., Growth of Mercuric Iodide in Spacelab III, 19th SAMPE Tech. Conf. 754 (1987).
[2] van den Berg, L., and Schnepple, W. F., Nucl. Inst. Meth. Phys. Res. **A283**, 335 (1989).
[3] Steiner, B. W., Kuriyama, M., and Dobbyn, R. C., Prog. Crystal Growth Char. **20**, 189 (1990).
[4] Kuriyama, M., Steiner, B. W., and Dobbyn, R. C. Ann. Rev. Mat. Sci. **19**, 183 (1989).
[5] Steiner B. W., Kuriyama M., Dobbyn, R. C., and Laor U., J. Res. Natl. Bur. Stand. (U.S.) **93**, 577 (1988).
[6] Steiner B. W., Laor, U., Kuriyama, M., Long, G. G., and Dobbyn, R. C., J. Cryst. Growth **87**, 79 (1988).
[7] Kuriyama, M., Steiner, B. W., Dobbyn, R. C., Laor, U., Larson, D., and Brown, M., Phys. Rev. **B38**, 12421 (1988).
[8] Steiner, B., Kuriyama, M., Dobbyn, R. C., Laor, U., Larson, D., and Brown, M., J. Appl. Phys. **66**, 559 (1989).
[9] Nicolay, Y. F., and Dupuy, M., Nucl. Inst. Meth. Phys. Res. **A283**, 355 (1989).
[10] Crouch, R. K., Fripp, A. L., Debnam, W. J., Carlson, F. M., and Simchick, R. T., Results from a Compound Semiconductor Crystal Growth Experiment in a Low Gravity Environment, SAMPE Electronics Materials and Processes Conf., Santa Clara, June 22-25 1987.
[11] Harman, T. C., J. Nonmetals **1**, 183 (1973).
[12] Butler, J. F., and Harman, T. C., J. Electrochem. Soc. **116**, 260 (1969).

About the authors: Bruce Steiner is a Scientific Advisor to the Director of the Materials Science and Engineering Laboratory of NIST; when the diffraction imaging work described here was carried out, Ronald C. Dobbyn, David Black, Harold Burdette, Masao Kuriyama, and Richard Spal constituted the Synchrotron Radiation Group of this Laboratory. Lodewijk van den Berg is a section head at EG&G Energy Measurements in Goleta, California; as a Payload Specialist on Spacelab III, he grew the space mercuric iodide crystal that was examined in this effort. Archibald Fripp is a senior scientist with the NASA Langley Research Center in Hampton, Virginia; Richard Simchick, a scientist with Lockheed Engineering and Sciences Company, works closely with Dr. Fripp at the NASA Langley Research Center in research applicable to the microgravity processing of electronic materials, using lead tin telluride as a model material. Ravindra B. Lal is a Professor of Physics at Alabama A&M University in Huntsville and was the principal investigator for an experiment flown on Spacelab III; Ashok Batra works closely with Professor Lal as a Research Assistant Professor in the terrestrial growth of triglycine sulfate crystals and the design and execution of experiments for their growth in microgravity. David Matthiesen and Brian Ditchek are materials scientists at GTE Laboratories who grow gallium arsenide teresterially and design and carry out experiments for its growth in microgravity.

Volume 96, Number 3, May–June 1991
Journal of Research of the National Institute of Standards and Technology

[J. Res. Natl. Inst. Stand. Technol. 96, 333 (1991)]

Root Projection of One-Sided Time Series

Volume 96 | Number 3 | May–June 1991

John A. Simmons

National Institute of Standards and Technology,
Gaithersburg, MD 20899

Until recently it has been impossible to accurately determine the roots of polynomials of high degree, even for polynomials derived from the Z transform of time series where the dynamic range of the coefficients is generally less than 100 dB. In a companion paper, two new programs for solving such polynomials were discussed and applied to signature analysis of one-sided time series [1]. We present here another technique, that of root projection (RP), together with a Gram-Schmidt method for implementing it on vectors of large dimension. This technique utilizes the roots of the Z transform of a one-sided time series to construct a weighted least squares modification of the time series whose Z transform has an appropriately modified root distribution. Such a modification can be employed in a manner which is very useful for filtering and deconvolution applications [2]. Examples given here include the use of boundary root projection for front end noise reduction and a generalization of Prony's method.

Key words: causal boundary roots; Gram-Schmidt algorithm; Prony's method; root projection; signal processing; time series; Z transforms.

Accepted: February 25, 1991

1. Introduction

Let y be a complex number. We define $\widehat{y}$ to be the "geometric sequence" vector with components $\widehat{y}_n = y^n$, $n = 0,..., N$. Then given a time series of length $N+1$ represented by the vector $a{:}N$, $a(y) = \widehat{y}^T a$; and the condition that y be a root of $a(y)$ is:

$$\widehat{y}^T a = 0. \tag{1}$$

Given a (possibly complex) time series $a{:}N$ and any set of complex roots of eq (1), $y_1,..., y_M$, $M \leq N$, we shall construct a linear projection to modify $a{:}N$ to another time series $b{:}N$ so that $b(y_\alpha) = 0$, $\alpha = 1,..., M$, and so that

$$(a-b)^T G (a-b)^*, \tag{2}$$

where "*" means complex conjugate, is a minimum for any positive definite Hermitian weighting matrix G. In other words $b{:}N$ is the series closest to $a{:}N$ in a weighted least squares sense, which has a chosen set of complex numbers among the roots of its Y transform. We refer to this process as *root projection* (RP).

In a slightly more general context consider a (possibly complex) time series $a{:}N$, a set of complex numbers $y_1,..., y_M$ with $M \leq N$, and a set of complex values $v_1,..., v_M$. We construct a time series $b{:}N$ whose difference from $a{:}N$ is minimal among those series with $b(y_\alpha) = v_\alpha$, $\alpha = 1,..., M$.

Because of the slight complication due to the complex numbers involved, we shall set and solve the above minimization problem using Lagrangian

multipliers. We want to find the time series $b{:}N$ such that:

$$b{:}N = \min_{x:N}\Big[(a-x)^T G^T (a-x)^* + \sum_{\alpha=1}^{M} \lambda_\alpha^* (x^T \widehat{y_\alpha} - v_\alpha) + \sum_{\alpha=1}^{M} \lambda_\alpha (x^{*T} \widehat{y}_\alpha^* - v_\alpha^*)\Big]. \quad (3)$$

Using index notation and differentiating eq (3) with respect $x_j{}^*$ and λ_α^*, we have:

$$\begin{aligned} &G_{ki}(a_i - b_i) = \lambda_\beta (y_\beta^*)^k \\ &b_k (y_\alpha)^k = v_\alpha. \end{aligned} \quad (4)$$

Setting $y_{\alpha k} \equiv (y_\alpha)^k$, $y_{\alpha k}^* \equiv (y_\alpha^*)^k$, and $G_{ij}^{-1} \equiv (G^{-1})_{ij}$; solving for b_i in the first of eqs (4), substituting into the second equation and solving for λ_β:

$$\begin{aligned} &\lambda_\beta = Q_{\beta\alpha}^{-1}(a_j y_{\alpha j} - v_\alpha), \text{ where} \\ &Q_{\alpha\beta} = y_{\alpha k} G_{kl}^{-1} y_{\beta l}^* \end{aligned} \quad (5)$$

whence:

$$b_i = a_i - a_j y_{\alpha j} Q_{\beta\alpha}^{-1} y_{\beta m}^* G_{mi}^{-1} + v_\alpha Q_{\beta\alpha}^{-1} y_{\beta m}^* G_{mi}^{-1}. \quad (6)$$

To make the connection to linear vector space theory, we start with complex N space, $\mathbb{C}^N$, having an inner product given by the positive-definite Hermitian matrix G:

$$(a,b) \equiv a^T G^T b^*, \quad (7)$$

and complex M space, $\mathbb{C}^M$, having an inner product given by the Hermitian matrix Q^{-1} (assumed, here, to be non-singular) with an inner product definition similar to eq (7) and with vectors such as v (with components v_α, $\alpha = 1,..., M$) [3]. We define the mapping Y from $\mathbb{C}^N$ to $\mathbb{C}^M$ by:

$$Y \equiv [Y_{\alpha i} = y_{\alpha i};\ \alpha = 1,..., M,\ i = 1,..., N]. \quad (8)$$

Y has an adjoint mapping $Y^\dagger$, from $\mathbb{C}^M$ to $\mathbb{C}^N$ given by the condition $(Y^\dagger v)^T G^T a^* = v^T (Q^{-1})^T (Ya)^*$, from which

$$Y^\dagger = G^{-1} Y^{*T} Q^{-1}. \quad (9)$$

In matrix notation eq (6) becomes:

$$\begin{aligned} &b = (I-P)a + Y^\dagger v \\ &P = Y^\dagger Y, P^2 = P, \end{aligned} \quad (10)$$

the projection property, $P^2 = P$, following directly from eq (6). It also follows directly from eq (6) that the range of $I-P$ is contained in the null space of Y^T, while the range of P is contained in the range of $Y^\dagger$. Since both $I-P$ and P as well as the null space of Y and the range of $Y^\dagger$ decompose $\mathbb{C}^N$ (the latter being an orthogonal decomposition), $I-P$ and P are orthogonal mappings onto the respective spaces. The mapping $I-P$ is, then, the desired root projection.[1]

2. A Gram-Schmidt Algorithm for Root Projection

In order to carry out root projection, we need an orthogonal basis for the range of $Y^\dagger$ in the unitary space with metric G. From eq (8) we see that this space is spanned by vectors of the form $G^{-1}\widehat{y_\alpha}$. For any two such vectors:

$$(G^{-1}\widehat{y_\alpha})^T G^T (G^{-1}\widehat{y_\beta})^* = (G^{-\frac{1}{2}}\widehat{y_\alpha})^T (G^{-\frac{1}{2}}\widehat{y_\beta})^*, \quad (11)$$

where

$$(G^{-\frac{1}{2}})^2 = G^{-1}. \quad (12)$$

Thus, an orthonormal basis chosen from the vectors $G^{-\frac{1}{2}}\widehat{y_\alpha}$ by applying the Gram-Schmidt process using the ordinary unitary inner product will determine an orthogonal basis for the range of $Y^\dagger$ in the unitary space with metric G by multiplying each vector by $G^{-\frac{1}{2}}$. However, when only a limited number of vectors need to be projected, the projection can be carried out more efficiently by:

i) transforming the vector a to $G^{\frac{1}{2}}a$
ii) projecting $G^{\frac{1}{2}}a$ in the Euclidean norm using the modified basis $G^{-\frac{1}{2}}\widehat{y_\alpha}$
iii) transforming the result back by multiplication with $G^{-\frac{1}{2}}$. (13)

The root projection process is especially simple when the set of y_α's is closed under complex conjugation (i.e., if y_α belongs to the set, so does y_α^*) for the case of simple time or frequency weighting. Because of the closure under complex conjugation, the $\widehat{y_\alpha}$ vectors can be replaced by the vectors consisting of their real and imaginary parts, $\Re(\widehat{y_\alpha})$ and

[1] A concise exposition of the linear algebra required for root projection, with $G = I$ can also be given in terms of the QR decomposition [4] (although we use the "transposed" form of that in [4]).

$\Im(\widehat{y_\alpha})$. By the time weighting case we mean that G^{-1} and $G^{-\frac{1}{2}}$ are diagonal (and, therefore, positive and real), so that no unitary transformation is required. In that case each of the rows is multiplied by the same factor. Thus, the basis for the range of $Y^\dagger$ can be obtained by applying the real form of the Gram-Schmidt process to the vectors $\Re(G^{-\frac{1}{2}}\widehat{y_\alpha})$ and $\Im(G^{-\frac{1}{2}}\widehat{y_\alpha})$. The root projection of a real time series, outlined in eq (13), will, in this case, also be real.

In the case of frequency weighting we mean scalar weights applied to the DFT components of a time series. The DFT process is, itself, a unitary transformation, and the orthogonalization process can be made real by the trick of putting $\Re(\tilde{a}_k)\sqrt{2}$ in place of the DFT component $\tilde{a}_k$ and putting $\Im(\tilde{a}_k)\sqrt{2}$ in place of the DFT component $\widetilde{a_{-k}}$. The real dot product of these series is the same as the unitary dot product of the DFT's. A diagonalized frequency weighting metric can then be applied to these transformed DFT series and the diagonalization process carried out. If projection is to be carried out on a real time series, it can be done in these transformed coordinates, otherwise, the new orthogonal DFT coordinates have to be reconstructed before doing the projection. Of course, the projected DFT series has to be reconverted to time series form after projection.

The real form of root projection can also be obtained by using QR algorithms such as given in LINPACK [4]. However, the implementation of the QR algorithms in LINPACK is CPU core consuming for large N and does not easily allow storage of the orthogonal matrix. These disadvantages are overcome using the Gram-Schmidt method given above. Employing a form of the Gram-Schmidt algorithm suggested by G.W. Stewart [5,6], a dynamically dimensioned Fortran 77 code was constructed allowing efficient use of a user selected buffer with the option of saving the orthogonalized basis vectors on disc for rapid repeated projection [7]. For y_α with modulus greater than one, projection was carried out using the root reciprocal to avoid overflow. In this case the series starts at $y_\alpha{}^{-N}$ and goes up to 1, a procedure equivalent to reversing the series to be projected.

3. Examples

We have principally used root projection for deconvolution. That topic is discussed in the third paper in this series [2]. Here we give four illustrative examples of root projection. The fourth example, applying root projection to Prony's method may have some practical use. However, a detailed study has not been conducted.

1. Figure 1 shows a normalized 101 point Gaussian with 60 dB dynamic range (ratio of center to edge points). In figure 2 we show the 800 point result of convolving the first 700 points of the experimental waveform of figure 2.4a in [1] with the Gaussian of figure 1. The noise pattern shown in figure 3 results from rounding off the time series in figure 2 to 8 bit accuracy. The rounded-off time series is not shown. The standard deviation of this noise is 6.02×10^{-4}.

A basis for all 800 point time series can be formed from the waveform of figure 2 together with the 799 geometric root vectors formed from the roots of the Y transform of that waveform. The noise vector of figure 3 must then be a linear combination of all these vectors, while the true signal

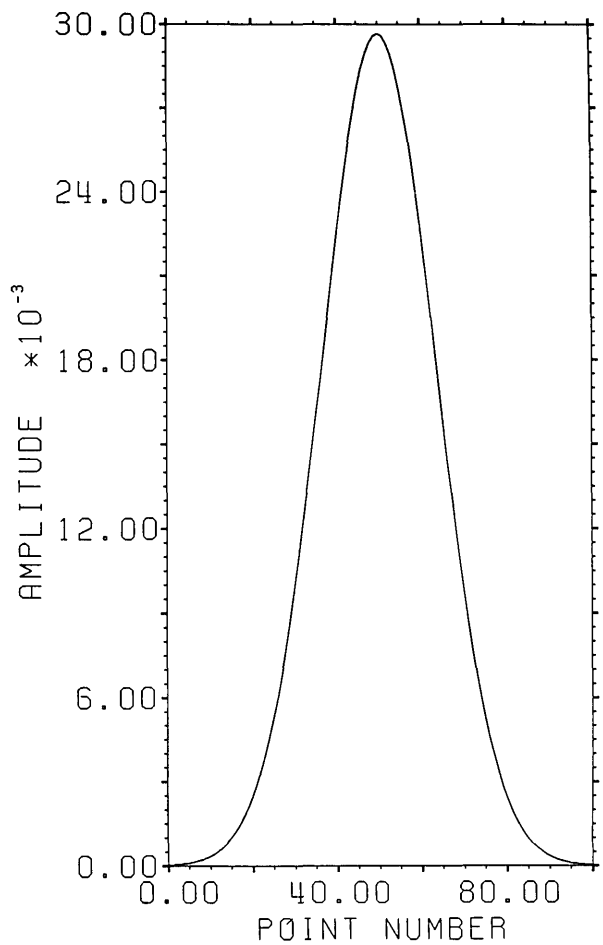

Figure 1. Normalized 101 point Gaussian with 60 dB dynamic range.

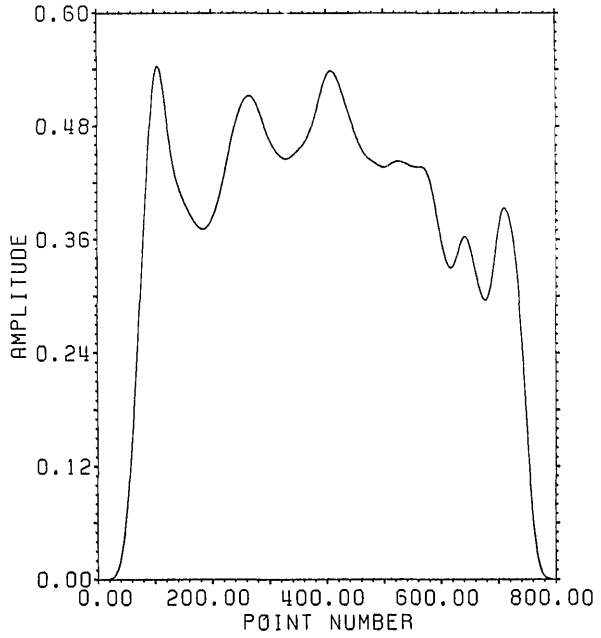

Figure 2. Convolution of the first 700 points of the time series in figure 2.4a of reference [1] with the 101 point time series of figure 1.

vector of figure 2, which is one of the basis elements, is orthogonal to all of the 799 geometric root vectors. If we use the geometric root vectors built from the Gaussian of figure 1 and apply root projection to the rounded-off series built from the series in figure 3, the noise of the resultant series will be reduced in magnitude, since it is orthogonal to the 100 geometric root vectors formed from the roots of the Y transform of the Gaussian. The standard deviation of that noise is 5.6×10^{-4}, almost exactly the theoretical estimate of $(\frac{7}{8})^{\frac{1}{2}}$ expected for Gaussian noise. If the 799 geometric root vectors from both the Y transforms of the Gaussian and the waveform of figure 2.4a in [1] are used for the projection, then the error appears as in figure 4—a mini-image of the correct time series—with a standard deviation of 2.0×10^{-5} (again approximately $(800)^{-\frac{1}{2}}$ of the original noise deviation). This time, of course, the noise is biased with a mean value of 4.9×10^{-5}. This curve also indicates the precision of the root finding and Gram-Schmidt routines used in these calculations.

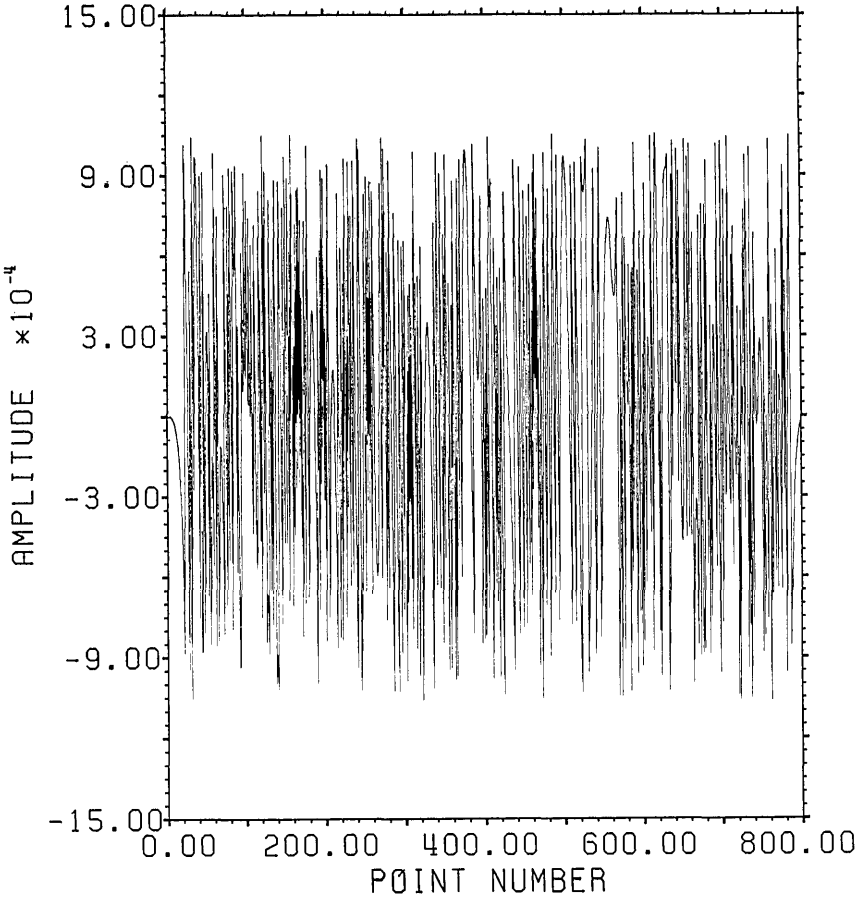

Figure 3. 8-bit roundoff noise for the series in figure 2.

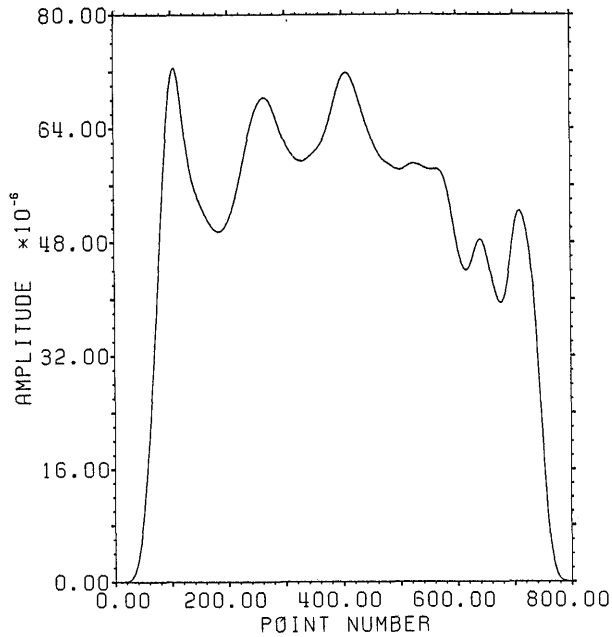

Figure 4. Error from projecting the series in figure 2 plus noise through all roots of the Y transform of the series in figure 2. Note that the remaining error is proportional to the original series.

2. The use of an extensive causal boundary root set for front end filtering is shown in figures 5a and 5b. To the Gaussian filtered curve of figure 2.3a in [1] we add the -40 dB noise distribution shown in figure 5a and apply root filtering using only the causal boundary roots from figure 2.1b of [1]. The noise after filtering showing extensive front end reduction is shown in figure 5b.

3. A more general example of root projection employing both time and frequency weighting is afforded by constructing an approximation to an optimal maximal-ripple lowpass filter. Figure 6 shows such a filter of 151 elements designed using the Remez exchange algorithm of McClellan et al. [8]. In this case the passband was set for 20% of the Nyquist frequency and the stopband for frequencies over 33% of the Nyquist frequency.

We start with a delta series shifted so that the value 1.0 lies at position #76 in the center of the time interval. To use unweighted projection and 101 roots evenly spaced from $-60°$ to $+60°$ on the unit circle would produce the familiar series obtained

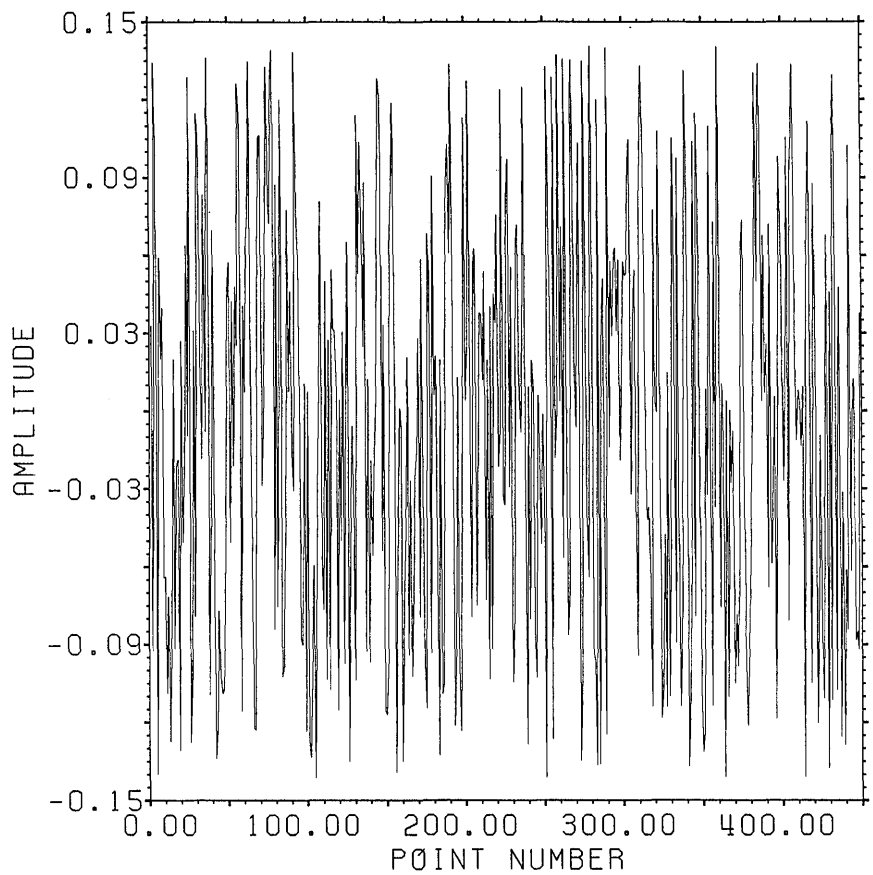

Figure 5a. Noise added to series in figure 2.3a of reference [1].

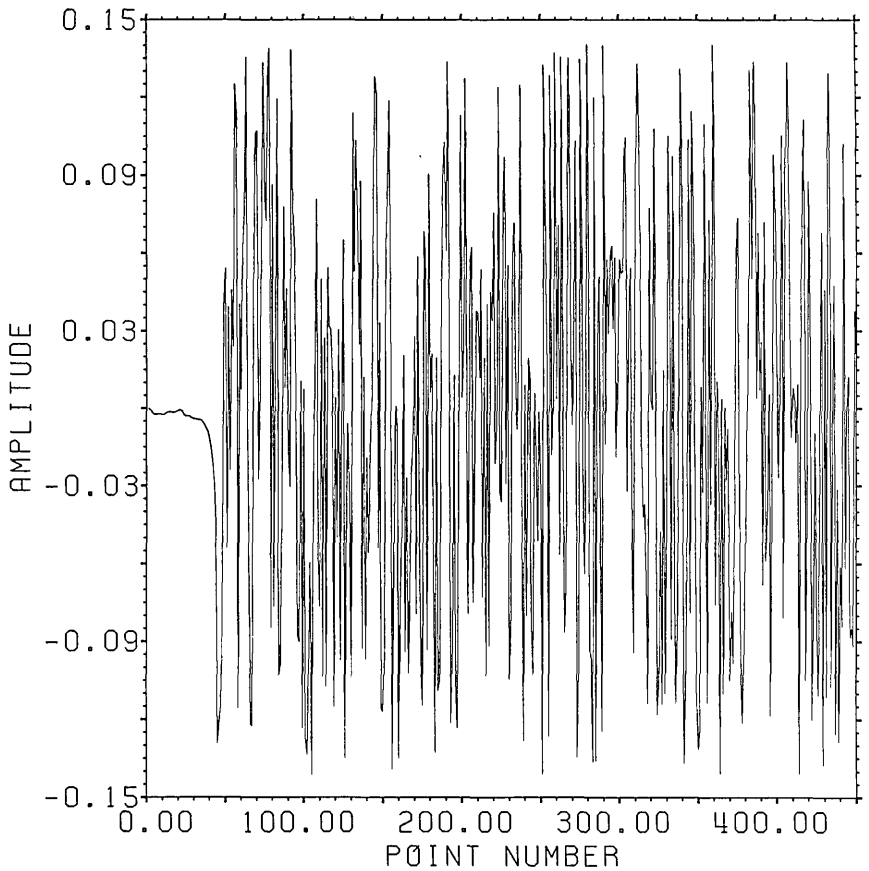

Figure 5b. Noise after filtering using causal boundary roots of the series in figure 2.3a of reference [1] showing extensive front end noise reduction.

from windowing the DFT of the shifted delta series. However, by using time weighted projection with the center symmetric time weight shown in figure 7 and adding two extra roots near each edge of the stopband region, one can produce a filter with 120 dB attenuation in the stopband and $\sim 2.5 \times 10^{-7}$ leakage in the time domain, but with 0.11 ripple in the passband. Without time weighting the filter produced has 64 dB attenuation in the stop band and 0.08 ripple in the passband, but with $\sim 1.0 \times 10^{-3}$ leakage in the time domain.

To remove the passband ripple from the weighted filter we subtract from it the shifted delta series. The resulting series should be close to zero in the passband region. We place 31 roots evenly in the passband and again place two extra pairs of roots near the edge of the band. Applying frequency weighted projection with a simple square frequency weight to hold the stopband attenuation in place, we obtain (after re-adding in the shifted delta series) the doubly root projected filter shown in figure 8, where it is compared with the unweighted doubly projected and optimal filters. The weighted series closely resembles the optimal series of figure 6 with leakage at the ends of the time interval of $\sim 8.0 \times 10^{-5}$ as opposed to $\sim 2.0 \times 10^{-3}$ for the unweighted filter. The leakage for the optimal filter is $\sim 6.0 \times 10^{-8}$. The attenuation in the stopband is 70 dB and the ripple in the passband is 1.0×10^{-4} for the weighted filter while the unweighted filter has 36 dB attenuation and 8.0×10^{-4} ripple. The optimal filter has 116 dB attenuation and 5.0×10^{-5} ripple. The spectra for the three filters are compared in figure 9.

The root patterns in the stopband for the optimal filter are shown in figure 10. As can be seen, they also cluster near the edge of the stopband, but they are not evenly spaced throughout the stopband. Similarly, the roots of the optimal filter minus the shifted delta series, shown in figure 11, exhibit the same uneven spacing and clustering in the passband region.

4. Root projection can be used to provide a "global" least squares generalization to the Prony method. Prony's method is applied to the case of

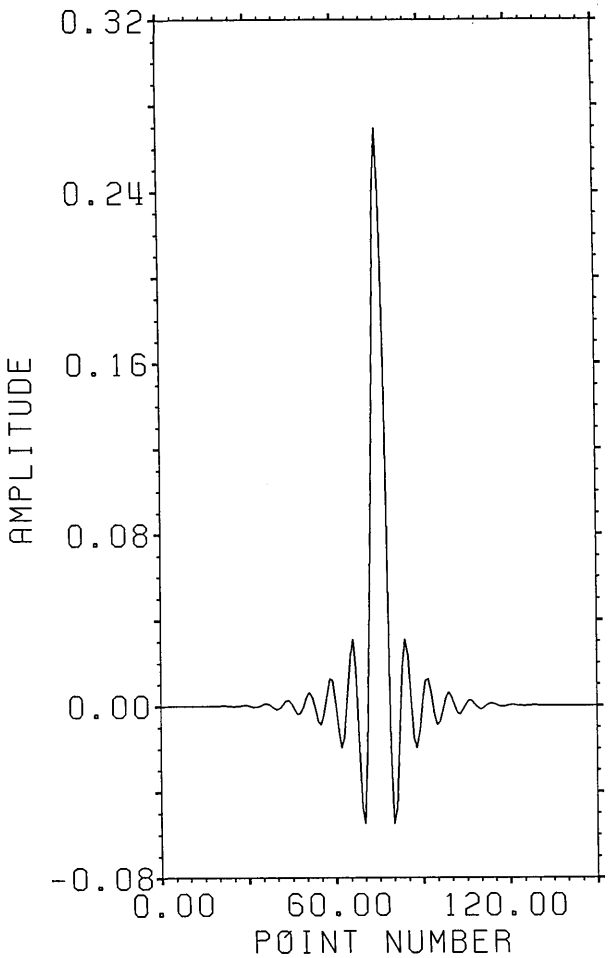

Figure 6. Optimal maximal ripple filter of 151 elements.

$N+1$ data points which one wants to represent as the sum of n decaying exponentials (Prony takes $N+1$ even and $n=(N+1)/2$):

$$F_k=\sum_{j=1}^{n}A_j y_j^k,\ k=0,\ldots,N. \tag{14}$$

Here y_j is thought of as a complex number of modulus less than 1, $y_j=e^{\beta_j T}$, real$(\beta_j)<0$, and T a sampling interval (See, e.g., [9]). Eq (14) can be rewritten, taking advantage of the cyclotomic polynomial expression, as

$$F(y)=\sum_{k=0}^{N}F_k y^k=\sum_{j=1}^{n}A_j\frac{1-(y_j y)^{N+1}}{1-y_j y}. \tag{15}$$

Combining the terms in eq (15) together to form a single fraction gives

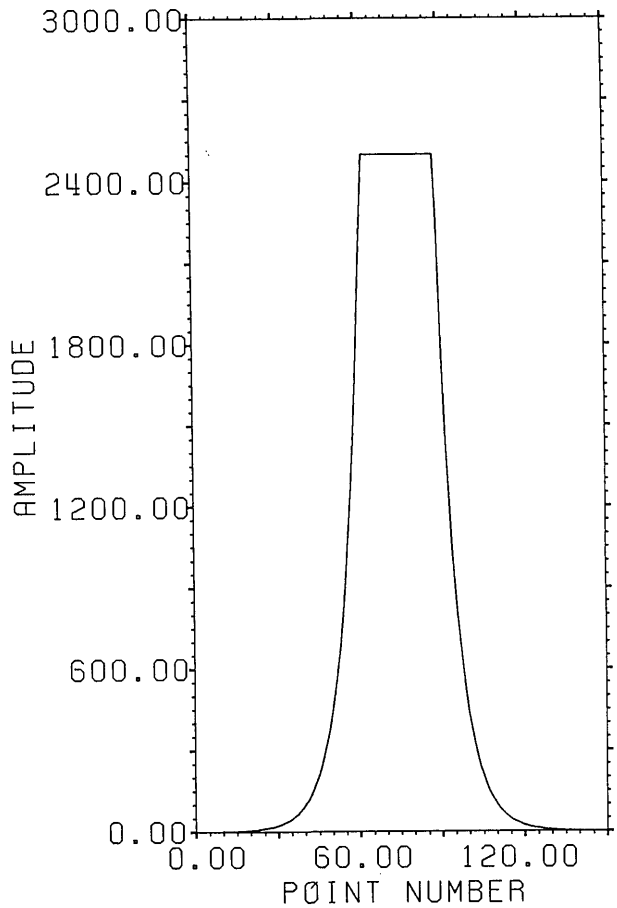

Figure 7. Time weights for approximating an optimal filter.

$$F(y)=\Big[A_1(1-y_2y)\cdots(1-y_ny)\Big(1-(y_1y)^{N+1}\Big)+\ldots$$

$$+A_n(1-y_1y)\cdots(1-y_{n-1}y)\Big(1-(y_ny)^{n+1}\Big)\Big]\Big/$$

$$[(1-y_1y)\cdots(1-y_ny)]. \tag{16}$$

Equation 16, then, has the form

$$F(y)=\frac{P(y)+y^{N+1}R(y)}{Q(y)}, \tag{17}$$

where $Q(y)$ is a polynomial of degree n whose roots are $(y_j)^{-1}$, and $P(y)+y^{N+1}R(y)$ is a polynomial of degree $N+n$ whose $N+1-n$ coefficients from that of y^n to that of y^N are zero. The roots of $Q(y)$ are referred to as the poles of F. Conversely one can show that if eq (17) holds with the corresponding coefficients zero, then setting

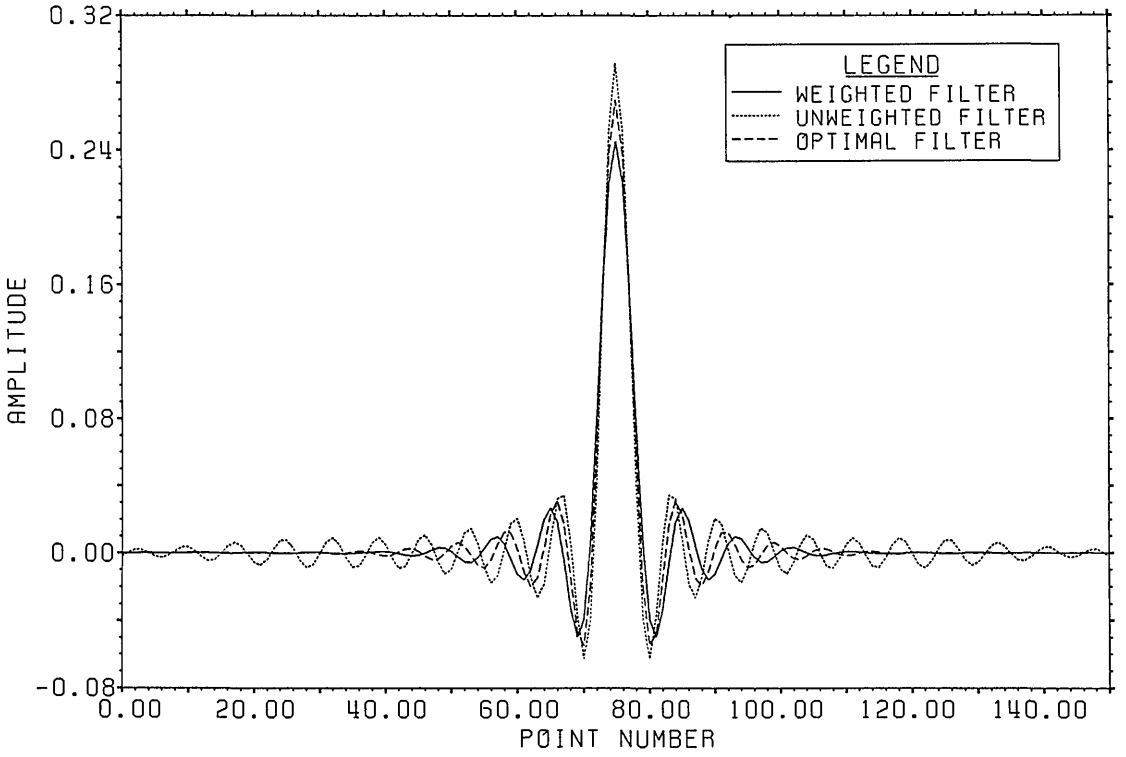

Figure 8. Comparative time plot showing the effects of using weighted projection for approximating an optimal filter.

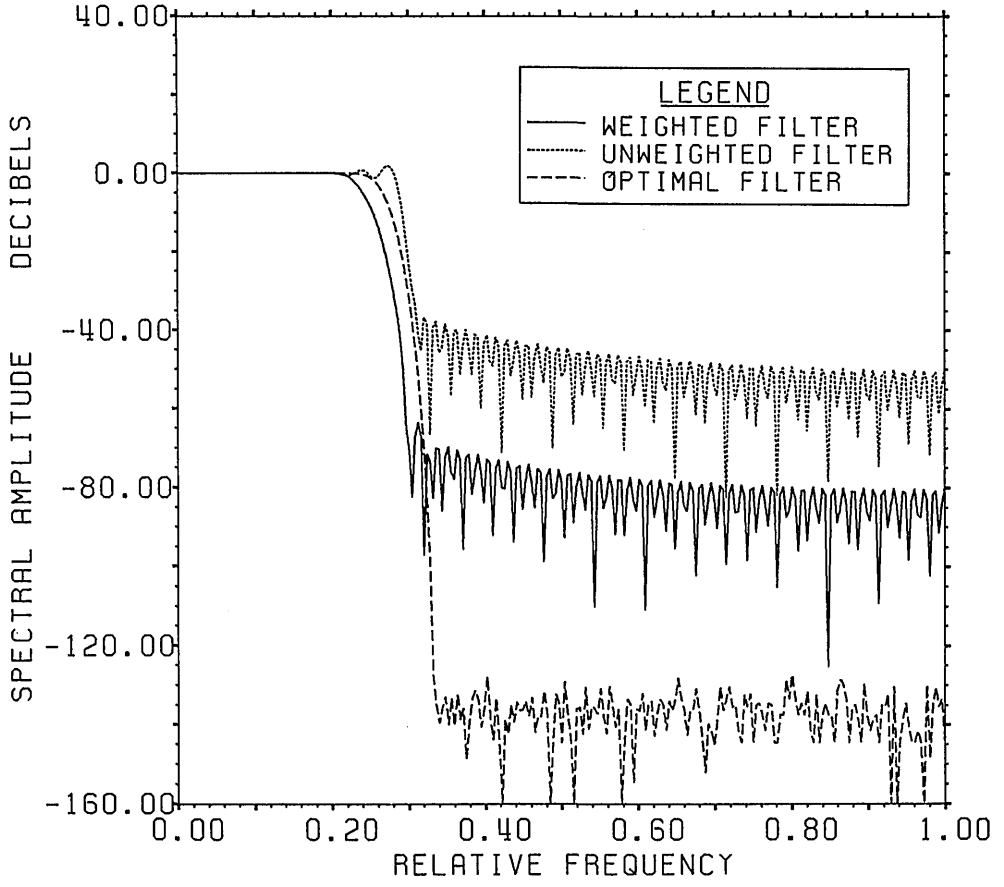

Figure 9. Comparative frequency plot showing the effects of using weighted projection for approximating an optimal filter.

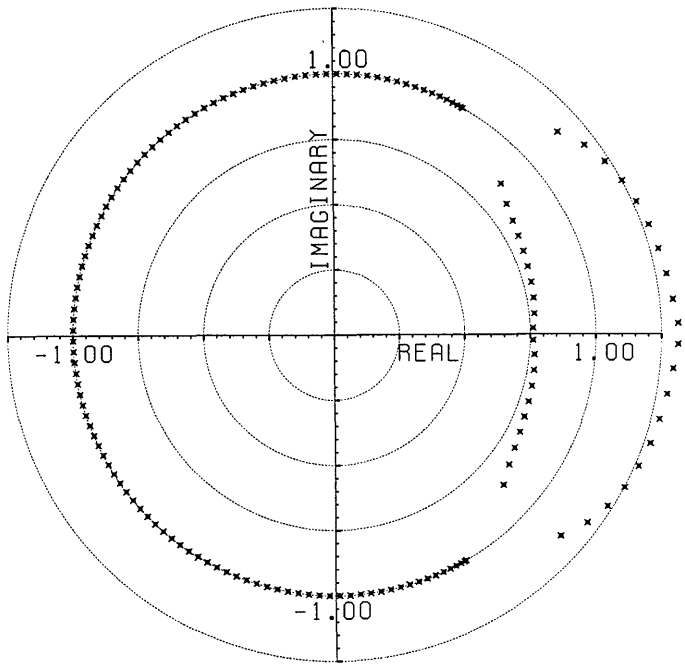

Figure 10. Roots of the *Y* transform of the optimal filter shown in figure 6 showing roots in the stopband.

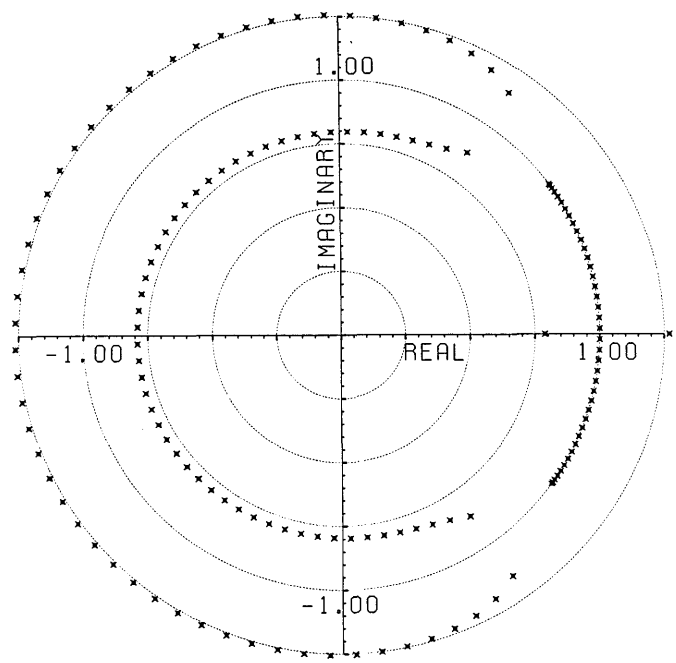

Figure 11. Roots of the *Y* transform of the optimal filter of figure 6 minus the shifted delta series showing roots in the pass band.

$$Q(y)=\prod_{j=1}^{n}(1-y_j y)$$
$$A_j=-\frac{P(1/y_j)}{Q'(1/y_j)} \tag{18}$$

yields the expression 14.

The goal, then, is to solve the algebraic equation:

$$Q(y)F(y)=P(y)+y^{N+1}R(y), \tag{19}$$

where P, Q, and R are all unknown. Only the integer n is given. This is generally a very ill-conditioned process, but when the y_j's are real or occur in conjugate pairs, the following root projection procedure is often successful:

a. Use an initial estimate for the coefficients, c_i, of $P+y^{N+1}R$ by setting $c_i=\kappa_1$ for $i=0,\ldots, n-1$, $c_i=0$ for $i=n,\ldots, N$, and $c=\kappa_2$ for $i=N+1,\ldots, N+n$; and use for the diagonal weighting matrix, W, the series w_l with $w_l=\lambda_1$ for $l=1,\ldots, n$, $w_l=1$ for $l=n+1,\ldots, N+1$, and $w_l=\lambda_2$ for $l=N+1,\ldots, N+n+1$, where κ_1, κ_2, λ_1, and λ_2 are user selected values.

b. Project the series made from $P+y^{N+1}R$ into the range of F using W and the geometric row vectors made from the roots of $F(y)$. The projected polynomial will have negligible values for the coefficients between $n+1$ and $N+1$ (depending on the noise in the data) and will be divisible by $F(y)$. Further, $P(y)$ can be read off from the projection and $Q(y)$ found by simple division (using FFT methods, for instance, as discussed in [2]).

As long as: *i*) the projection is stable, *ii*) the projected vector is not zero, and *iii*) the values of the weights are large enough to produce an answer within the noise for similar cases with no data noise, then the values of κ_1, κ_2, λ_1, and λ_2 have very little effect on the calculated y_j's and A_j's. In the ordinary Prony case there are as many degrees of freedom added in R as are restricted between P and R (one degree is an irrelevant multiplicative constant since F is expressed as a ratio).

This method has been tried for several examples with both real and complex poles with maximum to minimum root amplitude ratios up to 60:1 and with minimum root separation down to 0.01. Using double precision data (with $\kappa_1=10^8$, $\kappa_2=10^4$, $\lambda_1=10^{28}$, and $\lambda_2=10^{18}$) it works well up to about eight poles after which the positions of the larger poles start to degenerate (due to the ill-conditioning of the problem). The representation of the time series coefficients of F remains accurate to a relative error of about 10^{-12}, which represents approximately the cumulative accuracy of the algorithms involved. Using data given to 16 bit precision (around 5 place accuracy), the method is able to resolve about 3 poles (where $\kappa_1=10^3$, $\kappa_2=10^2$, $\lambda_1=10^{12}$, and $\lambda_2=10^6$) with the positions of the larger poles again beginning to degenerate first.

4. Summary

The idea of root projection (RP) was introduced to permit least-squares modification of a one-sided time series allowing a set of given complex numbers to be roots of the Y transform of the time series. The general framework of the method was presented, and techniques were given to adapt the method to the Gram-Schmidt algorithm. For time and frequency least-squares weighting, a dynamically dimensioned form of the Gram-Schmidt algorithm has been developed to carry out root projection. The code has been employed for root projection on time series with up to 1600 points and achieved better than 12 place accuracy. Illustrative examples of root projection were shown for noise reduction and filter construction as well as a least-squares extension of Prony's method.

Acknowledgments

The author wishes to thank D. P. O'Leary and G. W. Stewart for many valuable numerical suggestions and comments. This research was carried out under a joint NBS/DARPA program on NDE/RSP.

5. References

[1] Simmons, J. A., "Boundary Roots and Signature Analysis of Causal and Transient Time Series," submitted to SIAM J. on App. Math.
[2] Simmons, J. A., J. Res. Natl. Inst. Stand. Technol. **96**, 345 (1991).
[3] Halmos, P. R., Finite-Dimensional Vector Spaces, Springer Verlag (1974).
[4] Dongarra, J. J., Moler, C. B., Bunch, J. R., and Stewart, G. W., LINPACK User's Guide (2nd edition), Chapter 9, SIAM, 1981.
[5] Stewart, G. W., private communication.
[6] Rutishauser, H., Numerische Mathematik **16**, 205 (1970).
[7] Simmons, J. A., unpublished (Fortran '77 code available on request).
[8] McClellan, J. H., Parks, T. W., and Rabiner, L. R., IEEE Trans. Audio Electroacoust. **AU-21**, 506 (1971).

[9] Twomey, S., Introduction to the mathematics of inversion in remote sensing and indirect measurements, Elsevier/North Holland Press (1977) pp. 172–174.

About the author: *John A. Simmons is a mathematician in the Metallurgy Division of the Materials Science and Engineering Laboratory of NIST, and has worked in the elasticity and elastodynamics of defects in solids. This work has arisen from studies of the signal processing of acoustic emission in materials.*

Volume 96, Number 3, May–June 1991
Journal of Research of the National Institute of Standards and Technology

[J. Res. Natl. Inst. Stand. Technol. 96, 345 (1991)]

Deconvolution of Acoustic Emission and Other Causal Time Series

Volume 96 | Number 3 | May–June 1991

John A. Simmons

National Institute of Standards and Technology,
Gaithersburg, MD 20899

A new technique, root projection (RP), is given for quantitative deconvolution of causal time series in the presence of moderate amounts of noise. Deconvolution is treated as a well-conditioned but underdetermined problem and *a priori* information is employed to obtain comparable noise reduction to that achieved by singular value decomposition (SVD) techniques while providing more accurate frequency information about the inverse. Two detailed examples are given. The first gives noise analysis for alternate methods for deconvolution with a Gaussian kernel. The second example presents a model acoustic emission transducer calibration problem with typical noisy and incomplete output data. This example is treated by the use of a robust cross-cutting algorithm combining both the RP and SVD methods.

Key words: acoustic emission; causal time series; cross-cut algorithm; deconvolution; FFT deconvolution; Gaussian deconvolution; root projection; signal processing; singular value decomposition; time series; transducer calibration.

Accepted: February 25, 1991

1. Introduction

Because of the almost ubiquitous presence of Green's functions in linear physical theories, deconvolution (or inverse filtering) is the most commonly encountered special inverse problem associated with many characterization problems in NDE, both electromagnetic and elastic. A difficulty with inverse problems is their susceptibility to ill-conditioning; and, as is well-known, deconvolution problems can range in difficulty from those simply solvable by forward substitution to those for which no known method holds any hope of solution.

Some of the most powerful techniques that can be applied to deconvolution are those, such as regularization and singular valued decomposition (SVD), arising from general linear inversion methods. However, there are other methods available which take advantage of the special properties of the convolution process. We shall review briefly some of these methods, compare them with an SVD method especially adapted for deconvolution and present a new approach to deconvolution resting on the concept of root projection which, when combined with SVD, offers the possibility of approaching currently intractable deconvolution problems.

Two examples—both ill-conditioned—will be presented, and signal-to-noise concepts will be introduced to sharpen the SVD and root projection deconvolution (RPD) methods. These will then be compared with more standard deconvolution methods. The first example involves convolution with a Gaussian kernel in which the entire output is given, but corrupted by varying amounts of noise. In the second example the convolution kernel is derived from the actual response of a "standard" acoustic emission transducer in a calibration experiment. As

is typical in many such experiments, the entire output is not available. So one is faced with deconvolving a one-sided kernel from an incomplete one-sided time series output. To treat this problem, the SVD and RPD methods are combined into a more robust cross-cutting algorithm which provides both superior time and frequency information about the inverse.

2. Deconvolution Methods

Full convolution of two finite-length, real-valued, one-sided time series may be expressed as

$$\boldsymbol{a}:N * \boldsymbol{b}:M = (\boldsymbol{a} * \boldsymbol{b}):(M+N), \tag{1}$$

where $\boldsymbol{a}$ is a causal time series with elements $a_0,\ldots,a_N$, $\boldsymbol{b}$ is a causal time series with elements $b_0,\ldots,b_M$, and where $\boldsymbol{a}*\boldsymbol{b}$ is a causal time series with elements[1]

$$(\boldsymbol{a}*\boldsymbol{b})_n = \Sigma_k a_k\, b_{n-k},\ n = 0,\cdots,M+N. \tag{2}$$

The fact that $\boldsymbol{a}:\boldsymbol{b}$ or $\boldsymbol{b}:M$ are truncated at finite length makes no difference to the first $\min(M,N)$ terms in the output.

In the deconvolution problem one is given the series $\boldsymbol{a}:N$ and some part of the right hand series in eq (1), usually M terms, which is possibly corrupted by noise. For most applications one assumes that $M=N$, if need be by zero padding the shorter truncated series, and one writes the deconvolution problem in the form

$$\boldsymbol{a}:N * \boldsymbol{b}:N = \boldsymbol{v}:N \tag{3}$$

where $\boldsymbol{a}:N$ and $\boldsymbol{v}:N$ are assumed known and $\boldsymbol{\beta}:N$ is to be reconstructed as closely as possible. The series $\boldsymbol{a}:N$ is called the kernel series. In Y transform terminology eq (1) may be paraphrased as

$$a(y)\beta(y) = v(y) + y^{N+1}u(y), \tag{4}$$

where $\boldsymbol{u}:(N-1)$ is usually an unknown time series.

In the notation used in eq (1) with $N<M$, if $\boldsymbol{a}:N$ is thought of as an Nth order difference operator, then eq (1) becomes a high order difference equation with the first K terms associated with the boundary conditions at the causal boundary and the last $N-K-1$ terms associated with the boundary conditions at the transient boundary. Treating eq (3), then, has the interpretation of solving a one point boundary value problem, while eq (4) may be thought of as solving the two point boundary value problem where the influence of the transient boundary conditions on the solution may be studied and these conditions possibly restricted using *a priori* information.

In Y transform terms the inversion of eq (4) merely amounts to finding the coefficients of the Taylor's series of $\left(v(y)+u(y)y^{N+1}\right)/a(y)$ about the origin, and the first $N+1$ of these coefficients are the same as those of $v(y)/a(y)$ so long as $a_0 \neq 0$. From this point of view the expansion of $v(y)/a(y)$ about the origin is only valid within the radius of convergence of the Taylor's expansion. Since $v(y)$ is a truncated series and has noise, its zeroes will not include those of $a(y)$, so that this expansion is generally only valid to the innermost zero of $a(y)$.

Let the modulus of the smallest zero of $a(y)$ be r_0. If we use a circle of radius $r<r_0$, then we can write the Cauchy formula for the Taylor's expansion coefficients as

$$\beta_n = \frac{1}{2\pi i}\oint_{|y|=r} \frac{v(y)}{a(y)}\,dy. \tag{5}$$

Approximating this integral by numerical evaluation on r times the Mth roots of unity gives:

$$\beta_n = \frac{r^{-n}}{M}\sum_{k=0}^{M-1} \frac{v}{a}\left(re^{\frac{2\pi k i}{M}}\right)e^{-\frac{2\pi k i}{M}}. \tag{6}$$

This formula is exact when $0 \leq n \leq N < M$ if $\boldsymbol{v}:N = (\boldsymbol{a}*\boldsymbol{\beta}):N$. Otherwise, all coefficients $\beta_{n'}$, $n' \equiv m \pmod M$ of the Laurent expansion of $v(y)/a(y)$ will be aliased onto β_n.

If one defines the "radiused" time series $\widehat{\boldsymbol{ar}}:N$, where $(\widehat{ar})_n = a_n r^n$, $n=0,\cdots,N$, then eq (6) is merely a DFT inversion formula on the quotient of the DFT's of the radiused time series $\widehat{\boldsymbol{vr}}:N$ by $\widehat{\boldsymbol{ar}}:N$ where the nth term of the inverted series is multiplied by the expansion factor of r^{-n}. We refer to this technique implemented with the FFT algorithm as radiused FFT inversion. The factor r^{-n} in eq (6) expands noise exponentially rendering this algorithm generally useless in its raw form. The algorithm can, however, be employed in conjunction with root projection. That application will be discussed in detail shortly.

If one attempts to employ eq (6) for $r<r_0$ to reduce the noise buildup, one obtains the expansion of the wrong Laurent series form of v/a, a form

[1] We use here the modified Dirac bracket and Y transform notations employed in references [1] and [2].

yielding a non-causal time series whose terms depend on the residues of $v(y)$ at the roots of $a(y)$ inside r. In moving r from the inside to the outside of any one of these roots one changes the contribution due to that root from a causal time series with a factor of the residue times r^{-n} into an anticausal time series with terms of the type residue times r^{n-1}. In a DFT context these terms also wrap around through negative times to produce time aliasing, especially for large positive times. The rapidity of convergence of the partial fraction series associated with a given root y_α depends on that ratio y_α/r or r/y_α which is less than one.[2]

In matrix notation the deconvolution problem is almost exclusively formulated in the parallel form of eq (3):

$$\begin{bmatrix} a_0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ a_1 & a_0 & \cdot & & & \cdot & \cdot \\ \cdot & a_1 & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & & & \cdot & 0 & \cdot \\ a_{N-1} & \cdot & \cdot & \cdot & \cdot & a_0 & 0 \\ a_N & a_{N-1} & \cdot & \cdot & \cdot & a_1 & a_0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_{N-1} \\ \beta_N \end{bmatrix} = \begin{bmatrix} v_0 \\ v_1 \\ \cdot \\ \cdot \\ \cdot \\ v_{N-1} \\ v_N \end{bmatrix} \tag{7}$$

while eq (4) takes the form:

$$\begin{bmatrix} a_0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ a_1 & a_0 & \cdot & & & \cdot & \cdot \\ \cdot & a_1 & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & & & \cdot & 0 & \cdot \\ a_{N-1} & \cdot & \cdot & \cdot & \cdot & a_0 & 0 \\ a_N & a_{N-1} & \cdot & \cdot & \cdot & a_1 & a_0 \\ 0 & a_N & \cdot & \cdot & \cdot & \cdot & a_1 \\ 0 & 0 & \cdot & & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot & a_{N-1} & \cdot \\ \cdot & \cdot & & & \cdot & a_N & a_{N-1} \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & a_N \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_{N-1} \\ \beta_N \end{bmatrix} = \begin{bmatrix} v_0 \\ v_1 \\ \cdot \\ \cdot \\ \cdot \\ v_{N-1} \\ v_N \\ u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_N \end{bmatrix} \tag{8}$$

Equation (7), while completely determined, is generally ill-conditioned. Equation (8), on the other hand, while underdetermined, is generally well-conditioned and offers new algorithms for dealing with the deconvolution problem.

[2] In the ideal continuous case there is no truncation or aliasing problem associated with $v(y)$, only the noise problem. In that case the Fourier transform for $\beta(y)$ should be entire in the upper half plane, and the exponential noise buildup and non-causal representation arguments against the continuous Stokes' deconvolution formula for dividing the appropriate Fourier transforms are still valid [3].

The simplest method for solving eq (7) is by forward substitution, which is equivalent to formal division of the Y transforms $v(y)/a(y)$ starting at the constant term. This is generally the most unstable of inversion methods. The matrix equivalent of zero-padded and radiused FFT inversion is obtained by considering inverse-radiused circulant matrices.

From a real-valued time series $c{:}\mathcal{N}$ we build the $(\mathcal{N}+1)\times(\mathcal{N}+1)$ inverse-radiused circulant matrix C_r defined by:

$$(C_r)_{kl} = r^{l-k} c_{(\mathrm{mod}(k-l,\mathcal{N}+1))}, k,l = 0,\cdots,\mathcal{N}. \tag{9}$$

It is then a simple matter to show that the vectors

$$\widehat{r^{-1}e^{\frac{-2\pi mi}{\mathcal{N}+1}}}{:}\mathcal{N} = \begin{bmatrix} 1 \\ r^{-1}e^{\frac{-2\pi mi}{\mathcal{N}+1}} \\ r^{-2}e^{\frac{-4\pi mi}{\mathcal{N}+1}} \\ \cdot \\ \cdot \\ \cdot \\ r^{-\mathcal{N}}e^{\frac{-2\mathcal{N}\pi mi}{\mathcal{N}+1}} \end{bmatrix} \tag{10}$$

are eigenvectors for C_r. That is:

$$[C_r]\widehat{r^{-1}e^{\frac{-2\pi mi}{\mathcal{N}+1}}}{:}\mathcal{N} = \left[\left(\widehat{e^{\frac{2\pi mi}{\mathcal{N}+1}}}{:}\mathcal{N}\right)^T c{:}\mathcal{N}\right]\widehat{r^{-1}e^{\frac{-2\pi mi}{\mathcal{N}+1}}}{:}\mathcal{N}$$

$$= \tilde{c}_m\left(\widehat{r^{-1}e^{\frac{-2\pi mi}{\mathcal{N}+1}}}{:}\mathcal{N}\right), \tag{11}$$

where $\tilde{c}{:}\mathcal{N}$ is the $\mathcal{N}+1$ point DFT of $c{:}\mathcal{N}$. In addition, the radiused DFT provides a decomposition of the discrete δ function:

$$\left(\widehat{re^{\frac{2\pi mi}{\mathcal{N}+1}}}{:}\mathcal{N}\right)^T \widehat{r^{-1}e^{\frac{-2\pi mi}{\mathcal{N}+1}}}{:}\mathcal{N} = (\mathcal{N}+1)\delta_{mn}. \tag{12}$$

Then, the inversion of the equation:

$$C_r\,\beta{:}\mathcal{N} = v{:}\mathcal{N} \tag{13}$$

is easily carried out by eigenvector decomposition:

$$\beta{:}\mathcal{N} = \frac{1}{\mathcal{N}+1}\sum_{m=0}^{\mathcal{N}} \frac{\left[\left(\widehat{re^{\frac{2\pi mi}{\mathcal{N}+1}}}{:}\mathcal{N}\right)^T v{:}\mathcal{N}\right]}{\tilde{c}_m}\left(\widehat{r^{-1}e^{\frac{-2\pi mi}{\mathcal{N}+1}}}{:}\mathcal{N}\right). \tag{14}$$

To relate $c{:}\mathcal{N}$ to $a{:}\mathcal{N}$, we set $c_k = a_k r^k$, $k = 0,\ldots,N$, $c_k = 0$, $k = N+1,\ldots,\mathcal{N}$, which is the radiused form of the zero-padded kernel series. Then the inverse-radiused matrix C_r has the form shown in eq (15) of a

$$\begin{bmatrix}
a_0 & 0 & 0 & . & . & . & . & . & 0 & 0 & a_N r^{(N+1)} & . & . & . & a_2 r^{(N+1)} & a_1 r^{(N+1)} \\
a_1 & a_0 & 0 & 0 & . & . & . & . & . & 0 & 0 & . & & & . & a_2 r^{(N+1)} \\
. & a_1 & . & . & . & & & & & . & 0 & . & . & & . & . \\
. & . & . & . & . & . & & & & . & . & . & . & . & . & \\
. & . & & . & . & . & . & & & . & . & & . & . & a_N r^{(N+1)} & . \\
a_{N-1} & . & & & . & . & . & . & & . & . & & & . & 0 & a_N r^{(N+1)} \\
a_N & a_{N-1} & . & . & . & a_1 & a_0 & 0 & 0 & . & . & & & & 0 & 0 \\
0 & a_N & . & . & . & . & a_1 & a_0 & 0 & 0 & . & & & & . & 0 \\
. & 0 & . & . & & & & . & . & 0 & 0 & & & & . & . \\
. & . & . & . & . & & & & . & a_0 & 0 & . & & & . & . \\
. & . & & . & . & . & & & & a_1 & a_0 & . & . & & . & . \\
. & . & & & . & . & . & & & . & a_1 & . & . & . & . & . \\
. & . & & & & . & . & . & & . & . & . & . & . & 0 & . \\
. & . & & & & & . & . & . & . & . & & . & . & 0 & 0 \\
0 & 0 & . & . & . & . & . & 0 & a_N & a_{N-1} & . & . & . & . & a_0 & 0 \\
0 & 0 & . & . & . & . & . & 0 & 0 & a_N & a_{N-1} & . & . & . & a_1 & a_0
\end{bmatrix} \tag{15}$$

zero-padded convolution matrix with an extra exponentially damped upper right hand Toeplitz corner. Thus radiused FFT inversion solves the "damped" circulant rather than the convolution matrix problem. As seen from eqs (12) and (14), the answer obtained from this method will be correct for the convolution problem if $r^{(N+1)}$ times the last N terms of the answer are effectively zero. That happy circumstance is usually frustrated by the exponential buildup in the eigenvectors of eq (10) together with the presence of noise components in $\boldsymbol{v}{:}N$, as discussed previously in the Y transform context. However, referring to that discussion, one can conclude that if $r<1$ and if there are no zeroes of $a(y)$ with modulus less than r, then with sufficient zero padding FFT inversion will approach convolution inversion.

Least squares techniques are the standard way to deal with noise problems. Frequently they are formulated in terms of minimizing the norm $\|\boldsymbol{v}{:}N - \mathrm{A}\,\boldsymbol{\beta}{:}N\|$. However, since $\boldsymbol{v}{:}N$ contains signal-plus-noise, applying this condition too stringently forces induced domain noise into $\boldsymbol{\beta}{:}N$. What is required is a technique to separate the range signal from the range noise either by applying *a priori* information about the signal in the range or *a priori* information about the signal in the domain. In a simple sense the example given in reference [2] giving rise to figure 4 accomplished this in the range using root projection when the complete information was available in the domain. We shall extend that concept below.

Another way to deal with this difficulty is through regularization, in which the norm is modified by adding another norm—usually a scalar times a domain norm—onto the above range norm to adjust the range residuals to be of the order of magnitude expected from the noise. The question is: Which norm should one add? The techniques for choosing an appropriate norm and carrying out this process efficiently lie somewhat outside the framework of this work, so that we shall not provide detailed comparisons with regularization methods here.

For eq (7) the least-squares method is usually implemented by multiplying the equation by $A^{\diamond} = (A^{T})^{*}$ to give the equation

$$A^{\diamond} A\ \boldsymbol{\beta}{:}N = A^{\diamond}\ \boldsymbol{v}{:}N \tag{16}$$

where $A^{\diamond}A$ is a non-negative Hermitian matrix. The multiplication accomplishes two results: *i*) It forces the output vector into the range of A, thereby eliminating some of the components of the noise, and *ii*) It modifies the left hand matrix to a Hermitian form allowing eigenvector decomposition methods and algorithms to be employed. However, this approach suffers from two drawbacks: *i*) If A has a simple structure—e.g., A is a convolution operator—efficient algorithms which take advantage of this structure may be lost, and *ii*) The eigenvalues of $A^{\diamond}A$ are effectively the squares of those of A (in a meaning to be described directly), thus greatly increasing the ill-conditioned nature of the problem.

A closely related approach to the above is the singular value decomposition technique (SVD). In the SVD method one need not multiply the right hand side of eq (7) by $A^{\diamond}$. Rather, one represents A in the form

$$A = U^{\diamond} D V \tag{17}$$

where U and V are unitary matrices and D is a diagonal matrix whose diagonal elements are called the singular values of A. The SVD method, while usually slower to compute than the method arising from eq (15) has the advantage of finding diagonal elements which are the square roots of those stemming from $A^{\diamond}A$. This permits greater dynamic range for treating very ill-conditioned problems such as frequently occur when solving eq (7). For the deconvolution problem the SVD algorithm has been implemented using an efficient form of another algorithm due to Lanczos (in addition to the SVD decomposition) to permit a small dimensional approximation to the full SVD approach [4]. However, this Lanczos/SVD approach, while more efficient may not permit accurate inversion especially when N is much greater than the dimension of the Lanczos subspace.

We turn now to least squares algorithms associated with eq (8). Calling $A2$ the $(2N+1\times N+1)$ matrix in that equation, one can say that $A2$ describes an embedding of the generally $N+1$ dimensional range of A in $2N+1$ dimensional space. In terms of the Euclidean inner products for both $\mathbb{C}^{2N+1}$ and $\mathbb{C}^{N+1}$ given by $(\boldsymbol{u},\boldsymbol{v})=\boldsymbol{u}^T\boldsymbol{v}^*$, the adjoint of $A2$, $A2^{\dagger}$ is $A2^{\dagger}=A2^{\diamond}=A2^T$, since $A2$ is real [5]. Given a vector $\boldsymbol{c}:2N$ we can use $A2^{\dagger}$ to construct a least squares projection to map $\boldsymbol{c}:2N$ into the range of $A2$. Once in the range, we shall see that inversion is well-conditioned and simple.

The necessary and sufficient condition that $\boldsymbol{c}:2N$ lie in the range of $A2$ is that $\boldsymbol{c}:2N$ lie in the orthogonal complement of the null space of $A2^{\dagger}$ [5]. That is:

$$\text{If } A2^{\dagger}\boldsymbol{b}:2N=0, \text{ then } (\boldsymbol{b}:2N)^T\overline{\boldsymbol{c}:2N}=0. \tag{18}$$

To find a basis for the null space of $A2^{\dagger}$, we consider the geometric sequence vector $\widehat{y}$ where y is any complex number. These vectors are eigenvectors of $A2^{\dagger}$ satisfying

$$A2^{\dagger}\widehat{y_\alpha}:2N=(\widehat{y_\alpha}:2N)^TA2=a(y)\widehat{y_\alpha}:2N. \tag{19}$$

Thus, if $a(y)$ has distinct roots, we can identify the N dimensional null space of $A2^{\dagger}$ as spanned by

$$\widehat{y_\alpha}, \text{ where } a(y_\alpha)=0,\ \alpha=1,\cdots,N. \tag{20}$$

The condition 17 then becomes the relation that y_α are among the roots of $c(y)$, or in algebraic terms $a(y)$ divides $c(y)$. From this point of view the existence of CBR implies a set of "continuity" equations that must be satisfied by time series in the range of $a(y)$ under polynomial multiplication.

The root projection discussed in reference [2] provides us with the required mapping into the range of $A2$ [2]. Also, since there must exist a polynomial $b(y)$ of degree N such that $a(y)b(y)=c(y)$, we can find this polynomial simply by radiused FFT division with at least $2N+1$ points using any numerically reasonable radius near one.

Thus we can formulate the root projection deconvolution (RPD) algorithm as an alternative least squares approach to deconvolution:

i) Embedding $\boldsymbol{v}:N$ in $2N+1$ dimensional space

This is usually done by replacing an *a priori* range estimate with more accurate *a posteriori* range data where available. It can be done from the domain by convolving an *a priori* domain estimate with the kernel to give the *a priori* range estimate. Since the root projection of an *a priori* range estimate in this case is equal to itself, an equivalent method here is to subtract the *a priori* range estimate from the *a posteriori* range data, where available, and fill out the rest of the range with zeroes to produce a reduced problem with a zero *a priori* estimate. The *a priori* domain estimate can then be added back on to the inversion estimate from the reduced problem to give an upgraded inversion estimate. This latter approach can also be directly applied to introduce *a priori* information into the Lanczos/SVD method.

ii) Applying root projection

Here time weighted inversion is employed using weightings on the *a priori* and *a posteriori* parts of the range data reflecting the relative uncertainties of the *a priori* versus *a posteriori* parts of the data. A very high weighting factor applied to the *a priori* part of the data will cause the root projection to return a result with very small residuals in the *a posteriori* part of the range data. This means that the inverse fits both the range signal and the superimposed noise more closely as one increases the ratio of weights of *a priori* to *a posteriori* points in the time weight file.

iii) Inverting using FFT division

The entire FFT division operation is carried out on an $M\geq 2N+1$ dimensional space. The basis for this space is the set of DFT vectors built from the $(M-1)$st roots of unity. However, only N of the M FFT components of any domain vector are independent, since there are $M-(N+1)$ conditions re-

quiring all components after $N+1$ to be zero. In the range there are still only $M-(N+1)$ FFT components independent. This time there are N conditions requiring orthogonality to the geometric root vectors of the kernel and $M-(N+1)$ conditions requiring all image space components $(2N+1)$ to be zero. Thus, the basic range projection consists of projecting through the set of geometric root vectors supplemented by any simple basis of the $M-(N+1)$ dimensional space of vectors whose first $2N+1$ components are zero using the time (or frequency) weighted norm. The basic domain projection is through the $M-(N+1)$ dimensional space whose first $N+1$ components are zero. An independent orthogonal base for the range space is never built in root projection.

Few practical deconvolution problems can end at this point. Because of the presence of range noise, any inversion process will induce a domain noise which frequently obscures some or all of the features of the inverse. The critical problem when carrying out deconvolution consists of separating those aspects of the inverse which are "unequivocally" determined by the inversion process from those which are not. From the SVD or RPD point of view this means finding the signal-to-noise ratios in a series of orthogonal channels, accepting those aspects of the signal in the channels where the ratio is high and trying to supply *a priori* information in those where it is not. The SVD method provides an excellent means for optimal signal-to-noise separation. RPD, which works with frequency information, often provides an excellent representation for testing and including *a priori* information. One can add or remove various hypothetical features from an *a priori* estimate and test the consequence of these changes. Since with adequate weighting root projection always forces the range residuals back to within the noise limits, the result obtained will be a possible solution. However, essential solution features will be added and inessential ones taken away will not return.

Both RPD and SVD offer the possibility for determining the relative probability of a particular deconvolution estimate. However, we shall not discuss statistical algorithms in any systematic way here. Rather, we shall present two numerical examples principally to compare the accuracy and flexibility of the SVD, RPD, and radiused FFT division algorithms as inversion techniques and to present a cross-cut method for deconvolution combining Lanczos/SVD inversion (and filtering) with RPD and frequency space filtering.

In the first case discussed the noise we add is produced by rounding off the range data to a certain degree of accuracy, thus simulating an A/D conversion process. This noise while crudely uniform in frequency amplitude is not independent of the output, since one expects a negative correlation with the output derivative. In the second case the noise added was produced by a random number generator. Error will be measured by using a relative standard deviation. The ordinary standard deviation of one file from another will be divided by the generalized geometric mean of the non-zero terms of the reference file to give an order of magnitude correction to the error without being overly sensitive to "spikiness" in the reference file. Errors will only be compared for the same reference file.

3. Examples

3.1 Example I

We return to the example arising from the convolution of the 101 point Gaussian of figure 1 in reference [2] with the first 700 points of the experimental waveform of figure 2.4a in reference [1]. This example has the form of eq (1) with full range information; only noise has been added. This example illustrates one way of looking at $M+N$ dimensional deconvolution in terms of root filtering. The 8 bit noise added on in that example was an "arbitrary" element of $M+N+1$ dimensional space.

The projection condition placing the signal in the range of the kernel reduced the "power" of the noise by one eighth in this case. Another 700 degrees of freedom in the noise are connected to the roots of the Y transform of the answer, $\beta(y)$. Although usually impractical, each piece of *a priori* information describing a root of $\beta(y)$ removes one degree of freedom from the noise. When the roots of both the answer and the kernel are known, only the magnitude of the answer is undetermined. This is reflected in the fact that the only degree of freedom left for the noise is associated with the one-dimensional space generated by the convolution of the kernel and the answer, any element of which is orthogonal to the geometric series root vectors from both the kernel and the answer.

Two levels of range noise will be imposed for this example. We first consider the case where only 24 bit noise is added to the output, thus reducing the output to "single precision" accuracy.

Even with 24 bit accuracy the use of the forward substitution (or "real time deconvolution") method

for this problem produces errors in excess of 400% by term 100. The errors increase exponentially after that point.

Because most of the roots of the Gaussian lie on the unit circle, we carry out our analysis on a circle inside the unit circle. We choose a radius of 0.997 as reasonable for dealing with a series of length 700 and a zero-padded FFT of length 8192. Only two roots of the Y transform of the Gaussian lie inside this radius and the value of the Y transform of the noise at these roots is quite small so that in the case of 24 bit accuracy even radiused FFT inversion offers a good estimate of the answer if adequate zero padding is provided.

Figure 1 shows the range noise spectrum versus the spectrum of the range signal-plus-noise after root projection on the circle of radius 0.997.

Since the projected output lies in the range of the Gaussian under convolution, FFT division on a space of dimension 800 or greater is now accurate and each of the FFT components can be considered as lying in an orthogonal channel.[3]

Using the principle that throwing out a channel throws out both the signal and the noise in that channel, we have a simple criterion that we should reject all channels where the signal/noise ratio is less than one or, crudely, where the (signal-plus-noise)/noise ratio is less than $\sqrt{2}$. In this case that

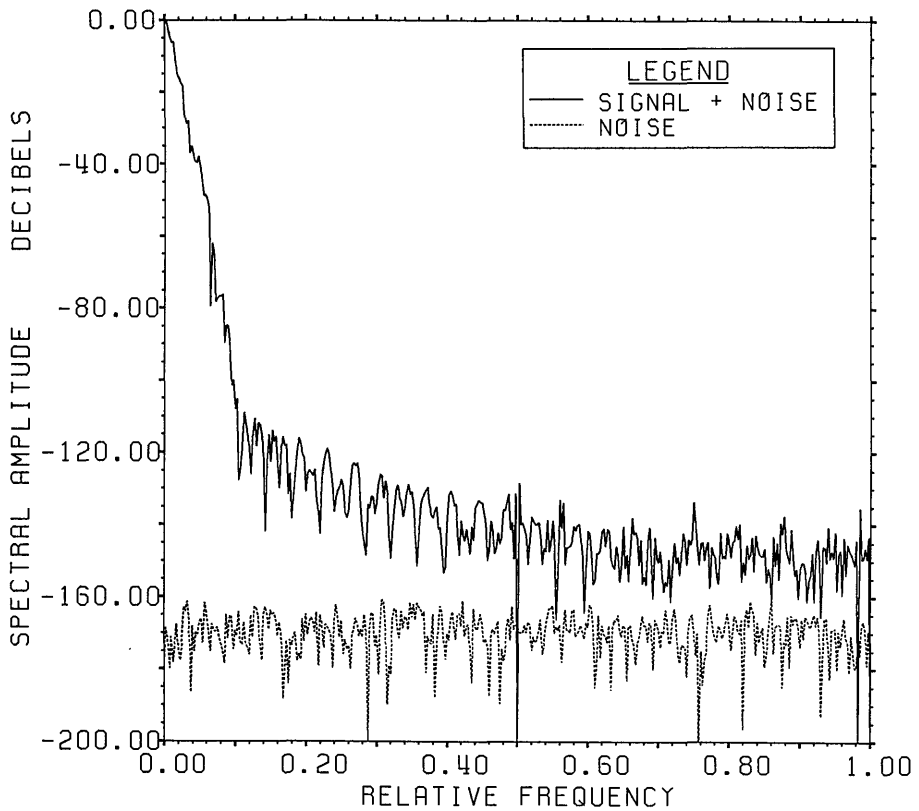

Figure 1. 800 point FFT on the circle of radius 0.997 of the range signal-plus-noise and noise for the convolution of the time series in figures 2.4a[1] and 1[2] with 24 bit roundoff noise added (Example Ia).

[3] The lack of independence of the FFT components in the image space means that projecting out some components will produce inverses no longer zero beyond point #700, thus no longer satisfying the multiplication of FFT's rule of the space of dimension M. Signal-to-noise analysis, then, is not rigorously correct in RPD on a channel by channel basis. To avoid this complication, filtering for RPD will be carried out in the domain space after the inversion. This change of dimension before filtering implies *a priori* information about the carry-over of "relevant" frequency features between FFT representations of different cardinality.

means essentially every channel is acceptable so that no filtering of the inverse should be necessary.

Figure 2 shows a comparison of the errors of the estimated inverse versus correct answer for RPD, radiused FFT inversion, and SVD inversion using a subspace of dimension 100. As can be seen, the RPD error is the smallest, merging with the FFT results for higher frequencies. The spectrum of the RPD error is, of course, exactly the quotient of the spectra of the range noise with that of the Gaussian kernel (padded to 800 points which was the FFT dimension used for RPD). The significantly larger error associated with the best Lanczos/SVD estimate arises from the use of only a 100 dimensional Lanczos subspace. The algorithm chosen uses the signal-plus-noise range vector as one of elements of the Lanczos subspace, but i)The range basis obtained may not be adequate to expand the noise or ii) The range basis may not separate the signal and the noise well, iii) The domain basis may not be adequate to expand the answer accurately, and iv) The matrix used to represent the convolution matrix has to be truncated at the last row perturbing it from its correct form on the whole space.

In this case the error appears to stem equally from the last two of these causes. The relative standard deviation (r.s.d.) from the answer caused by inadequate domain basis is 1.6% while the r.s.d. for the best SVD estimate is 2.2%. Tne r.s.d. in expanding the noise is negligible. The two errors stand in the ratio about $\sqrt{2}$ which would be expected from equal independent contributions. By comparison the r.s.d. of the error of the RPD estimate is 0.16% while that of the FFT estimate is 0.25%. The relatively large error in the Lanczos/SVD estimate for this example remains no matter how small the roundoff error; only when the inversion error becomes small relative to the induced domain noise does the Lanczos/SVD method become effective. The SVD inversion estimate is shown in figure 3; the FFT and RPD estimates are visually identical to the correct answer.

The ill-conditioned nature of the deconvolution process becomes apparent when the output has more noise. A comparison of range signal-plus-noise to noise after RPD is given in figure 4 for the

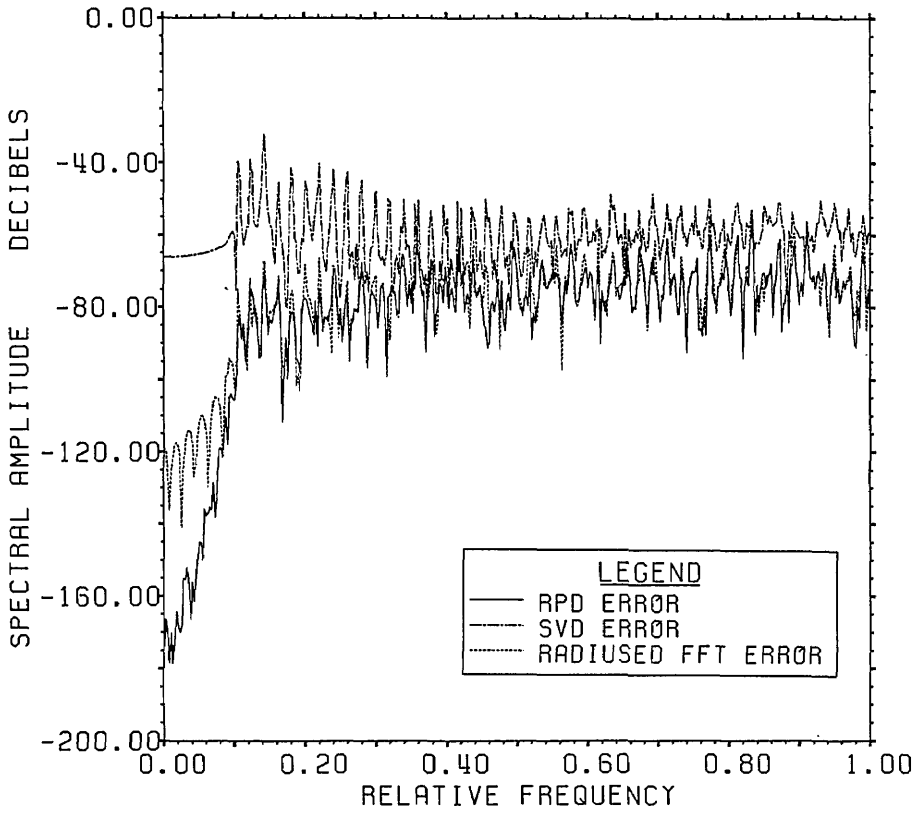

Figure 2. Comparison of spectra of errors for 3 deconvolution methods for Example Ia. An 800 point FFT was used on the circle of radius 0.997.

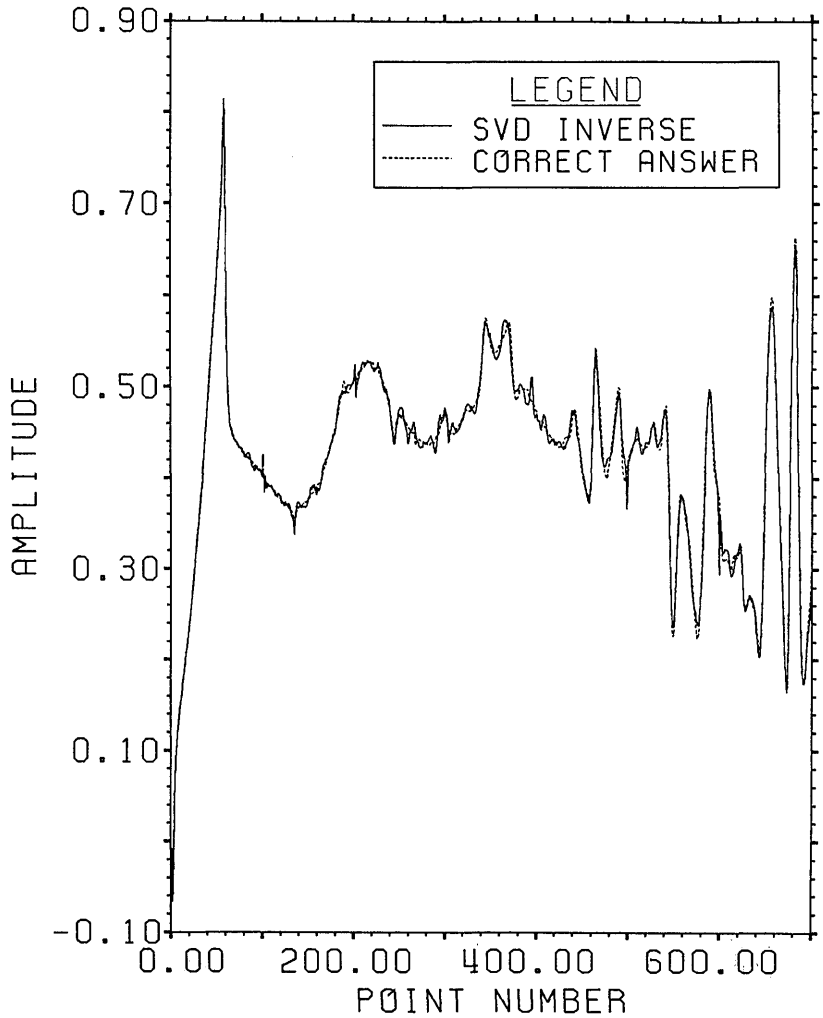

Figure 3. Example Ia: Overplot of the SVD inversion estimate and the correct answer.

case where the output has been truncated to 8 bit accuracy. The same figure made before root projection would appear visually to be substantially the same. The noise level is lowered about 0.6 dB by RPD. These spectra were also carried out on the circle of radius 0.997. Because the time curves were not renormalized after radiusing, the average noise level reads slightly below the -64 dB expected level for unradiused noise. As can be seen the (signal-plus-noise/noise) ratio is clearly greater than $\sqrt{2}$ until about 0.065 of the frequency band, dropping to less than one beyond 0.1.

Figure 5 shows the comparison of the errors for the estimated inverses from radiused FFT inversion (again padded to 8192 points) and inversion following RPD. In this case radiused FFT inversion produces incorrect answers, not only in the estimate for the inverse, but in the range upon reconvolving the "inverse" with the kernel. The values of the Y transform of the noise at the two roots lying inside the radius 0.997 have now greatly perturbed the radiused FFT division result.

The errors for the inverse estimated by RPD are determined by the ratio of the 800 point FFT of the root projected noise to the 800 point FFT of the zero-padded Gaussian. One sees that root projection has preserved the low-frequency features of the signal which are contained in the channels of high signal/noise ratio.

Even without knowing the true noise distribution we can estimate the meaningful channels and build a time domain approximation from them. Figure 6 shows the complete spectrum of the root projected inverse. We can use obvious *a priori* information that all channels beyond 0.08 of the frequency range, where the steep rise begins, are dominated by noise. We use an 801 point optimal filter with a pass band of 0.0 to 0.065 and stopband of 0.08 to

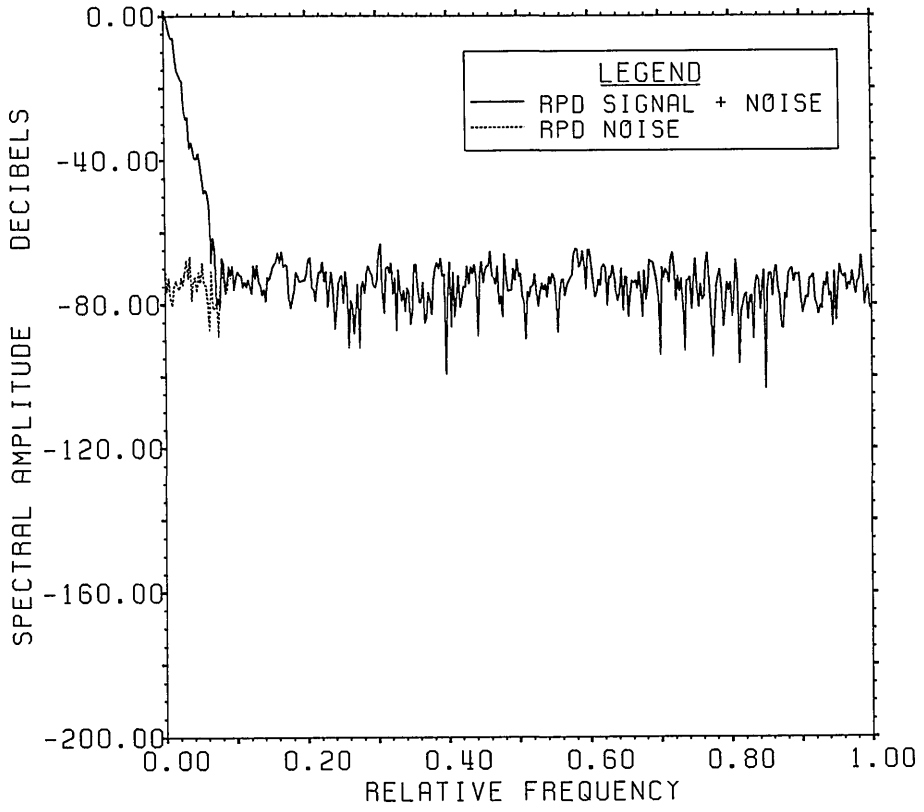

Figure 4. 800 point FFT on the circle of radius 0.997 of the range signal-plus-noise and noise after RPD for the convolution of the time series in figures 2.4a [1] and 1 [2] with 8 bit roundoff noise added (Example Ib).

1.0 with attenuation of 90 dB, convolve this with the complete RPD inverse and "trim the ends" to select out the 700 points from point 401 to point 1100. Figure 7 gives the graph of that filtered estimate from RPD. It has an overall r.s.d. of 10.5%, mostly due to errors near the front and back. Throughout much of the first 400 points, the r.s.d. is less than 3%.

A signal-to-noise analysis similar to that for RPD can be carried out for SVD inversion. In this case the channels in which one carries out the analysis are no longer associated with the familiar frequency eigenvectors but are built by the SVD process and are peculiar to the kernel series. However, one has the same number of range dimensions as domain dimensions which makes analysis of the filtering process simpler in SVD than in RPD.

Figure 8 presents the actual signal and noise spectra for the Lanczos/SVD representation in the case of 8 bit range data accuracy. The signal generally falls with decreasing singular values while the noise remains relatively constant at the expected value of -64 dB except for a rise for small singular values (large dimension number). The fraction of the noise which can be represented by only 100 of the 800 range eigenvectors is $(799.995/800)^{\frac{1}{2}}$ of the total noise. This highlights one of the features associated with the Lanczos/SVD method, notably the linkage between the range data and the SVD basis whose first element is the signal-plus-noise data vector. Since the SVD basis vectors have a special form, however, it can be difficult to obtain information on the noise statistics for detailed analysis.

The plots of signal-plus-noise and inverse SVD spectra are given in figure 9. As can be seen by comparing figures 8 and 9, it is harder to find a best guess from the available data shown in figure 9 than it was for root projection. The technique of truncating above a singular value, which leads to a residual closest to the expected noise residual, is completely in error here. It gives rise to more than 70 terms in the expansion and produces an estimate full of noise with about 1000% r.s.d. from the correct answer. However, based on an expected signal-plus-noise cutoff value of about -61 dB, a reasonable guess might be to include all compo-

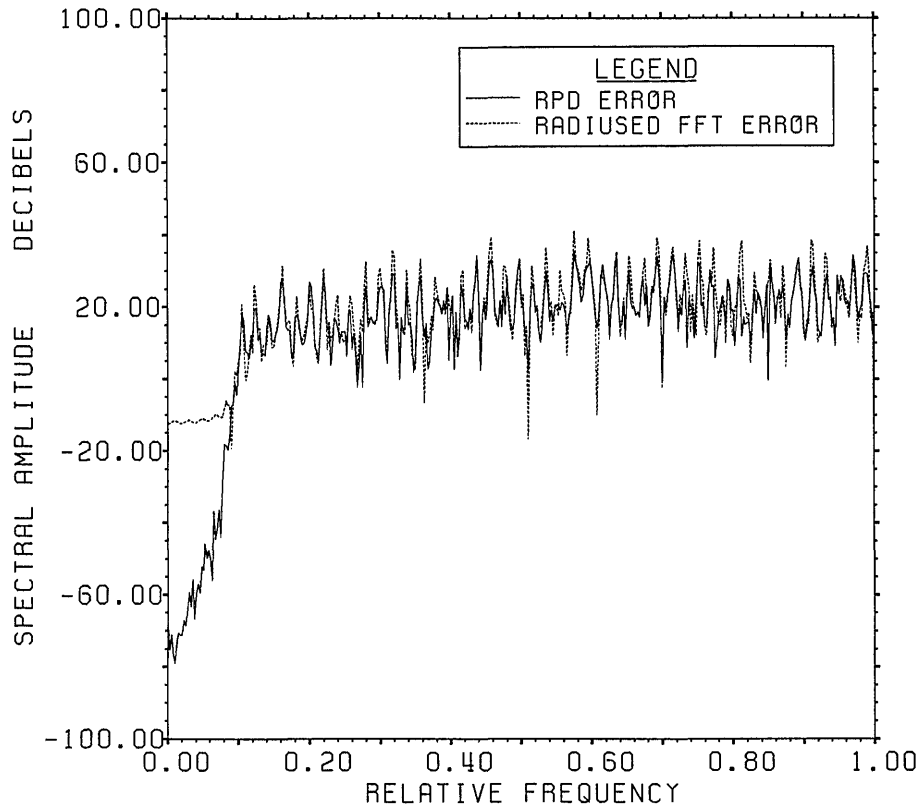

Figure 5. Comparison of errors for root projection and radiused FFT inversion for Example Ib. An 800 point FFT was used on the circle of radius 0.997.

nents from dimension 1 through 59.[4] The use of this "reasonable" guess produces the estimate shown in figure 10. The result is remarkably similar to that obtained by filtering of the root projection inverse. It has an overall r.s.d. of 10.4%.

3.2 Example II

The second example concerns deconvolution in the presence of incomplete range information. The kernel for this example is the experimental curve shown in figure 2.4a of reference [1] which was the answer in the previous example [1]. This time only 400 output data points are given as shown in figure 11. The error added was −40 dB random noise, which produces an r.s.d. of about −41 dB. The noise is clearly visible in the figure. The answer, shown in figure 12, represents a possible transducer response function. The output data then represents the response of the transducer to the same stimulus that produced the kernel response of figure 2.4a in reference [1] for the calibrating capacitive transducer.

As would be expected in a problem with a rising transient front end, the forward substitution approach is inapplicable, producing relative errors of 5000 on the third point and overflowing soon thereafter. The result of applying FFT division using an 800 point FFT is shown in figure 13. The range r.s.d. for this curve is 158% indicating a completely erroneous result even though the correct answer has less than 200 non-zero points and 400 points were given. If zero padding is added to the FFT (8192 point FFT) the range r.s.d. drops to 44% rather than the 0.8% needed to be within the expected range noise error. The domain r.s.d. against which the SVD and RPD estimates will be compared below is 1976%. Noting that only two roots of the kernel lie below 0.979, one can apply radiused FFT division with a radius of 0.975 to produce a

[4] One could also use a S/N cutoff at −61 dB; the resultant estimate in this case is visually identical to that chosen. However, we have found by experience that something like a Gibbs phenomena also occurs in SVD expansions so that including all components up to a given dimension number tends to give slightly less oscillations in the estimates than using a S/N cutoff.

A regularization type approach to this problem would be to use all channels for the inversion, but to replace each singular value, λ, by $\lambda+\epsilon$ where ϵ is some appropriately chosen small constant.

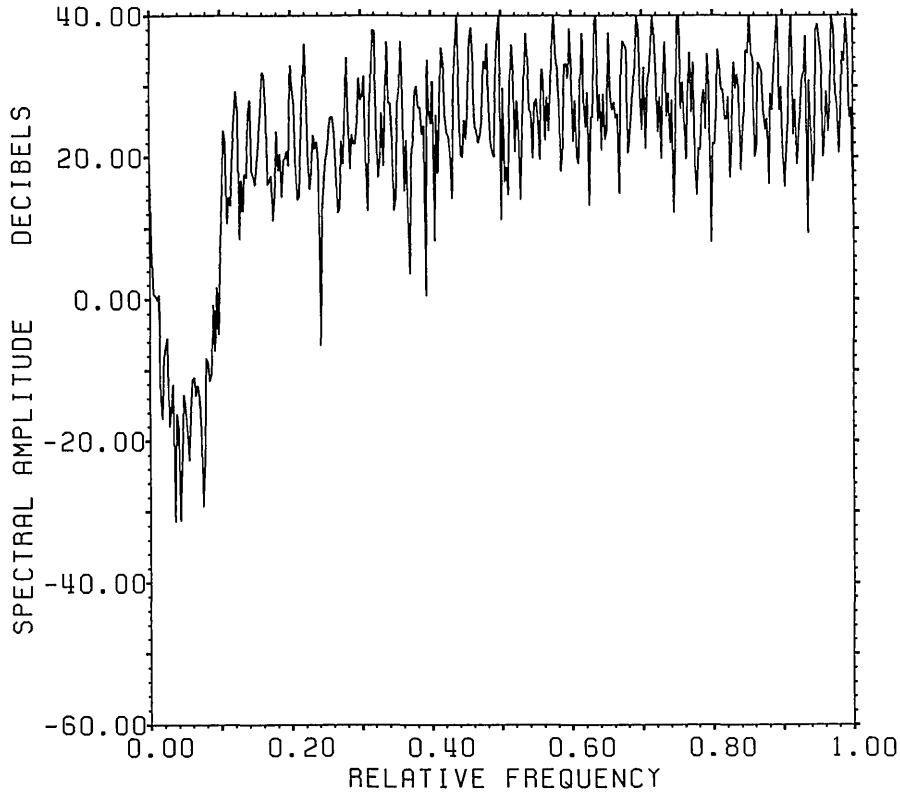

Figure 6. Spectrum from 800 pt FFT (with radius = 1.0) of estimated inverse for projection showing evident break between signal and noise dominated regions.

range r.s.d. of 1.1%. The range r.s.d. is almost within acceptable limits, but the domain errors are substantially greater than those produced by either Lanczos/SVD or RPD.

To deal with this model problem for which one has only limited *a priori* information (mostly that the answer should not be filled with high frequency noise), we introduce the cross-cutting deconvolution algorithm (CCD) shown schematically in figure 14. This algorithm uses the Lanczos/SVD and RPD methods in tandem to produce a more robust estimate which is at least as good as the particular choice of either estimate and often better than both of these methods. Lanczos/SVD and RPD with optimal filtering are first applied independently to produce first estimates to the deconvolved inverse. Typically, with a kernel whose spectrum is dominated by low frequency components, the SVD estimate will show some of the most prominent high frequency features, but will have reduced low frequency fidelity.

The RPD estimate, on the other hand, will tend to have good low frequency fidelity but reduced resolution of some of the high frequency features and possibly greater end noise. The outputs from each of these algorithms, conservatively filtered to avoid extraneous features, is then fed into the other algorithm as an *a priori* estimate. Each of these two methods uses a different orthogonal decomposition to separate signal from noise and the reuse of the output from one of these algorithms provides essentially no improvement in the estimate (although RPD may be rerun to reduce end noise). However, what one algorithm may discard as noise can contain useful signal when decomposed using the other algorithm. The second, cross-cut estimates are then frequently significant improvements upon the initial estimates. They each, of course, contain induced domain noise which may be partially independent so that the average of the estimates can be expected to provide some overall improvement over the second approximations.

Figure 15a shows the signal-plus-noise and domain spectra and figure 15b the actual signal and noise spectra upon applying Lanczos/SVD inversion to the model problem data. Based on figure 15a with an approximate signal-plus-noise value of −37 dB, reasonable choices might lie between selecting all components from 49 up to 80. As one increases the dimension number past 49, the actual

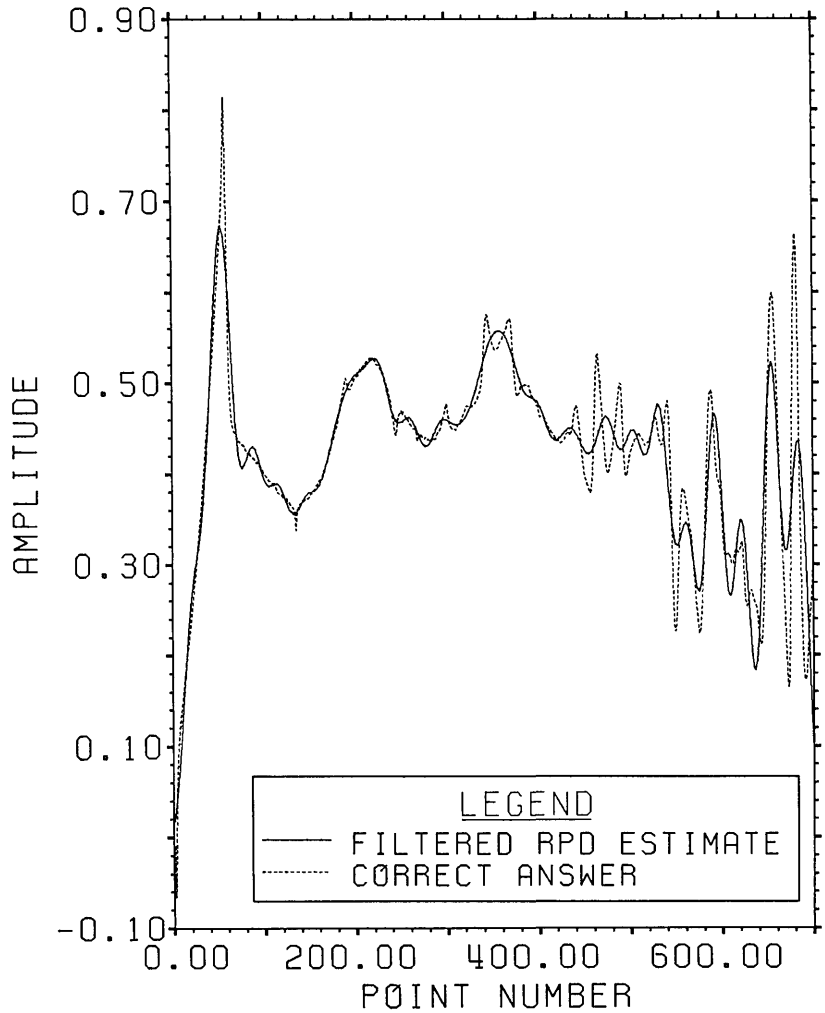

Figure 7. Filtered root projection estimate for Example Ib.

r.s.d. decreases more or less steadily (cf. fig. 15b) from 103% to 65% at dimension number 77. However, as the domain noise increases (and due also to increasing "high frequency" characteristics of the SVD base vectors) more extraneous features become added to the signal. In this case because the eigenvalues are not too small in the 49 to 80 range vis-a-vis those below 49, the "risk" is not too great of introducing a great deal of "high frequency" domain noise by inadvertently adding a channel with excess noise. The only guideline then for choosing among these candidate estimates is *a priori* information about the signal.[5] We choose the estimate made up from all dimension numbers through 62. This estimate has an overall r.s.d. of 75%. No matter which of these estimates is used as a starting point, the final output of the cross-cutting algorithm varies only between 52% and 54% r.s.d.

The RPD techniques for finding the initial approximation from incomplete range data differ somewhat from those of the SVD method. The potentially large amount of noise that can be introduced by the *a priori* part of the range estimate required by the RPD method can preclude a signal-to-noise analysis similar to that used for SVD. Instead we examine the a posteriori data separately to estimate bounds on the useful part of the frequency range and re-run RPD on windowed raw estimates to remove excessive end noise.

Figure 16a shows the FFT of the 400 points of *a posteriori* data after they have been windowed using a 400 point window built from a maximal ripple

[5] The Lanczos dimension used throughout this paper is 100. However, a special run was made for this example using Lanczos dimension 200. Although the error in representation of the answer was much better in the 200 case, the range noise was sufficiently large that only small improvement occurred in the r.s.d. of the estimates. The main improvement was in the signal-plus-noise diagram which showed a distinct plateau now making the choice of dimension number 77 as a cutoff almost apparent.

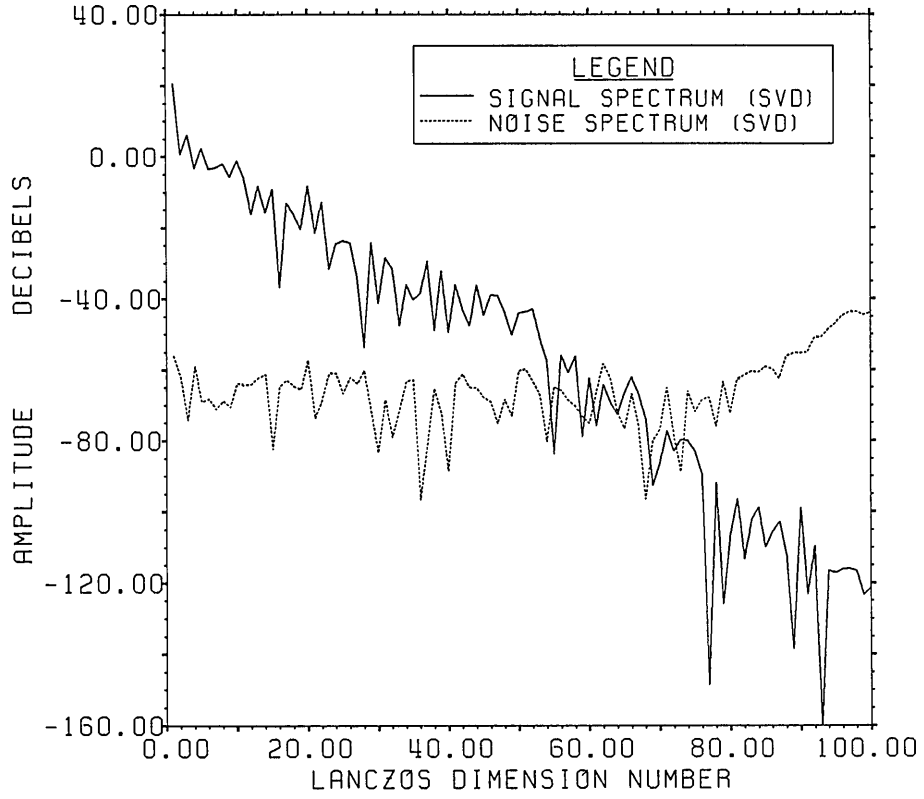

Figure 8. The true signal and noise spectra for Example Ib using the SVD basis (not the Fourier basis) on a Lanczos subspace of dimension 100.

filter to smooth the frequency distribution. The figure shows essentially no information above the −40 dB noise level beyond 30% of the Nyquist frequency. Any information about the inverse above the relative frequency of 0.3, then, must either be *a priori* information or information obtained from the SVD estimate. In other cases there will be information over the noise level through high frequency or other "difficult" parts of the frequency range where the kernel has a small spectral amplitude.

With no *a priori* information available we employ a simple extrapolation procedure to create the input range signal for the first RPD approximation. Figure 16b shows the normalized signal-plus-noise and domain spectra after RPD inversion of the given data with the endpoint values continued out to the end of the range data and then windowed. A time weighting of 1:16 (one in the given data and 16 in the extrapolated data) was used to bring the range residuals to well within the expected −41 dB r.s.d. over the initial data. When the weight factor is increased much beyond 20 in this example, exponentially increasing noise terms begin to show up in the inverse and negligible further information seems to be obtainable. In other less noisy or ill-conditioned cases the weight factor can be increased to several thousand.

Figure 16c shows the spectra of the true signal and range noise for the first RPD approximation. The range signal is merely the product of the 800 point FFT's of the kernel and the answer. The range noise, on the other hand, is not simply the −40 dB *a posteriori* noise, but is here completely dominated by the *a priori* noise which, because of the smooth extrapolation used, has very large low frequency components. Without *a priori* information, there seems to be no simple way of estimating the actual range noise which can be dominated by such errors. The initial procedure for RPD in this case, then, is to choose as broad-band a filter as possible producing an estimate which doesn't show excessive high frequency noise and to use the general *a priori* information that the inverse should not have a rising high frequency spectrum. No matter which filter is chosen, however, one can expect extensive low frequency errors. How they manifest themselves is shown in figure 17.

Figure 17 shows the SVD and RPD first estimates. As described above the overall r.s.d. for the SVD estimate is 75%. The overall r.s.d. for the

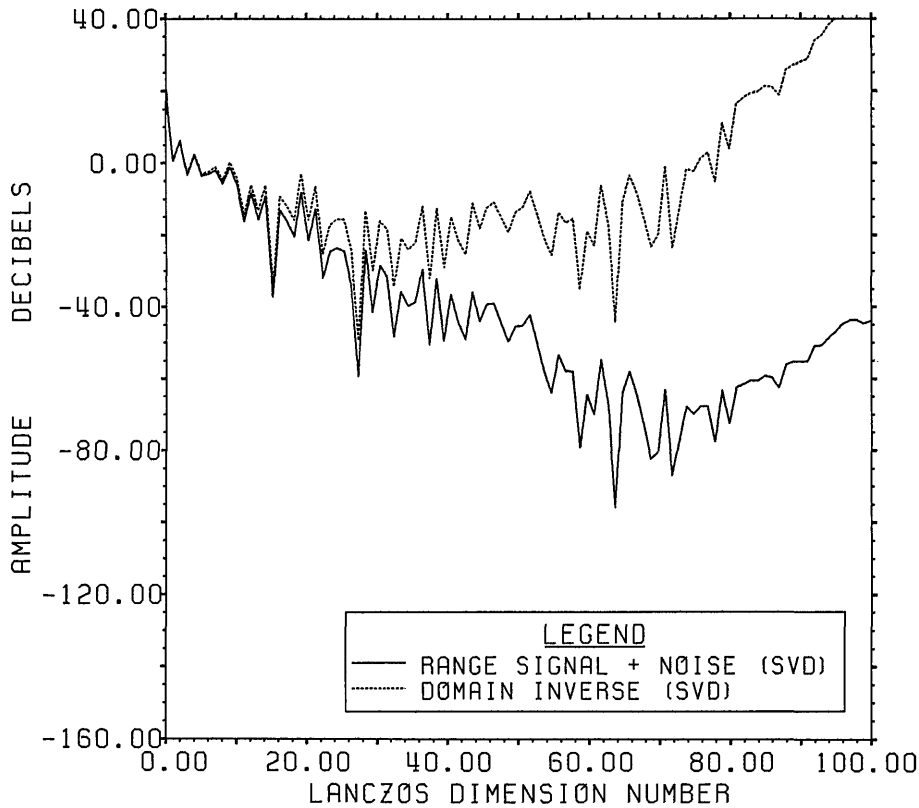

Figure 9. Range signal-plus-noise and estimated inverse spectra for Example Ib using the appropriate SVD bases built in the Lanczos subspace of dimension 100.

RPD estimate is 471%. Examination of the two curves shows this difference to be due to the much greater end error in the RPD estimate. Both of these methods tend to produce larger errors at the end of the time domain since the end values of the answer are involved in fewer terms of the *a posteriori* data. RPD, particularly without a reasonable *a priori* range guess, produces such errors due to leakage of the large *a priori* errors into the *a posteriori* part of the time domain.[6] It's easy to confirm in this case that there is almost no valid data beyond point 140 in either estimate. This is done for RPD by using the RPD estimate of figure 17 windowed to 140 points as an *a priori* estimate and reinverting. A lower time weighting ratio of 1:4 is all that's needed in this case to produce good residuals since the *a priori* range data are more accurate. After filtering the resultant estimate is almost indistinguishable from the original for the first 140 points and has only small oscillations past point 140. This indicates that the large end oscillations in the initial estimate were forced by *a priori* range errors. The overall r.s.d. of the error of this RPD estimate is 60%.

A similar process to show that there is almost no valid data beyond point 140 of the SVD estimate can also be carried out. However, there is nothing intrinsically "wrong" with the estimate produced by the RPD process which returns range residuals almost within the expected noise limits (some increase in residuals is due to filtering). It is only that the *a priori* range data chosen forced features in the answer which were irrelevant to the essential features required by the *a posteriori* data. In fact one can establish a different "model" problem in which the correct answer is the curve given as the RPD estimate in figure 17. Application of SVD to 400 points of range data from this second "model" problem with similar noise added produces estimates similar to those of the model problem as we have presented it with only slightly increased end oscillations. The overall r.s.d. of the error of the similar 62 component estimate from the answer to

[6] Errors also tend to occur very near the causal time boundary in both methods when extensive projection is employed. One can give a heuristic interpretation of this front end error for RPD based on the fact that the RPD estimate is low-pass filtered. By reducing the frequency range by a factor of M, say, lower resolution occurs in the time domain which shows up most clearly in the first M points of the time domain.

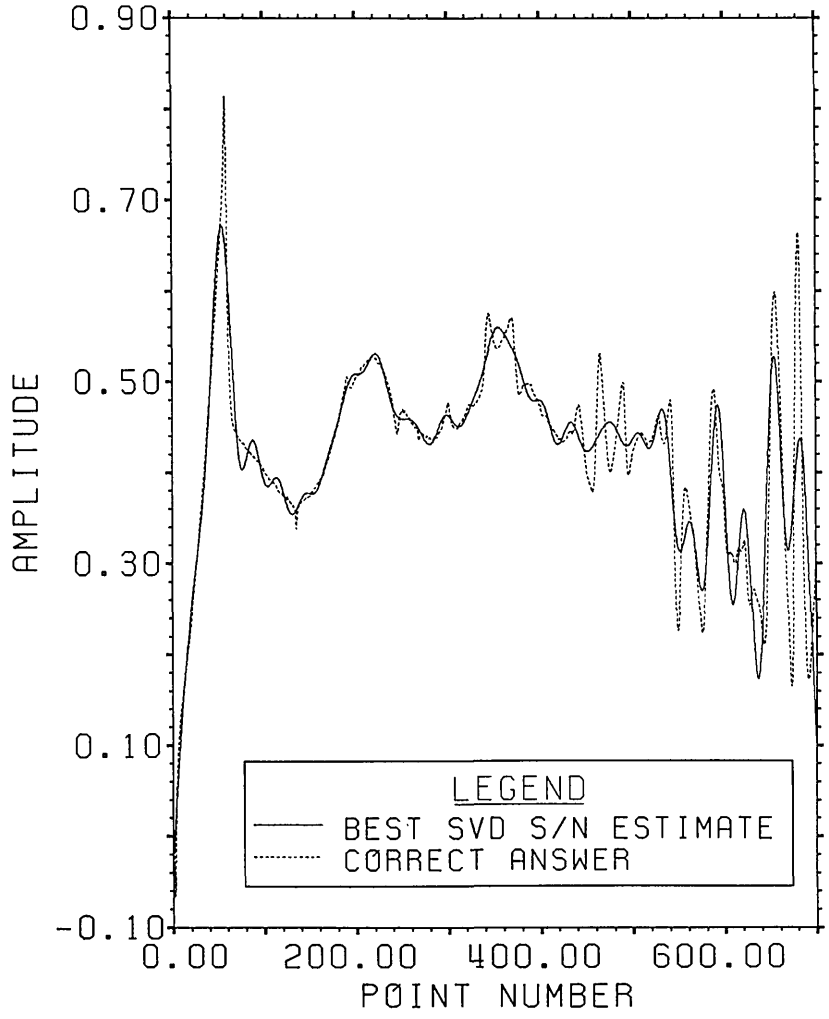

Figure 10. Best Lanczos/SVD estimate for Example Ib using S/N projection.

the alternate problem, for instance, is 366%. The extra data needed to identify the end features of the alternate answer are contained in the last 399 points of missing data.

In most problems with 50% incomplete range data one can still obtain accurate information about the answer through most of the time domain involved. Typical examples we have checked show good accuracy for 60% to 95% of the time range. An assessment of the importance of missing end data may be possible in RPD using knowledge of the positions of the roots in the root transform of the answer vis-a-vis those of the transform of the kernel.

The second stage of the cross-cut algorithm uses each of the two first stage outputs as *a priori* inputs to the alternate algorithm. For this example, since we have established that little meaningful data lie beyond point 140, we will use the windowed first estimates rather than the full estimates as *a priori* inputs. For the SVD process, then, the windowed RPD first estimate is convolved with the kernel and the first 400 points of the resultant time series are subtracted from the *a posteriori* data. Lanczos/SVD deconvolution is then performed on this reduced problem and the estimated inverse added back onto the *a priori* estimate to produce the second estimate.

Figure 18a shows the signal-plus-noise and inverse SVD spectra and figure 18b shows the true signal and noise spectra for this reduced SVD problem. With a -37 dB signal-plus-noise approximate cutoff, a reasonable choice from figure 18a would be a cutoff at 12 vectors, which in this case is confirmed by the curves in figure 18b. The r.s.d. for the second SVD estimate is then 56%.

Figure 19a shows the signal-plus-noise and inverse RPD spectra and figure 19b shows the actual range signal and noise spectra for the reduced RPD problem which has as its *a priori* input the

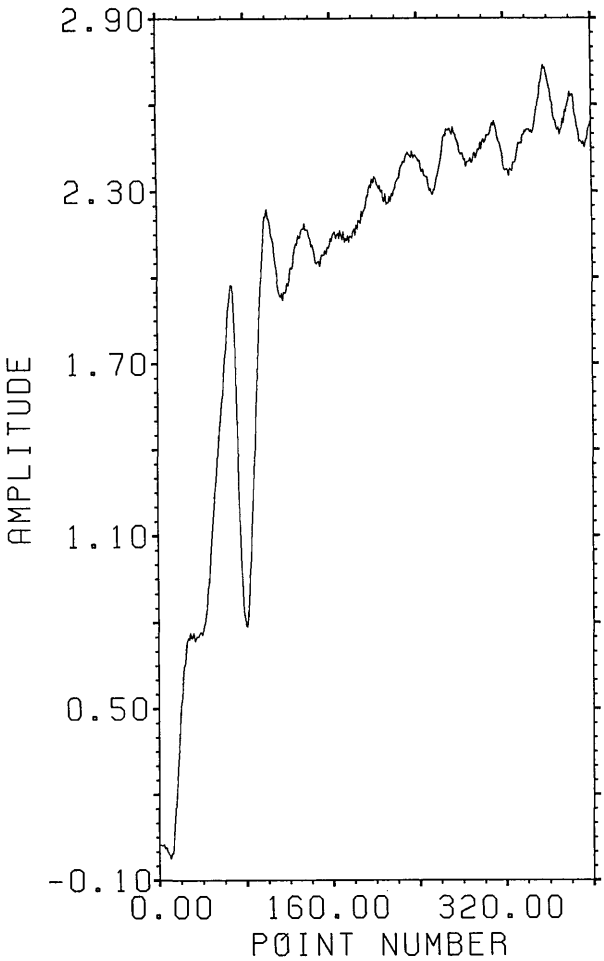

Figure 11. Synthetic output data with noise for model transducer calibration problem (Example II).

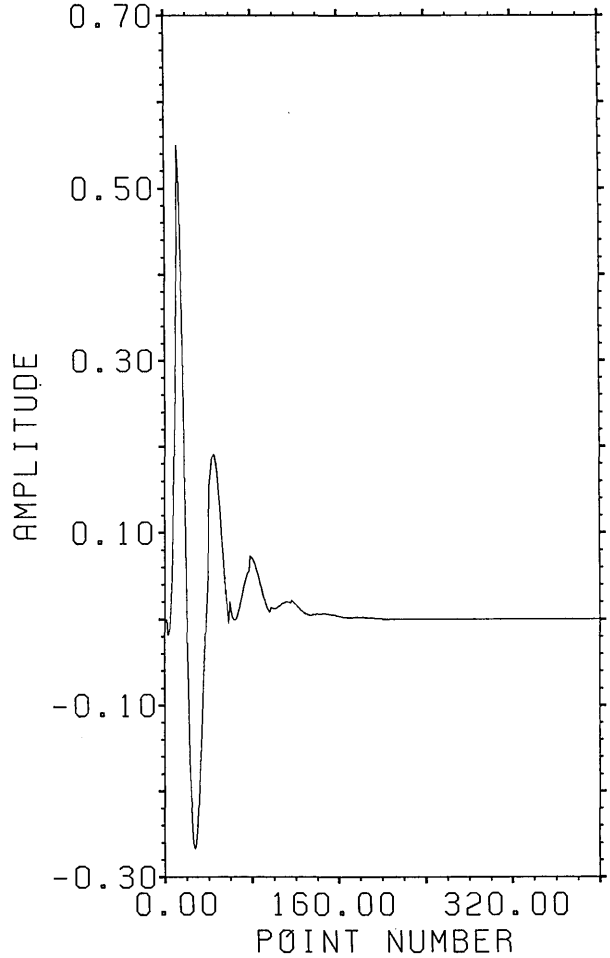

Figure 12. Correct answer for Example II.

windowed first estimate from the Lanczos/SVD algorithm. The extensive high frequency *a priori* noise level is no longer present, and the apparent cutoff from figures 19a and 19b is between 0.16 and 0.2 relative frequency. We choose the same filter (transition between 0.16 and 0.3 relative frequency) for this stage as in the first stage, since no high frequency noise is apparent. This second RPD estimate has an overall r.s.d. of 60%.

The final step in the cross-cut algorithm is to average the two estimates arising from the second approximations. The resultant average is shown compared with the correct answer in figure 20. It has an overall r.s.d. of 52%. Since the geometric mean of the answer is 0.014, the actual domain noise is −42.5 dB or 2.5 dB less than the added range noise. Figure 21a gives a comparison of the spectrum of the cross-cut estimate, windowed to 200 points to smooth the spectrum slightly, with that of the answer. Figure 21b gives a similar comparison of the phases. As can be seen from these figures, almost all the error in the estimate is at frequencies above 0.1.

4. Summary

A comparative development has been given of a number of related deconvolution methods including a new algorithm based upon root projection. Radiused FFT division, singular valued decomposition (SVD) and root projection deconvolution (RPD) were compared in detail in two examples. The first example involved deconvolution of a 101 point Gaussian filter from a relatively complex 800 point waveform in the presence of noise. Two levels of noise were used. The concept of signal-to-noise filtering was introduced and applied to both SVD and RPD. It was shown to be particularly powerful when used in conjunction with the SVD method. In both cases RPD provided the most ac-

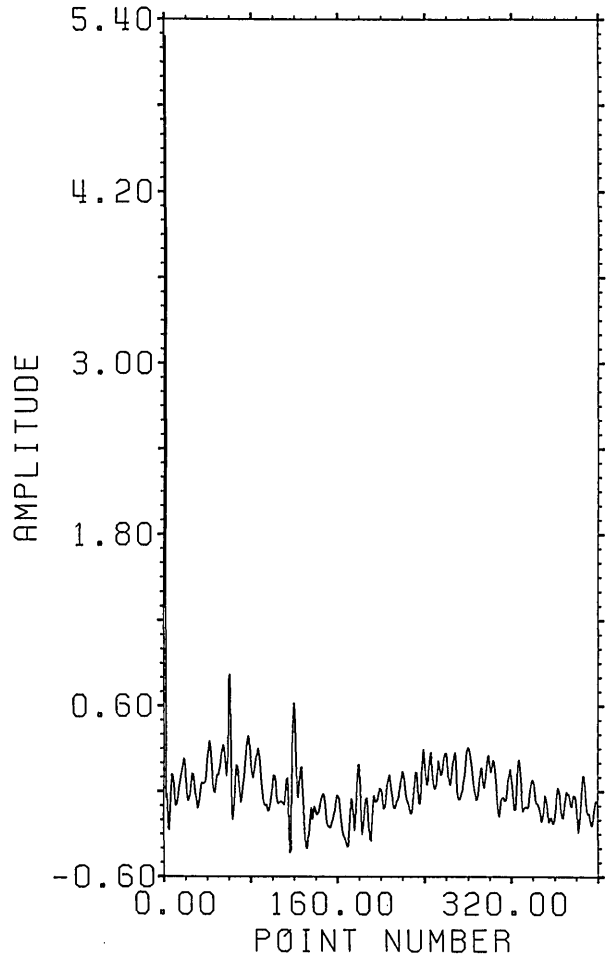

Figure 13. Result of using 800 point FFT division to estimate answer for Example II.

curate inversion estimate although this accuracy was matched by the SVD method for the higher noise level. The second example was a model transducer calibration problem with typically incomplete range data. A robust cross-cut deconvolution (CCD) algorithm utilizing the orthogonal decompositions of both the RPD and SVD methods was used in this example. The deconvolved estimate showed virtually no errors for frequencies below 10% of the Nyquist frequency.

Root transforms were applied in this work only for one dimensional time series. The study of the comparable Riemann surfaces for higher dimensional series or the use of boundary roots in finding signatures for the Radon transforms of such series have yet to be studied.

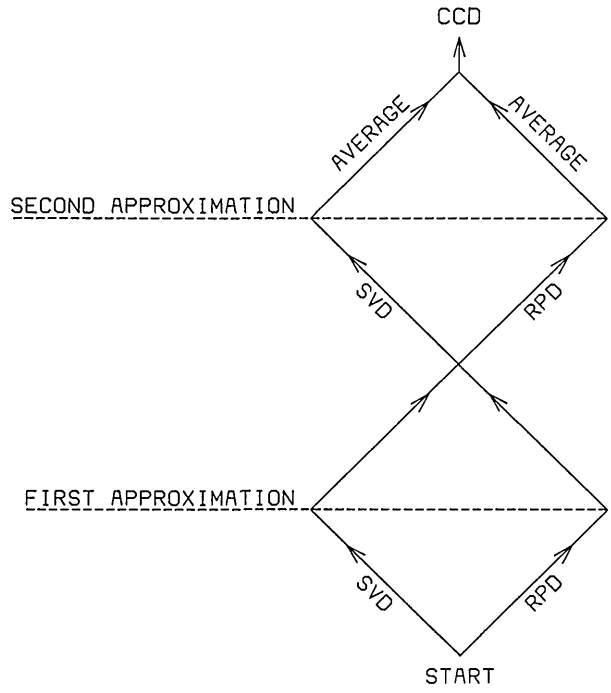

Figure 14. Schematic diagram for cross-cut deconvolution (CCD) algorithm.

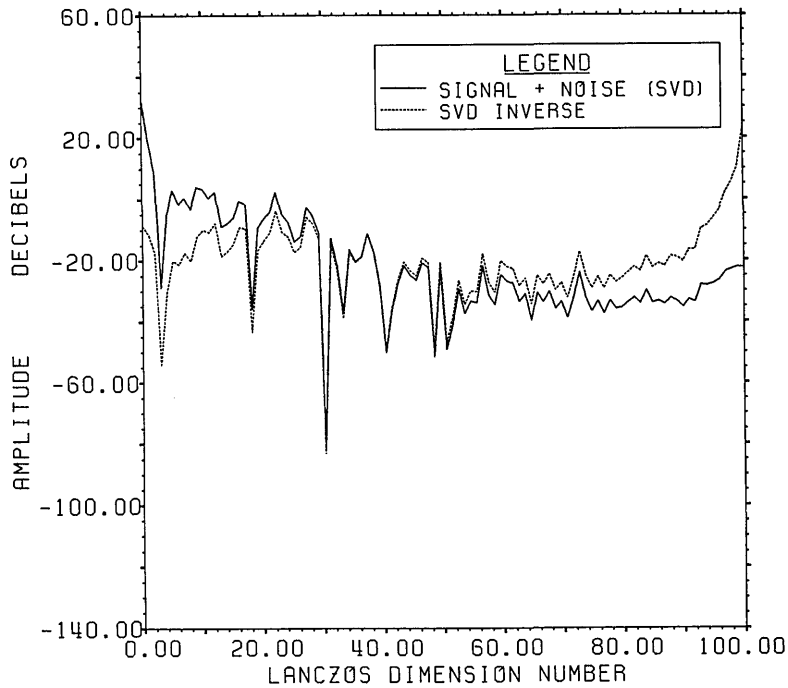

Figure 15a. Range signal-plus-noise and estimated inverse spectrum for Example II using the appropriate SVD bases (not the Fourier bases) built in the Lanczos subspace of dimension 100.

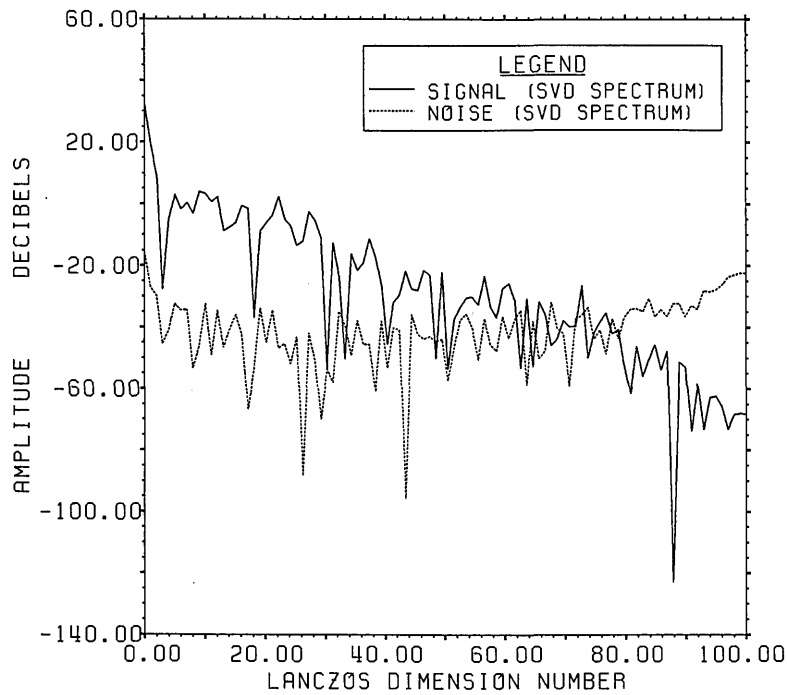

Figure 15b. The true signal and noise spectra for Example II using the SVD range basis on the Lanczos subspace of dimension 100.

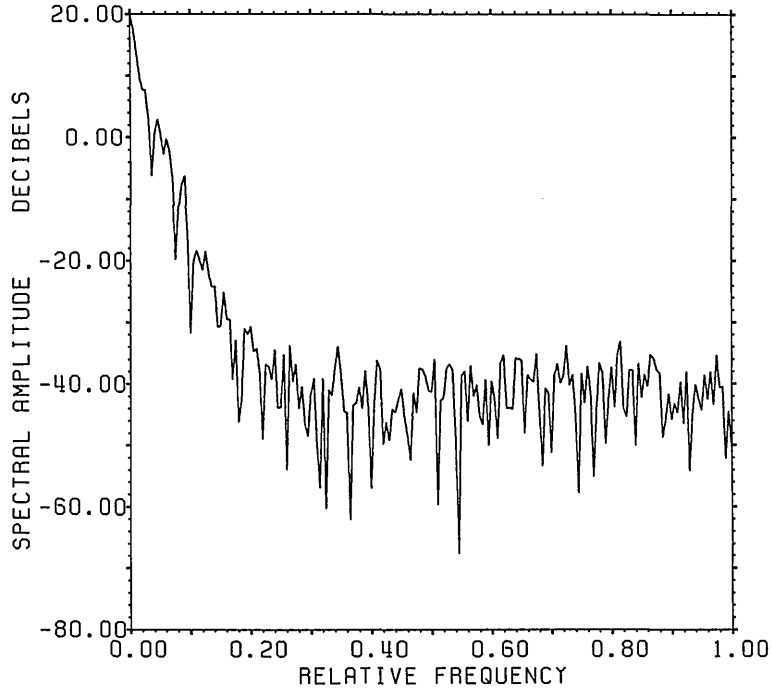

Figure 16a. Normalized 400 point FFT of the windowed output data of figure 11.

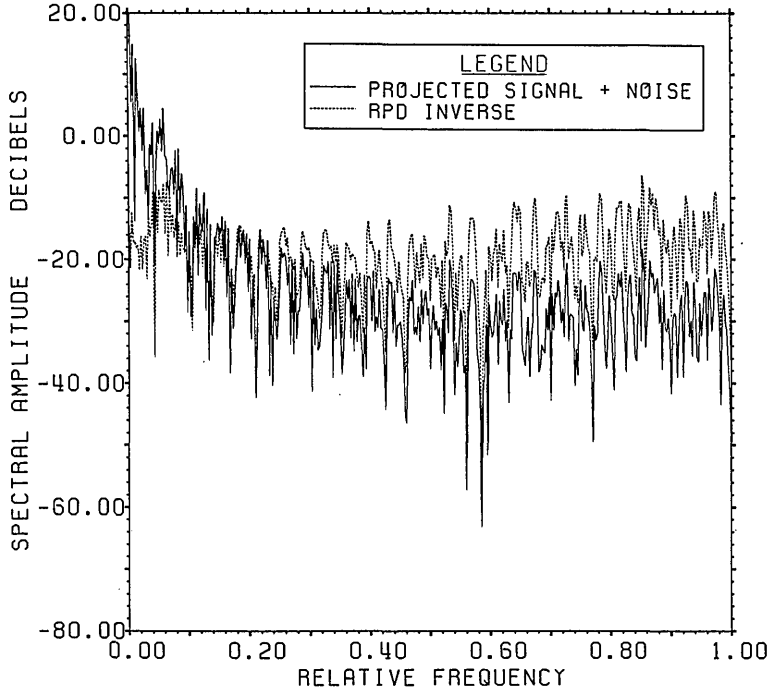

Figure 16b. Signal-plus-noise and estimated inverse spectra for example II using an 800 point FFT on the result of root projection where the data of figure 11 were simply end extrapolated out to 800 points.

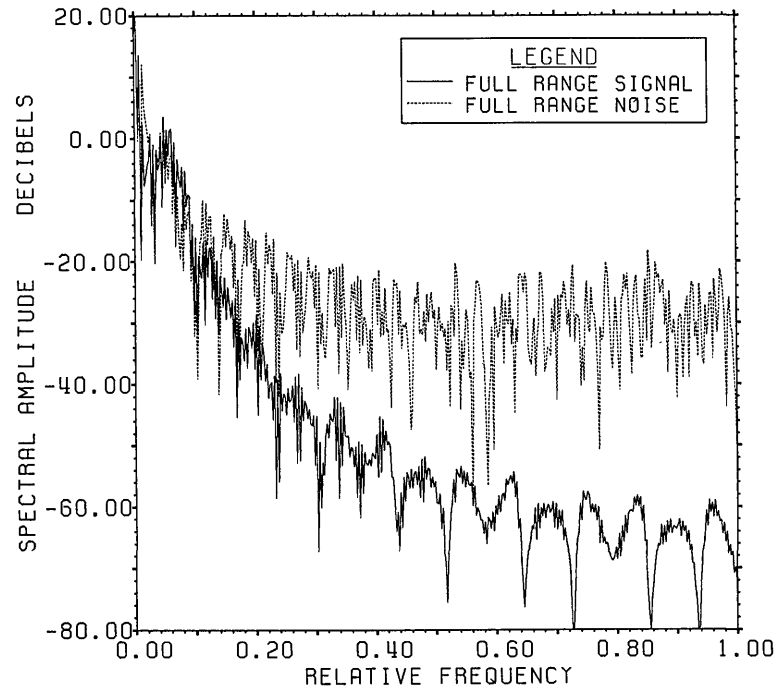

Figure 16c. Spectra of 800 point FFT of true range signal for Example II and noise from initial RPD inversion.

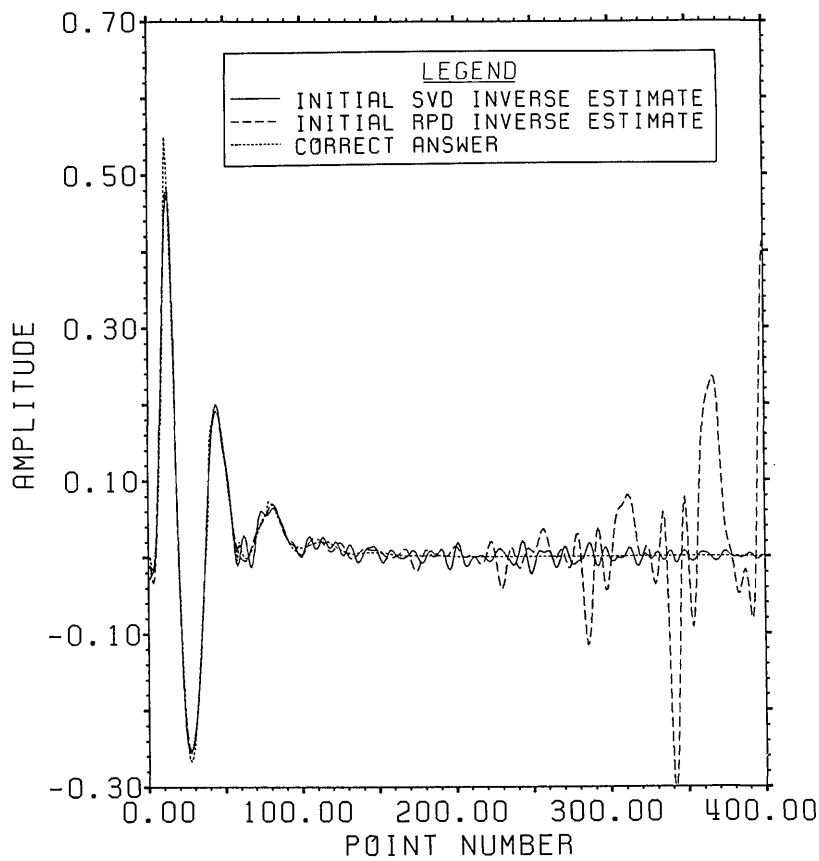

Figure 17. Initial SVD and RPD estimates for Example II.

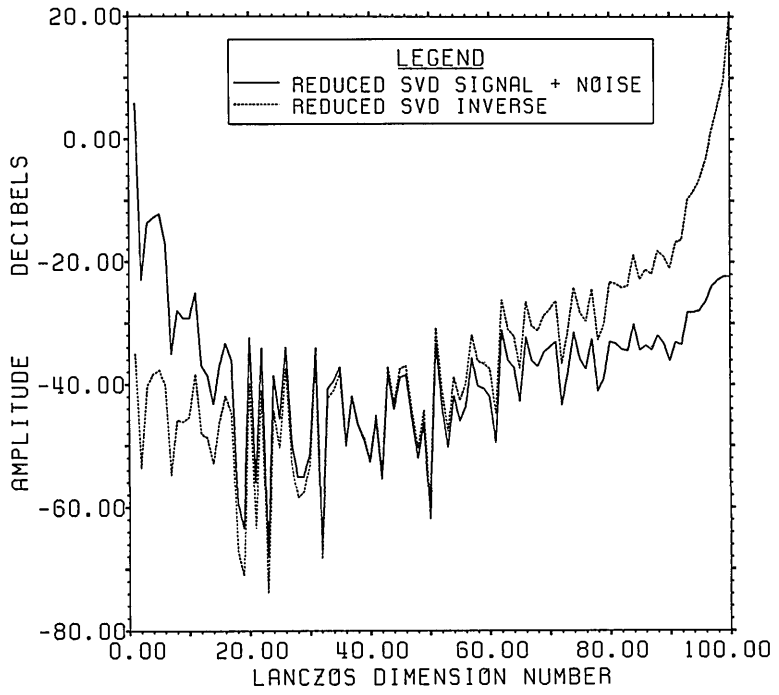

Figure 18a. Range signal-plus-noise and estimated inverse spectra for the second stage of the cross-cut algorithm in Example II. The appropriate SVD bases built in the Lanczos subspace of dimension 100 are used. The input is the first stage RPD estimate windowed to 140 points.

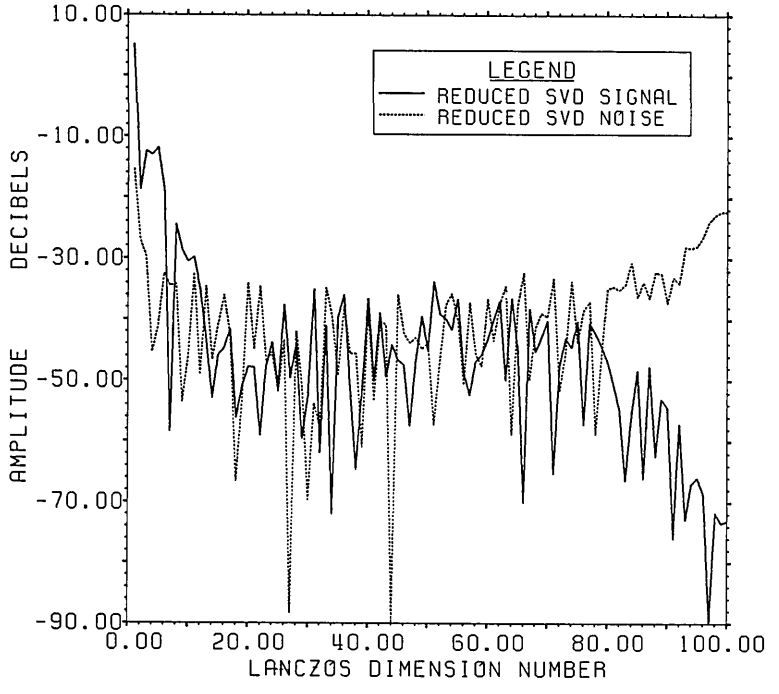

Figure 18b. The true signal and noise spectra for the second stage of the cross-cut algorithm in Example II using the SVD range basis on the Lanczos subspace of dimension 100.

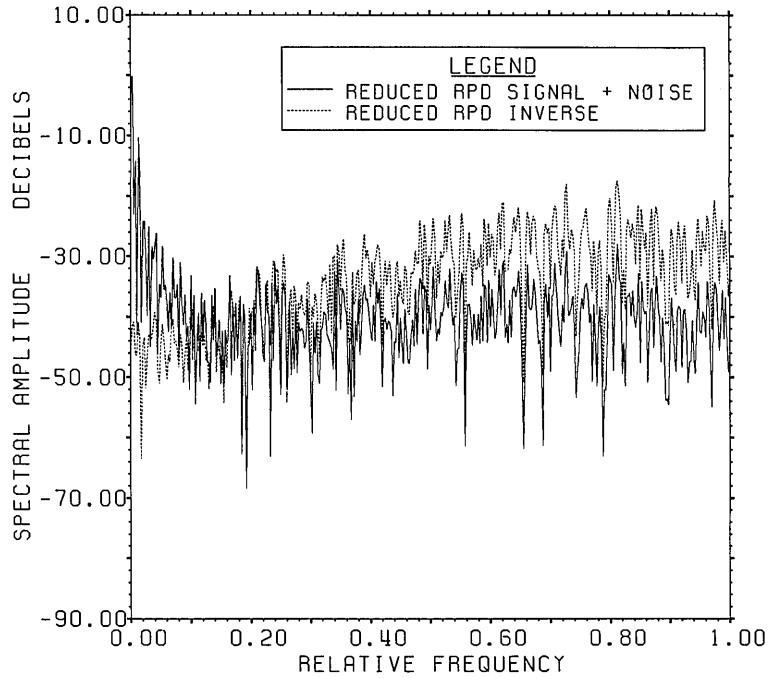

Figure 19a. Signal-plus-noise and estimated inverse spectra for the second stage of the cross-cut algorithm in Example II. An 800 point FFT is used on the result of root projection from the windowed first stage SVD input.

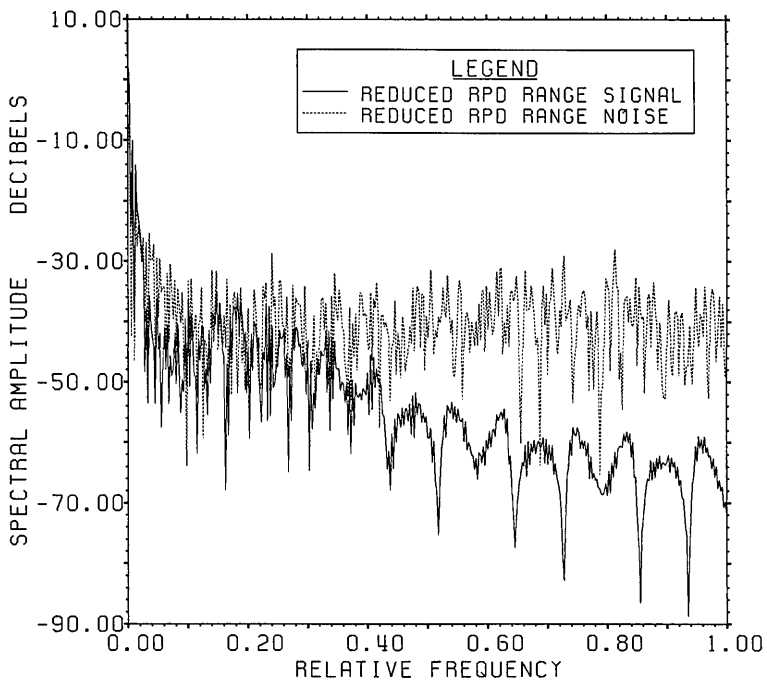

Figure 19b. The true range signal and noise spectra for the second stage of the cross-cut algorithm in Example II. An 800 point FFT is used on the result of root projection from the first stage SVD input.

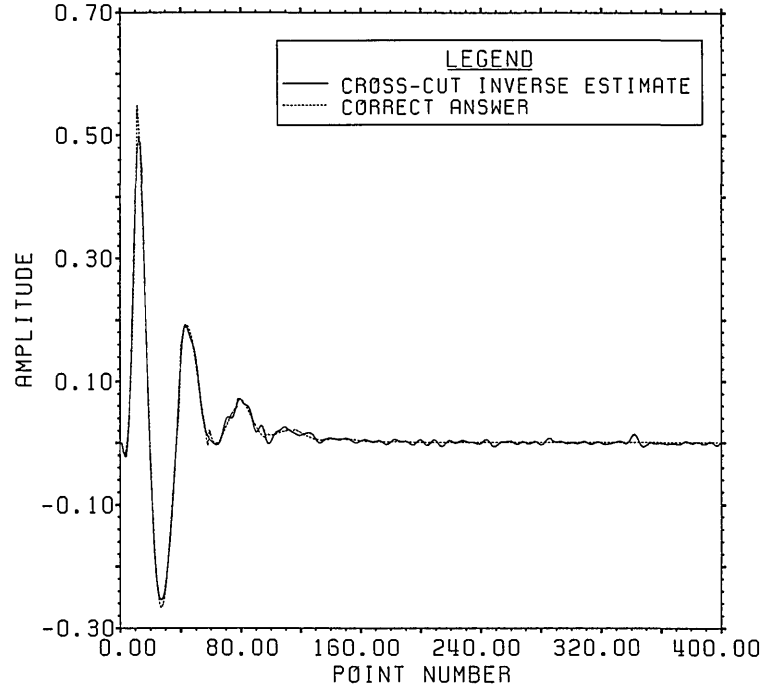

Figure 20. The CCD estimate for Example II.

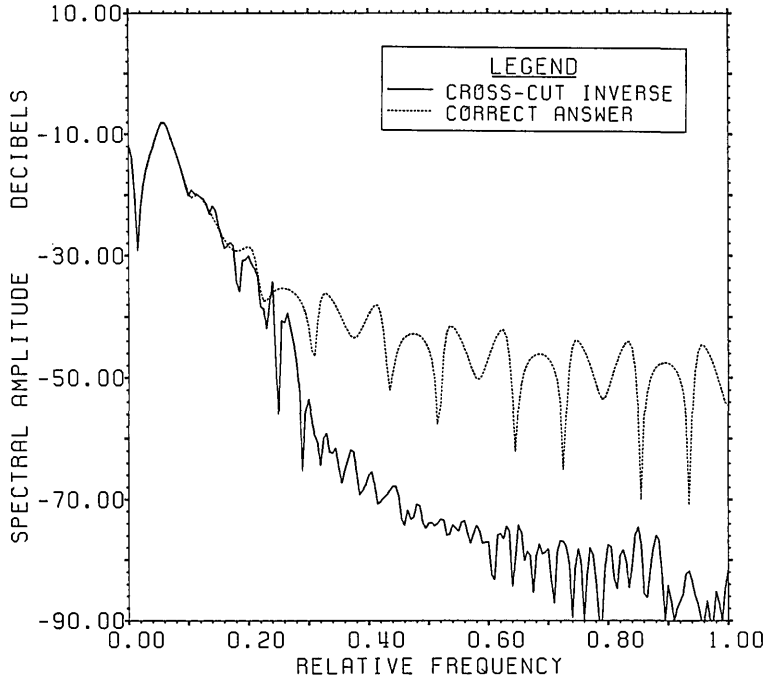

Figure 21a. Spectrum of the CCD estimate for Example II.

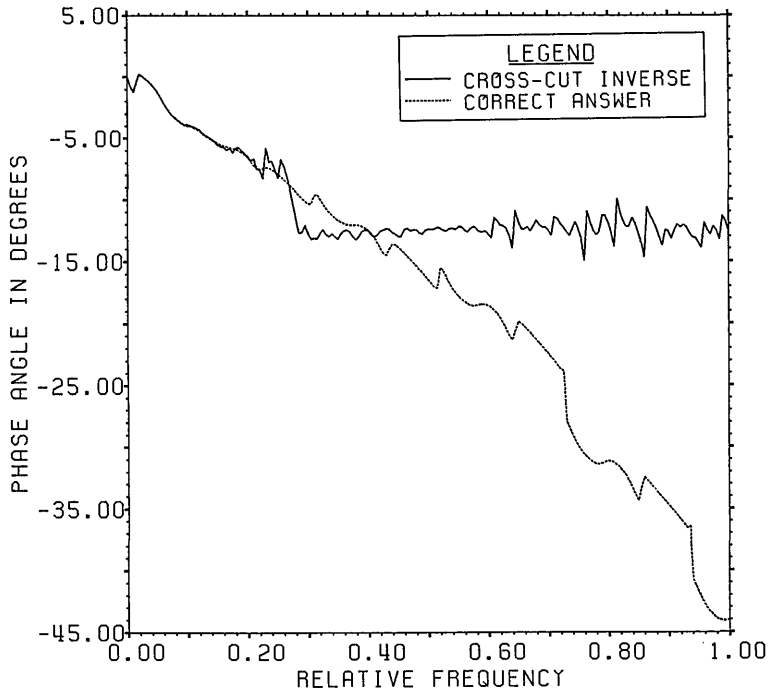

Figure 21b. Phase of the CCD estimate for Example II.

Acknowledgments

The author wishes to thank D. P. O'Leary, R. B. Clough, G. W. Stewart, and H. N. G. Wadley for many valuable suggestions and comments. This research was carried out under the joint NBS/DARPA program on NDE/RSP.

5. References

[1] Simmons, J. A., Boundary Roots and Signature Analysis of Causal and Transient Time Series, submitted to SIAM J. Math.
[2] Simmons, J. A., J. Res. Natl. Inst. Stand. Technol. **96**, 333 (1991).
[3] Shilov, G. E., and Gelfand, I. M., Generalized Functions, Volume I, translated by E. Saletan, Academic Press, New York (1964) pp. 160–165.
[4] O'Leary, D. P., and Simmons, J. A., SIAM J. Sci. Stat. Comput. **2**, 474 (1981).
[5] Halmos, P. R., Finite-Dimensional Vector Spaces, Springer Verlag (1974).

About the author: *John A. Simmons is a mathematician in the Metallurgy Division of the Materials Science and Engineering Laboratory of NIST, and has worked in the elasticity and elastodynamics of defects in solids. This work has arisen from studies of the signal processing of acoustic emission in materials.*

Volume 96, Number 3, May–June 1991
Journal of Research of the National Institute of Standards and Technology

News Briefs

General Developments

NIST SIGNS AGREEMENT WITH STATE FIRE MARSHALS GROUP

NIST and the recently established National Association of State Fire Marshals have signed an agreement to formalize a number of cooperative fire safety activities already in place and to set up several new ones. NIST has long worked with fire marshals in many states, including the investigation of major fires. Recently, a group of marshals and NIST staff worked together to define national fire research needs and priorities. As part of the new agreement, the two organizations will establish a database accessible by computer network containing technical information necessary to run fire models such as NIST's HAZARD I. Also, the agreement calls for a deputy fire marshal to work at NIST as a way to transfer technology to the fire community. Other activities will include joint fire safety research, especially on issues of national interest such as smoke toxicity.

INTEGRATED COOPERATIVE PROGRAM EXPANDED

NIST has expanded a cooperative program to bring together federal agencies and private organizations to develop prototype systems in Open Systems Interconnection (OSI), Integrated Services Digital Network (ISDN), and computer security. Program goals include facilitating the development and commercial availability of OSI, ISDN, and security mechanisms and services; providing a facility for users, developers, and vendors to jointly define, develop, and test systems in telecommunications, network management, and security; and coordinating funding from federal agencies. Program information and guidelines are available by writing the Integrated OSI, ISDN, and Security Program, B154 Technology Building, NIST, Gaithersburg, MD 20899. Fax requests to 301/948-1784. General program information is available from Patricia Noonan Edfors, B154 Technology Building, NIST, Gaithersburg, MD 20899, 301/975-3758.

NEW COMPUTER SECURITY SERIES PUBLICATION ISSUED

The first publication in a new series on computer security, entitled Bibliography of Selected Computer Security Publications: January 1980-October 1989 (NIST/SP-800/1) has been issued. Compiled by Rein Turn and edited by Lawrence E. Bassham III, the 200-page document serves as an excellent resource for individuals interested in computer security issues. The citations are listed under 10 categories: general, management, foundations, access control, trusted systems, database security, communication and network security, cryptography, privacy, and pre-1980 publications. Included in the bibliography are appendices with the addresses of all journals and magazines referenced in the bibliography and a list of key words. Available from the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402. Order by SN #003-003-03060-1 for $11 prepaid.

USERS' GUIDE TO DIODE LASERS

NIST and University of Colorado researchers have produced a guide for atomic physicists who want to use diode lasers in their work. The review article appeared in a recent issue of Review of Scientific Instruments. Traditionally, physicists have used dye lasers to tune to particular atomic transitions. But semiconductor diode lasers have many advantages: they are smaller, more efficient, less expensive, and very reliable. They have reasonable power, and their wavelength coverage is steadily increasing. "It is now possible to have a diode laser system which will produce more than 10 mW of tunable light

with a bandwidth of 100 kHz for a cost of less than $1,000," the researchers report. They discuss basic diode laser characteristics, advantages, and limitations. They also explain how to purchase the correct laser, tune it to the desired frequency, and control the laser's frequency. The guide, paper No. 68-90, can be obtained from Jo Emery, Div. 104, NIST, Boulder, CO 80303, 303/497-3237.

PRODUCING HYDROGEN SLUSH

NIST researchers have constructed an auger-type generator for producing hydrogen slush, the preferred fuel for the National Aerospace Plane. Until now the freeze-thaw method has been the most common for producing hydrogen slush, but it has several disadvantages not present with the auger-type method. For example, freeze-thaw produces hydrogen slush in batches whereas auger is a continuous process. The NIST-developed generator features a large (approximately 18 cm diameter) auger with a supercritical helium flow loop that simulates the performance of a helium refrigerator. Paper No. 6-91 discusses design and construction and is available from Jo Emery, Div. 104, NIST, Boulder, CO 80303.

APPARATUS FOR MEASURING COEFFICIENT OF FRICTION

NIST researchers have developed a novel apparatus to measure the coefficient of friction in oxygen environments. It will be used specifically to study problems with ball bearing wear in the high-pressure oxygen turbo pump of the space shuttle main engine. Although designed to last 7 1/2 h, these bearings often fail after 45 min of use. The NIST apparatus can handle loads up to 900 kg, and friction measurements can be made in inert atmospheres or in flowing oxygen at temperatures from 80 to 1030 K. To meet the various requirements, two measurement cells, one hot and one cold, were developed. One cell, made of high-temperature nickel alloy and cooled by ambient air, is used from room temperature to 1030 K. The other cell, made of high-purity nickel and cooled by nitrogen gas, operates from 80 to 523 K. Paper No. 8-91 describes the apparatus and is available from Jo Emery, Division 104, NIST, Boulder, CO 80303, 303/497-3237.

OUTREACH PROGRAMS HELPING INDUSTRY

NIST's Advanced Technology Program (ATP) will help industry to develop and apply generic technology it needs to commercialize new products and manufacturing processes, said the NIST Visiting Committee on Advanced Technology in a 31-page report to the Secretary of Commerce. The committee was established by the Omnibus Trade and Competitiveness Act of 1988 and given the responsibility to review and make recommendations regarding the general policy, organization, budget, and programs of NIST. Its 1990 Annual Report of the NIST Visiting Committee on Advanced Technology gives special attention to NIST's new industrial outreach programs, including the ATP and the Manufacturing Technology Centers program. The report also reviews progress in the institute's science and engineering programs, discusses NIST's industrial relations policies, and examines the agency's budget picture and organizational plan. Available by sending a self-addressed, stamped envelope to Dale Hall, A527 Administration Building, NIST, Gaithersburg, MD 20899.

STM BUILDING OF ROOM-TEMPERATURE NANOSTRUCTURES

The vision of custom-building atomic-scale structures—perhaps "nanomachines" or ultra-high-performance electronic devices—by individually manipulating atoms with a scanning tunneling microscope (STM) has inspired a number of recent experiments. The most dramatic results to date have involved positioning individual xenon atoms on a very cold (−269 °C) surface. Practical applications of that technique are somewhat limited because it requires that the surface be maintained at cryogenic temperatures, and only works with a limited class of atoms. NIST researchers recently demonstrated a novel room-temperature procedure for manipulating atoms or molecules on surfaces that promises much more general application. By generating an electric field at the tip of the STM probe, they induced cesium atoms adsorbed on typical semiconductor substrates to collect under the probe tip, forming novel structures that would not otherwise occur by simple adsorption. Details are reported in the March 8 issue of Science.

DEVICE PROMISES BETTER TRACKING OF DRUGS, VIRUSES

A prototype device that shows promise as a reliable detector of drugs, pollutants, bacteria, and viruses in humans has been developed at NIST. Called the liposome-based flow injection immunoassay (LipoFIIA) system, the device uses a component of the human immune system to measure the amount of a specific chemical compound from complex mixtures such as blood and urine. Automated, reusable, and fast, the system improves on the

performance of many other commercially available instruments. An analysis takes minutes instead of the hours or even days required by other methods. The LipoFIIA system is as much as 100 times more sensitive than comparable devices, allowing analyses at levels as low as parts per billion. It uses liposomes—submicroscopic, balloon-like globules—to detect how much of a given substance is in a biological sample. Antibodies, similar to the proteins the human body uses to recognize and latch onto foreign substances, also play a major role. The system recently passed clinical trials at the National Institutes of Health.

IGNITION CHARACTERISTICS STUDIED

At the request of the National Aeronautics and Space Administration, NIST is studying the ignition and combustion characteristics of selected alloys in pressurized oxygen. A number of serious accidents and fires have involved the ignition and combustion of metals in oxygen systems. As the operating pressure in these systems increases, the potential for disaster also increases due, in part, to the increased reaction rate of materials with oxygen. To reduce this potential, the ignition and combustion characteristics of metals must be known. In the past, NIST researchers have studied the characteristics of the iron-based alloy UNS S66289 and the nickel-based alloy UNS N07718. Most recently, NIST reported on the characteristics of the nickel-based alloy UNS N07001 in Ignition Characteristics of the Nickel-Based Alloy UNS N07001 in Pressurized Oxygen (NISTIR 3947). Available from the National Technical Information Service, Springfield, VA 22161. Order by PB #91-144428 for $17 prepaid.

PROTOCOL DEVELOPED TO ANALYZE ALTERNATIVE REFRIGERANTS

NIST researchers have developed a protocol for scientists studying alternatives to fully halogenated chlorofluorocarbons (CFCs). These CFC fluids, used for many years as refrigerants and propellants, must be replaced because they are thought to be contributing to ozone depletion in the atmosphere. In searching for alternative fluids, extensive experimental measurements of the replacements' physical properties are needed to evaluate their performance. Sound interpretation of these measurements requires that all components of the fluids be known. Impurities in the samples can plague an experimental measurement and ruin interpretation of results. The protocol, Strategy for Chemical Analysis of Alternative Refrigerants (NIST TN 1340), provides for verifying sample identity, characterizing major impurities, and determining the concentrations of impurities. A protocol and apparatus for high-temperature, high-pressure thermophysical property measurements are also described. Available from the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402. Order by stock no. 003-003-03062-7 for $5 prepaid.

"EXPECT" AUTOMATES INTERACTIVE PROGRAMS

"Expect," a software tool for automating interactive programs, is available on request from NIST. Expect, originally designed to automatically log in NIST's Automated Manufacturing Research Facility computers, runs on UNIX systems but can also control non-UNIX computers and networks. A NIST scientist says the software "uses a script—much like a movie script—to simulate a human interacting with a computer. An important difference is that, unlike a movie script, Expect scripts can describe alternatives and then 'play' differently each time as they adapt to the situation." Since April 1990 over 1,600 sites have requested Expect. Requests should be sent to Don Libes, A127 Metrology Building, NIST, Gaithersburg, MD 20899 or library@cme.nist.gov.

GRANTS ANNOUNCED TO ADVANCE INDUSTRIAL TECHNOLOGY

Commerce Secretary Robert A. Mosbacher recently announced the first grant awards under the Commerce Department's Advanced Technology Program (ATP), which he said could lead to the birth of revolutionary products and processes in key U.S. industries and help boost the country's trade and competitiveness. Eleven new research and development programs were selected for funding under the ATP in fundamental industrial technologies including improved manufacturing techniques for electronics (such as x-ray lithography); optical recording; a variety of hardware and software technology for computers; high-temperature superconductivity; machine tool control; and novel laser designs. The awards will provide approximately $9 million in first-year grants to initiate nearly $100 million in R&D programs over the next 5 years. The research and development programs selected for funding are:

Volume Holographic Mass Storage Subsystem—Microelectronics & Computer Technology Corporation

- Nonvolatile Magnetoresistive Semiconductor Technology—Nonvolatile Electronics, Inc.
- Short-Wavelength Sources for Optical Recording—National Storage Industry Consortium
- Tunable Deep UV and VUV Solid-State Laser Source—Light Age, Inc.
- New User-Interface for Computers Based on On-Line Recognition of Natural Handwriting—Communication Intelligence Corporation
- Printed Wiring Board Interconnect Systems—National Center for Manufacturing Sciences, Inc.
- Advanced Manufacturing Technology for Low-Cost Flat Panel Displays—Advanced Display Manufacturers of America Research Consortium
- Fabrication and Testing of Precision Optics for Soft X-Ray Projection Lithography—AT&T Bell Laboratories
- Solid-State Laser Technology for Point-Source X-Ray Lithography—Hampshire Instruments, Inc. & McDonnell Douglas Electronic Systems Co.
- Advanced Compensation Techniques for Enhancing Machine-Tool Accuracy—Saginaw Machine Systems, Inc.
- Advanced Thallium Superconductor Technology—E.I. du Pont de Nemours & Co.

TWO SITES NAMED FOR TECHNOLOGY TRANSFER CENTERS

The Industrial Technology Institute in Ann Arbor, MI, and the Kansas Technology Enterprise Corporation of Topeka, KS, were selected to establish regional manufacturing technology transfer centers. The two organizations now will negotiate cooperative agreements with NIST for approximately $1.5 million each for the first year, which they will match. "The establishment of Michigan and Kansas centers means that small and medium-sized businesses in the two areas soon will have improved access to the advanced manufacturing technology and processes they need to compete and prosper," Commerce Secretary Robert A. Mosbacher said. "The new centers complement the efforts of centers in Ohio, South Carolina, and New York, which are already working with the Commerce Department to modernize America's industrial sector through technology transfer." The NIST manufacturing technology centers act as clearinghouses between industry and sources of manufacturing technology.

COMMENTS SOUGHT FOR MACHINE TOOL STANDARD

Comments are invited on a proposed American Society of Mechanical Engineering (ASME) standard to provide machine tool builders and users a uniform method to evaluate performance of metal-cutting machining centers. A NIST mechanical engineer who heads the ASME committee that drafted the voluntary industrial standard, says "A single national standard will give buyers a more credible way to selecting manufacturers. Manufacturers, at the same time, will have a clear way for verifying their improvements in the marketplace." "Methods for Performance Evaluation of Computer Numerically Controlled Machining Centers" is based on research at NIST, Lawrence Livermore National Laboratory, the University of North Carolina (Charlotte), and the University of Florida (Gainesville). Contact Denver Lovett, Fabrication Technology Division, NIST, Gaithersburg, MD 20899, 301/975-3503.

ELECTROMAGNETICS PUBLICATIONS LISTED

Two new bibliographies list all publications (1970-July 1990) by staff in two NIST divisions. A Bibliography of the NIST Electromagnetic Fields Division Publications (NISTIR 3945) deals with measurement methods and standards for antennas; dielectric properties; electromagnetic interference and susceptibility; microwave power, impedance, and attenuation; near-field antenna measurements; noise; remote sensing; time domain reflectometry; and waveform measurements. Metrology for Electromagnetic Technology: A Bibliography of NIST Publications (NISTIR 3946) covers measurement methods and standards for laser systems; optical fibers and communication equipment; cryoelectronics; magnetics; superconductors; and other unusual electrical engineering materials. Available from the National Technical Information Service, Springfield, VA 22161. Order NISTIR 3945 by PB #91-132241 for $23 prepaid and NISTIR 3946 by PB #91-132266 for $17 prepaid.

MULTIMEDIA COURSEWARE PUBLICATION ISSUED

The federal government and other U.S. organizations are likely to invest billions of dollars to develop multimedia training materials for use in computer-based interactive training systems. Multimedia Courseware in an Open Systems Environment: A Federal Strategy (NISTIR 4484), discusses the federal strategy for creating an environment in which high-quality portable courseware is

available as commercial off-the-shelf products competitively supplied by vendors. The strategy comes from the Department of Defense Portable Courseware Project, which requires standard software interfaces. Available from the National Technical Information Service, Springfield, VA 22161. Order by PB #91-143362 for $17 prepaid.

1992 OSI WORKSHOP SCHEDULES RELEASED

NIST announced the 1992 workshop schedules for implementors of Open Systems Interconnection (OSI). The 1992 meeting dates for the workshops to reach implementor agreements on OSI computer network protocols are March 9-13, June 8-12, Sept. 21-25, and Dec. 14-18. Hosted by NIST and held in Gaithersburg, MD, the workshops will cover protocols in seven layers of the ISO Reference Model. Attendance is limited by space requirements. Registration is on a first-come, first-served basis, and a fee will be charged for attending the workshops. For registration information, contact Brenda Gray, 301/975-3664; individuals with technical questions may contact Tim Boland, 301/975-3608.

NIST/ASM COMPLETE ALLOY PHASE DIAGRAM PROGRAM

Marking the completion of a successful $10 million program between government and industry, officials from ASM International presented the second edition of the Binary Alloy Phase Diagrams to NIST recently. The three-volume compendium, designed to help engineers and materials users worldwide, contains more than 3,000 alloy phase diagrams—"road maps" describing what happens when two elements are combined at various temperatures. The data represent the results of a 10 year effort by ASM and NIST to assess the quality of all previous work in the field. Under a 1981 agreement, NIST was responsible for providing quality control and technical guidance to more than 50 technical experts worldwide. The new edition replaces a smaller one, which had become an internationally accepted source of critically evaluated standard reference data on alloy phase diagrams. For information, contact the Alloy Phase Diagram Program, ASM International, Materials Park, OH 44073, 216/338-5151.

AUTOMATION DECISIONS EASIER WITH AUTOMAN 2.0

AutoMan 2.0, a new version of the personal computer software package that supports complex automated manufacturing investment decisions, should make life easier for industry managers. The revised program enables users to see graphically the results of investment alternatives. AutoMan 2.0's sensitivity analysis feature depicts how investment alternatives would be rated if the importance of any criterion, such as lead time, return, life-cycle cost, or product quality, was changed. The software includes several starter decision models with criteria already specified. Users can apply these models or develop entirely new ones with up to 49 impact criteria. NIST's Office of Applied Economics developed AutoMan 2.0 in collaboration with NIST's Automated Manufacturing Research Facility and the U.S. Navy. Available from the National Technical Information Service, Springfield, VA 22161. Order by PB #91-506568 for $50 prepaid, plus $3 shipping.

FIPS PUBLICATION 160 APPROVED FOR C

The Commerce Secretary approved a Federal Information Processing Standard (FIPS) for the programming language C, adopting voluntary industry specifications (ANSI X3.159-1989). This standard specifies the form and establishes the interpretation of programs written in the C programming language. The purpose of the standard is to promote portability of C programs for use on a variety of data processing systems. The standard is used as the reference authority in developing compilers, interpreters, and other forms of high-level language processors. It is also used by computer professionals who need to know the precise syntactic and semantic rules adopted by ANSI. Federal standards for high-level programming languages permit federal agencies to exercise more effective control over the production, management, and use of the government's information resources. For technical details, contact L. Arnold Johnson, A266 Technology Building, NIST, Gaithersburg, MD 20899, 301/975-3247.

COMPUTER GRAPHICS METAFILE TEST SERVICE DEBUTS

On May 1, the Computer Graphics Metafile (CGM) Test Service begins a 1 year trial program. The service will analyze a CGM file to test if it meets requirements that allow the transfer of pictures among different graphical software systems, graphical devices, and computer graphics installation. The testing will benefit both users and vendors by increasing confidence in the performance and quality of CGM products and in successfully interchanging graphics pictures. The two requirements used are the Federal Information Processing Standard (FIPS) 128, "Computer Graphics Metafile," and the Computer-aided Acquisition

and Logistic Support (CALS) Application Profile (MIL-D-28003). An information pack containing details on prices, the test service, and how to submit files is available from Lynne S. Rosenthal, A266 Technology Building, NIST, Gaithersburg, MD 20899, 301/975-3353.

NEW DIRECTORY OF STANDARDS ACTIVITIES AVAILABLE

Standardization Activities of Organizations in the United States (SP 806) summarizes the standards activities of more than 750 organizations in the United States, including federal agencies and approximately 425 private-sector groups. The largest section contains an alphabetical listing of 637 nongovernment organizations that develop standards or contribute to the standardization process by working with other organizations, or are sources of documents and information. The format provides quick access to information on the type of organization, scope of activities, whether standards are voluntary or mandatory, availability, and key words. Entries for 77 federal agencies, departments, and other organizational components that develop standards are included, as well as a section on sources for information, a subject index, acronyms and initials, former names of some organizations, and names of those in a previous directory no longer involved with standards. Available from the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402. Order by stock no. 003-003-03070-8 for $31 prepaid ($38.75 foreign).

MICROWAVE POWER MEASUREMENT TO BE IMPROVED

NIST expects to announce within a year a new premium calibration service for coaxial microwave power sensors. Advance notice is being provided since users of this service will need to have a thermistor-type bolometer mount built to NIST specifications. This is necessary because some essential design characteristics of commercially available general-purpose thermistor mounts make them incompatible with a microcalorimeter. The special mounts, which may have either a Type N or APC-7 connector, will be measured directly in the NIST microcalorimeter. This will give the customer a reference standard equal to that used by NIST for calibration transfers. By using the proper transfer technique, the customer will be able to calibrate other power sensors with an uncertainty equal to the standard service offered by NIST. Measurement uncertainty for the new service is expected to be approximately one-half to one-third the presently stated uncertainty of 1.8 percent at 18 GHz. The present service, with some accuracy improvement, will continue to be available. For more information, contact Fred Clague, Division 813.01, NIST, Boulder, CO 80303, 303/497-5778.

106 COMPANIES VIE FOR 1991 BALDRIGE AWARD

NIST announced that 106 companies have applied for the 1991 Malcolm Baldrige National Quality Award. That number includes 38 manufacturing firms, 21 service companies, and 47 small businesses. Last year, 97 companies applied. The award, named for the late Commerce Secretary Malcolm Baldrige, was established by legislation in August 1987. It promotes national awareness about the importance of improving quality management and recognizes quality achievements of U.S. companies—though the award is not for specific products or services. Firms applying must undergo a rigorous examination. On-site visits for those passing an initial screening will take place in September and will be followed by an announcement and award ceremony in the fall. In previous years, the awards have been presented by the president during a ceremony in Washington, DC. The award program is managed by NIST with the active involvement of the private sector. A maximum of two awards may be given annually in each of three categories: manufacturing, service, and small business.

FEDERAL LAB ACCREDITATION PROGRAMS GUIDE AVAILABLE

The Directory of Federal Government Laboratory Accreditation/ Designation Programs (SP 808) is designed to help users from government, commerce, and industry locate federal laboratory accreditation programs and the organizations designated by the agencies to assist them in carrying out their responsibilities for testing products and services. Thirty-one lab accreditation programs are listed, as well as 13 federal programs with limited types of assessment. Entries are organized by agency, department, or independent commission and contain a program description, date initiated, authority, fields of testing accredited or designated, products affected, program requirements, availability of publications, and accreditation criteria. A limited number of copies are available. Send a self-addressed mailing label to Standards Code and Information Program, A633 Administration Building, NIST, Gaithersburg, MD 20899, 301/975-4031. A list of other standards-related and certification directories also is available.

IONIZING RADIATION CALIBRATION PROGRAM AVAILABLE

NIST has established a program to accredit laboratories to perform ionizing radiation calibration services under the National Voluntary Laboratory Accreditation Program (NVLAP). The program is designed for users of instruments to ensure the protection of workers in medical facilities, defense plants, and nuclear fuel-cycle operations. The program was requested by the Departments of Energy and Defense and other federal labs that provide the types of secondary calibration services included in the program. Laboratories will be accredited by NIST for 1 year and can maintain accreditation by demonstrating compliance with NVLAP criteria through on-site assessment every 2 years and annual proficiency testing. For information, contact Nancy M. Trahey, Chief, Laboratory Accreditation Program, A124 Building 411, NIST, Gaithersburg, MD 20899, 301/975-4016, fax: 301/975-3839.

INTERCOMPARISON OF POWER MEASUREMENTS CONDUCTED

Precise measurement of optical power is crucial to proper performance of modern optical communication systems. Power measurements need to be traceable to national standards, and standards from various countries should show good agreement. In a recent intercomparison of power standards among laboratories in the United States (represented by NIST), United Kingdom, Germany, and Australia, the maximum difference between any two laboratories was 0.36 percent and the maximum deviation of any one laboratory from the average ranged from 0.19 to 0.26 percent. Limited to wavelengths and power levels useful for optical fiber measurements, the intercomparison was made of responsivity scales at wavelengths of 1300 and 1500 nm—the wavelengths of interest to the optical fiber telecommunication community. The intercomparison served as a pilot study for a larger intercomparison now under way, which involves 14 countries. Paper No. 13-91 gives the conclusions of the pilot study. Available from Jo Emery, Division 104, NIST, Boulder, CO 80303, 303/497-3237.

FOURTH GENERATION LANGUAGES REPORT ISSUED

Use of fourth generation languages (4GL) has increased in data processing organizations, especially where end-users take on more programming responsibility. (4GL generally refers to nonprocedural, end-user-oriented computer languages.) Since no standards exist for 4GL, managers who select a given 4GL need a method to determine how well it will meet organizational, application, and user requirements. This 60-page report, Functional Benchmarks for Fourth Generation Languages, contains functional benchmarks (as opposed to performance benchmarks), which consist of a testing method and descriptions of tests to evaluate a particular 4GL. There are two levels of test result evaluation: the ability of the 4GL to perform a task; and the ease of performing it. Available from the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402. Order by stock no. 003-003-03071-6 for $3.25 prepaid.

LASER DAMAGE SYMPOSIUM PROCEEDINGS PUBLISHED

The proceedings of the 21st Annual Symposium on Optical Materials for High-Power Lasers (Boulder Damage Symposium), held Nov. 1-3, 1989, are now available. Laser Induced Damage in Optical Materials: 1989 (SP 801) describes sessions held on materials and measurements, surfaces and mirrors, thin films, and fundamental mechanisms. The emphasis was on new frontiers and developments, particularly on materials for high-power apparatus, and primarily in the wavelength range from 10.6 μm to the ultraviolet. Highlights include surface characterization, thin-film substrate boundaries, and advances in fundamental laser-matter threshold interactions and mechanisms. Full text and illustrations of all papers presented at the symposium are included in the 670-page book. Available from the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402. Order by stock no. 003-003-03061-9 for $30 prepaid.

ACTIVE GLASS INTEGRATED OPTICAL DEVICE DEMONSTRATED

NIST scientists have designed, fabricated, and demonstrated the successful operation of an active glass integrated photonic device that demonstrates the potential for new active multifunction components fabricated in rare-earth doped glass. This technology combines many of the attractive aspects of the well-established rare-earth doped optical fibers with the dense packing and integration possibilities provided by methods that have already been worked out for silicon integrated circuits. The device is known as a 1-by-2 Y-branch splitter. The NIST active splitter compensates for the loss of signal amplitude that occurs in passive devices when the signal from a single channel is split into two.

Splitters are used in optical broadcast networks to fan the output from a single channel to N output waveguides. A feed optical fiber attached to the input waveguide launches light into it; other fibers are attached to the N output channels. The splitter is designed to disperse light uniformly from the input fiber to the output fibers. In the passive splitters now in use, the signal amplitude per channel is reduced by 1/N. The NIST active splitter is fabricated in neodymium-doped glass and provides for signal gain through separate optical pumping. Demonstration of the splitter operation has shown that the 3 dB splitting loss otherwise expected in a 1×2 splitter can be compensated for fully. In addition, the NIST device has been operated as a multi-output integrated photonic laser. In this mode, the Y-branch splitter is folded on itself through optical feedback to form two phase-locked lasers operating simultaneously. NIST is extending these concepts to more complex devices and developing metrology to support them; for example, a nondestructive interferometric technique has been developed to measure precisely the optical path length of the component branches.

NEW STRUCTURES AND PROPERTIES DATABASE AND ESTIMATION PROGRAM RELEASED

A NIST scientist has completed a major new personal computer software product, with an affiliated database that has just been made available for distribution by the Standard Reference Data Program.

NIST Standard Reference Database 25, Structures and Properties Database and Estimation Program, presents a program that allows the user to use either a mouse or the keyboard to draw a molecular structure or substructure on the screen. Evaluated thermochemical data on the species can then be retrieved from a database that contains information on 5,000 chemical compounds. If data are not available for the compound of interest, the program includes software that can calculate values of enthalpies of formation, vapor pressures, and boiling points from widely used estimation procedures. Properties are estimated solely from the molecular structures drawn by the user, and the features of the program include automatic perception of rings and long-range interactions, determination of symmetry number corrections, and the computation of equilibrium constants for user-created chemical reactions. The program, which requires 3.2 megabytes of disk space, 640 K RAM, and EGA, VGA, or Hercules graphics adapter, also allows for data searches by chemical formula, name, and Chemical Abstracts Registry Number. The cost is $240.

ELLIPSOIDAL-MIRROR ANALYZER AT SURF-II

NIST recently commissioned an ellipsoidal mirror analyzer (EMA) at the NIST-SURF-II synchrotron light source. This instrument is a unique charged-particle energy analyzer that produces two-dimensional images of the emission from solids. This "display-type analyzer" shows the intensity of emission as a function of the angle from surface normal.

The heart of the instrument is an ellipsoidal mirror: a concave mirror that is elliptical in one plane and spherical in the perpendicular direction. A charged particle emitted from a sample placed at one of the foci reflects back from the mirror to the other focal point with its angular information intact. This reflection and an additional retarding-field provide the energy analysis. The particles are detected using pulse-counting techniques that provide a two-dimensional image.

The EMA permits simultaneous measurement of the energy, angle, and mass of an emitted particle; the latter coming from the use of time-of-flight techniques. By collecting all the pulses, angle-integrated spectra can be obtained. In the imaging mode, the "angle-resolved" spectra are unique. These images give a cross-sectional (k-space) view of the occupied states at a particular energy.

Initial images of electron and ion emission have been obtained from Ru(0001). The Ru 4*d* bands give images with dramatically different symmetries, depending on the initial binding energy of the state imaged. Present effort is focused on improving image transfer rates to give the enhanced signal-to-noise necessary to image monolayers of adsorbed molecules.

LASER ATOMIC LENS

Researchers at NIST have recently shown analytically that nanometer-sized spots of atoms can be produced by focusing an atomic beam with a laser. In a soon to be published paper they examine the focusing effects of a hollow, tube-shaped ("donut"-mode) laser beam on an atomic beam. When the laser beam is focused, such that the inside diameter attains a minimum at some point in space, it was found that this region acts as a lens for a neutral atom beam traveling down the axis of the laser beam.

The origin of the force exerted by the laser lies in the dipole moment induced in the atoms by the oscillating laser field. This dipole moment feels a force when the laser intensity has a gradient, as is present inside the hollow laser beam. In order for the force to be sizeable, the laser frequency must be tuned near a resonance in the atomic absorption spectrum.

One of the most intriguing properties of this laser-atomic lens is that, if the dimensions of the laser and atomic beams are kept small (of order 1 μm), spot sizes as small as 1 nm (0.001 μm) can in principle be obtained. In this situation, one of the major contributions to the spot size is diffraction of the atoms arising from their De Broglie wavelength. Such small spot sizes suggest a number of possible applications, such as microscopic atomic deposition, atomic microscopy, and precision measurements.

NEW X-RADIOGRAPHY FACILITY

NIST has installed a new, high-current, high-energy, industrial radiography source. The new facility, available to all NIST researchers as well as outside industry users, greatly expands NIST on-site capabilities for industrial radiographic applications. The high energy (420 kV) and high current (10 mA) will permit radiographic imaging through 10 cm of steel. In a recent application for the American Dental Association Paffenbarger Research Center, zirconium alloy ingots were radiographed to select specimens free of internal casting defects prior to machining.

The x-ray tube is mounted on a movable gantry to permit various beam orientations. The source is rated for 100-percent duty cycle and provides for automatic timing of exposures from 1 to 100 min. It may also be operated manually for any period. The stable, high-current device is also well suited for x-ray fluorescence studies in basic physics.

MEASURING GRAIN ORIENTATIONS IN THE SEM

A system for determining the crystallographic orientation of grains as small as 0.2 μm in any polycrystalline material has been installed on a scanning electron microscope at NIST. Orientation information increases understanding of the structure and properties of a number of materials systems, for example, high T_c superconductors, interfaces in electronic packaging systems and fiber-reinforced composites, and thin films. The specimens require very little specialized preparation beyond the production of a clean surface. Electron backscatter diffraction patterns are recorded in real time using a phosphor screen and low-light video camera. Analysis of the orientation information is carried out using software written here over the last year. Given the crystal structure of the material, the software is capable of automatically indexing the pattern once the positions of three or more zone axes have been measured. Additional hardware and software currently being installed will perform digital image processing and produce patterns of much higher quality, which will improve the accuracy of the technique.

HIGH-STRENGTH COMPOSITE INTERMETALLIC ALLOYS

In aerospace applications, there is an urgent need for low-density structural materials with high-temperature strength and low-temperature ductility. Titanium aluminides provide several alloy phases that show promise in these directions. By producing a composite of two such alloy phases, NIST scientists have developed a new intermetallic heterophase alloy with expected superior mechanical properties.

This development is based on the idea of combining two titanium aluminide ternary phases having complementary properties. The two phases were found to be in thermodynamic equilibrium with each other at elevated temperatures, thus providing stability in use. The phases combine to produce a high-strength composite microstructure. A variety of fine microstructures can be achieved by the use of different cooling and heating schedules.

REFERENCE DATA FOR FIRE MODEL VALIDATION

Fire models are becoming accepted for product design and building code use. The establishment of the accuracy of such models is thus as important as their creation. The NIST fire program, the international leader in fire and fire hazard modeling, has just completed documentation of five sets of real-scale fire tests, over 125 tests in all. These involve a variety of combustibles burned in enclosures ranging from a single room to a seven-story building. An upcoming report will include this information as well as a brief history of room fire testing, delineation of the model validation process, a discussion of the key measurements in validation testing, and

guidance for comparing experiment with prediction. The data from these tests will be available from NIST (via INTERNET), enabling fire modelers around the world to test their models uniformly.

NEW AUTOIGNITION MEASUREMENT METHOD

NIST scientists have developed and tested a new apparatus for making short-duration autoignition measurements of hydrocarbon fuels under conditions where the fuel/air stoichiometry, the nature of the hot metal surface, and the contact time are well-controlled. The work was sponsored by the U.S. Air Force Engineering and Services Laboratory. Finding serious problems with current closed-container methods, NIST scientists developed a flow-through device that provides much more reliable data on the importance of fuel structure effects. A key innovation is the ability to use various evenly heated metal foils as the igniting surface. They performed over 1,100 autoignition temperature determinations for 15 fuels, three metal surfaces, and three stoichiometries. The measured temperatures generally decrease slightly for larger hydrocarbons (with the C_2 hydrocarbons being quite low) and for richer mixtures. For the metal surfaces, the decreasing order of autoignition temperatures is nickel > 304 stainless steel > titanium. The data are consistent with literature suggestions that branched alkanes should be more resistant to autoignition than the linear isomers. A report on the work concludes with a discussion of the implications for engine knock and recommendations for further study.

FEDERAL INFORMATION PROCESSING STANDARDS (FIPS) REVISED

The Secretary of Commerce approved a revision to FIPS 120, Graphical Kernel System (GKS), which will be published as FIPS 120-1. FIPS 120, which adopts American National Standard Graphical Kernel System (ANS GKS), ANSI X3.124-1985, was modified to add a requirement for validation of GKS implementations acquired by the federal government. The standard specifies a library or toolbox package of subroutines to produce and manipulate two-dimensional pictures.

Also approved was a revision to FIPS 54, Computer Output Microform (COM) Formats and Reduction Ratios, 16 mm and 105 mm, to be published as FIPS 54-1. Effective July 1, FIPS 54-1 adopts American National Standard for Information and Image Management—Specifications for 16 mm and 35 mm Roll Microfilm, ANSI/AIMM MS14-1988, which specifies the image arrangement, size, and reduction ratios for 16 and 105 mm microforms generated by computer output microfilmers.

FIPS 9-1, Congressional Districts of the United States, provides the structure of numeric codes for representing congressional districts and similar areas defined for the various Congresses of the United States. A reissue of FIPS 9 incorporating technical changes, FIPS 9-1 is used in the collection, processing, and interchange of coded data by federal agencies.

NIST PUBLISHES VALIDATED PROCESSOR LIST

A new publication, NISTIR 4500, Validated Processor List, identifies COBOL, FORTRAN, Ada, and Pascal programming language processors that have a current validation certificate and those SQL language processors that have a registered test report. The list also includes GOSIP conformance testing registers. Processors scheduled for validation or processors having a current validation certificate or test report may be offered or delivered by vendors in response to requirements set forth in solicitations by federal agencies. The list is updated and published quarterly.

HIGH-TEMPERATURE JOSEPHSON JUNCTIONS FOR CRYOELECTRONIC INTEGRATED CIRCUITS

NIST has designed and fabricated a form of Josephson junction from high-critical-temperature superconductor material. This step completes the dem-onstration of capability to fabricate all the components needed for cryoelectronic integrated circuits such as ultrasensitive magnetic field detectors and microwave detectors and constitutes a significant step toward fabrication of more complex circuits such as high-speed digital processors, analog signal processors, and even voltage standard arrays. Because high-T_c superconductors do not require cooling with the expensive and difficult-to-obtain liquid helium required for low-T_c superconductors, capability to fabricate operating integrated circuits from high-T_c superconductors will promote their practical application and is expected to give rise to a range of commercial products. The NIST development contributes to a growing perception that successful commercialization will be possible.

In addition to the Josephson junctions, in the form of microbridges of high-T_c superconductor/normal metal/high-T_c superconductor, NIST has fabricated insulated crossovers between high-T_c layers, contacts between resistors and high-T_c elements, and contacts between one high-T_c element and another. The microbridges are formed by in situ deposition of YBCO and silver-gold alloy from different directions onto a substrate having a 100 to 500 nm step previously formed by ion milling. The step shadows the superconductor, leaving a gap which is crossed by the unshadowed normal metal. Device length, critical current, and resistance are controlled by forming the Josephson weak link on the edge of the step.

NIST ANNOUNCES SPECIAL-TEST SERVICE FOR HIGH-ACCURACY ELECTRICAL INSTRUMENTS

Special test measurement services to support precision multifunction calibrators, digital multimeters (DMMs), and low-voltage ac-dc transfer standards are now offered by NIST. These services have been developed in response to requests from users of a new class of highly accurate, stable, commercial DMMs and multifunction/multirange calibrators that have appeared in the marketplace in the past several years. Organizations frequently require calibration by NIST when they purchase these instruments to use them as laboratory standards for calibrating other instruments.

NIST measurement services for multifunction, multirange instruments had been available previously for the ac voltage function from 100 mV to 1 kV. Following the development and implementation of the necessary physical standards and measurement techniques, these services have been broadened to support the ac current, dc voltage and current, and dc resistance modes found on many of the new calibrators and DMMs. In addition, the ac voltage range has been expanded to lower voltage values between 1 and 100 mV, and the scope of low-voltage ac calibrations has been broadened to include ac-dc difference devices. The NIST calibration system is automated, employing standards that themselves are calibrated against higher-echelon standards maintained by NIST. The accuracies and ranges for these multifunction calibrations cover the capabilities of instruments known to be commercially available through early 1991.

NIST AND AIR FORCE APPROVE MOU ON INTELLIGENT CONTROL SYSTEMS

The Air Force Manufacturing Technology Directorate (MTD) and NIST have signed a Memorandum of Understanding (MOU) for the development of an open system architecture for real-time intelligent control systems. This MOU establishes the framework for cooperative research whose goal is the improvement of the national competitiveness of the aerospace manufacturing industry.

Under this MOU, MTD and NIST will identify and evaluate specific problems and research opportunities related to intelligent control systems, develop improved or new practices and standards, and conduct research to support these activities. The first joint efforts under this MOU involve research being conducted by NIST on the Air Force's Next Generation Controller (NGC) Program. This program has adopted the NIST NASREM control architecture for the NGC. NIST is assisting the Air Force in mapping the control modules for an intelligent machine tool controller into the NASREM architecture.

SURFACE ROUGHNESS OF TURNED PARTS IN REAL TIME

As part of real-time process-control strategy being implemented in the Quality in Automation (QIA) project, an investigation has been carried out at NIST to study the feasibility for pulsed ultrasound to resolve average roughness values and discrete asperities of less than 1 μm. A range of ultrasonic frequencies and beam dimensions was applied to turned sample surfaces using both scattering and profiling techniques. As a result, the ability to resolve an average surface roughness R_a value of less than 0.5 μm on a stationary 50 mm-diameter turned aluminum part has been demonstrated.

ELECTRON TRANSFER REACTIONS OF CYTOCHROME *C* ON MOLECULAR MONOLAYERS

A NIST scientist and university colleague have developed a new experimental strategy for studying the long-distance electron transfer reactions of the protein cytochrome *c*. Protein electron transfer is vitally important in controlling the bioenergetics of photosynthesis and respiration. In their approach, thin (25 Å), compact molecular monolayers terminated by carboxylic acid groups are formed by self-assembly on gold electrode surfaces. These

monolayer surfaces bind cytochrome *c* molecules irreversibly, thereby holding them a fixed distance from the gold electrode. The interfacial electron transfer rate from the gold electrode across the monolayer film to cytochrome *c* is then measured electrochemically. The researchers estimate an electron transfer distance of ~30 Å from both the electron transfer kinetic data and structural models of the monolayer-cytochrome *c* complex. The significance of this work is that the distance dependence of electron transfer can now be studied rigorously and easily by varying the thickness of the self-assembled monolayer. Moreover, the experiments show that self-assembled monolayers can be used to molecularly engineer surfaces to control the adsorption of biomolecules. This work is detailed in the Journal of the American Chemical Society 1991, 113, 1847-1849.

HIGH-TEMPERATURE SUPERCONDUCTOR THIN FILM

Thin films of high-temperature superconducting oxides are likely to play an extremely important role in many technologies, and research on the production of thin films is proceeding at a rapid pace. Scientists at NIST are currently producing thin films of $DyBa_2Cu_3O_7$, a superconducting oxide with a transition temperature of 89-90 K, by evaporation of the constituent elements in an ultra-high vacuum environment. This method, similar to the molecular beam deposition methods employed in the semiconductor industry, provides the control necessary to allow one to assemble the thin films in an almost atom-by-atom manner. In addition to studies employing a wide variety of high-resolution electron-microscopies (transmission and scanning electron microscopies), this growth method is coupled to in situ electron spectroscopic diagnostics for investigating the electronic and atomic structure of the oxides as grown. This combined approach will provide NIST scientists with the tools necessary to study in detail the growth of these materials, the correlation between growth conditions and superconducting properties, and chemical aspects of possible processing steps necessary for the success of high-temperature superconducting technologies.

RADIOCARBON (^{14}C) MEASUREMENTS USED TO TRACE ATMOSPHERIC ORGANIC POLLUTANTS

NIST researchers have succeeded in apportioning the significant contributions of residential wood combustion and motor vehicle emissions to organic aerosols collected during the winter in Boise, Idaho. This work was a collaboration with the EPA as part of their Integrated Air Cancer Project. The origin of these aerosols was important in terms of visibility degradation and their potential for causing human health problems, since this material may have contained mutagenic compounds. The technique used by NIST scientists includes micro-analytical methods to prepare organic extracts of fine particulate material for accelerator mass spectrometry (AMS) ^{14}C measurements. The group has advanced the ^{14}C AMS measurement process to a sensitivity measured in sample size to about 30 μg carbon. Through the use of NIST Standard Reference Materials, a high level of quality assurance has been established throughout the process of isolating and measuring the organic fraction.

HOW DO YOU MAKE A METAL OUT OF ATOMS?

As manmade structures become increasingly smaller and smaller, eventually we will be dealing with materials consisting of only a limited number of atoms. In this limit, where quantum effects become the rule, an intriguing question of importance in possible future atomic size devices is: will very small metal atom structures still exhibit metallic characteristics? Would they conduct electricity like normal metals? NIST scientists have found that one- and two-dimensional structures of cesium atoms on III-V semiconductor surfaces exhibit insulating characteristics; metallic properties are not present until three-dimensional structures form.

The ability to probe the electronic properties of atomic-size structures has become possible with the spectroscopic abilities of the scanning tunneling microscope. This ability allows NIST scientists to examine with atomic resolution the spectrum of electronic states near the Fermi level of nanometer-sized structures. This is accomplished by recording the tunneling current flowing into and out of electron energy levels of a specimen as a function of voltage applied between the tip and sample; the voltage controls which energy levels are examined. Cesium structures on GaAs and InSb(110) surfaces were studied because NIST

scientists discovered that, as the cesium density increases, this system exhibits such structures as a one-atom-wide linear chain, a one-atom-thick two-dimensional film, and finally, a three-dimensional overlayer. Tunneling spectra on these structures show a band gap in the electronic energy levels that narrows in going from the linear chains to the two-dimensional film. Because of the band gaps these structures are insulating, even though they are composed of "metal" atoms. True metallic characteristics, as evidenced by an absence of a band gap, were not observed until three-dimensional structures of cesium were formed. This work has spurred experimental and theoretical interest in both U.S. and European communities in trying to understand the insulating behavior of these intriguing systems.

NIST DEVELOPS RADIOACTIVITY STANDARDS FOR HOLMIUM-166 AND RHENIUM-188 BONE-SEEKING PHARMACEUTICALS

Holmium-166 and rhenium-188 are bone-seeking radionuclides of therapeutic interest for improved cancer treatment. Holmium-166 is under investigation by drug manufacturers, while rhenium-188, available from Oak Ridge National Laboratory, is receiving considerable attention from the same manufacturers. NIST standards are needed for these nuclides to allow proper assays for the investigational new drugs and for use in new drug applications to the FDA.

Both radionuclides were standardized for activity by the method of $4\pi\beta$ liquid-scintillation efficiency tracing with tritium. For applications in nuclear medicine, the clinical users of a radionuclide require accurate half-life and decay scheme information. Therefore, measurements were made on the photon emission rates of the principal x and gamma rays using semiconductor detectors. Half lives measured with the NIST primary ionization chamber were 26.77 ± 0.01 h for the holmium-166 and 17.01 ± 0.01 h for the rhenium-188. Calibrations for both radionuclides are now available from NIST.

TENSILE CREEP OF SILICON NITRIDE

High-temperature deformation is one of the important properties in determining lifetime and reliability of advanced ceramics. NIST researchers have recently completed the tensile creep data collection process on a candidate heat engine silicon nitride. Substantial transient effects have been observed. A creep rupture analysis shows all the data fit a Monkman-Grant type curve, wherein the time-to-failure depends only on the creep rate. Preliminary lifetime predictions indicate this material will withstand an applied stress of approximately 50 MPa for the DOE-prescribed temperature and time limits of 1370 °C for 1 year. Analysis of all the data currently available on silicon nitride shows that it also fits the Monkman-Grant formulation and, in fact, fits within a narrow band, regardless of composition or the presence of reinforcement such as whiskers.

COMPUTER MODEL FOR THE ANALYSIS OF CORROSION MEASUREMENTS

Scientists at NIST have developed a computer model for the analysis of electrochemical corrosion measurements. Currently, curve fitting to an equivalent electric circuit is usually employed for analysis, but a physical interpretation of the electric circuit in terms of the corrosion processes simulated is often difficult. For the NIST model, a system of equations describing the physical processes is used to calculate the response of the electrochemical cell to the perturbations imposed by an experiment. The curves calculated by the computer can be compared to the experimental results. The main advantage of this approach is that interpretation is straightforward, allowing a direct assessment of the processes that determine corrosion behavior. This work is continuing with the aim of treating more complicated systems.

MAGNETIC SEMICONDUCTORS SUPERLATTICES

A research effort involving researchers from NIST and the University of Notre Dame, in which neutron diffraction has been used for the study of magnetic semiconductors, has shown for the first time the dependence of the crystalline and magnetic structure on epitaxial strain in alternating layer superlattices of MnSe with ZnSe and MnSe with ZnTe. In these materials the zincblende crystalline structure, which does not exist in the bulk, is stabilized. It has been discovered that when MnSe is grown with ZnTe, the epitaxial strain "tunes" the tetragonal distortion so that new incommensurate antiferromagnetic structures are stabilized. In addition, the strain alters the first-order phase transition in MnTe so that it becomes second order. These materials are of technological interest for electro-luminesence and tunable infrared sensors.

COMBUSTION IN MICROGRAVITY

Researchers at NIST recently developed a theoretical model and computer simulation of ignition and flame spread over a thin paper in a microgravity environment. The model consists of complex heat and mass transport, flow motion, and chemical reactions in the gas phase and in the condensed phase. At present such calculations are only feasible without gravity. The results indicate that ignition can occur in air but there is no transition to flame spread. The transition can occur in 35 percent oxygen concentration or higher. This project is supported by NASA's Microgravity Science Program and the results can be used for fire safety in the Shuttle and in the Space Station.

NIST INVESTIGATES DESIGN OF HEAT EXCHANGERS FOR AIR CONDITIONERS WITH REFRIGERANT MIXTURES

NIST researchers built and incorporated into a central residential air-to-air heat pump a new type of refrigerant-to-air heat exchanger designed specifically for refrigerant mixtures. In tests of the full heat pump, efficiency was increased by 14 percent compared to the unit with the conventional R-22 refrigerant. NIST scientists had previously demonstrated more than 30 percent increases in efficiency with similar equipment and refrigerant mixtures; however, all previous research focused on refrigerant-to-water heat exchangers where counterflow is accomplished easily to take advantage of the temperature changes when refrigerant mixtures boil or condense at constant pressure. The challenge for conventional air-to-air heat pumps that predominate the residential market was to accommodate the counterflow design without large pressure drops in the air stream and the attendant large fan power required. A cross-counterflow design was used along with the binary mixture R22/R114. A modified design procedure was also developed and verified in the course of this research to allow equipment to properly size these types of heat exchangers in future equipment.

NIST DEVELOPS INTERDISCIPLINE INTEGRATION PROCEDURES FOR ISO STANDARDS IN PRODUCT DATA EXCHANGE

A NIST scientist recently published integration procedures used in the specification of product data requirements for computer-aided information systems. Traditionally, the specification of product data requirements has been segregated by discipline boundaries inhibiting the exchange of information and impeding such approaches to product development as concurrent engineering. Providing an integrated specification of generic product data requirements establishes the foundation upon which data exchange and sharing are possible both across disciplines and throughout the life cycle of a product. The integration procedures enable the users of computer-aided systems to contribute to the development of an integrated specification that accommodates their common data requirements. The integration procedures have been adopted by the Project Management Advisory Group of the International Organization for Standardization (ISO) Technical Committee on Industrial Automation Systems (TC 184) Subcommittee on Manufacturing Data and Languages (SC4) responsible for ISO 10303, informally referred to as STEP (Standard for the Exchange of Product Data).

ADVANCED ACCESS CONTROL SYSTEM DEVELOPED

NIST has developed the Smartcard Access Control System (SACS), which provides more effective security than conventional password-based access control. SACS requires users to have a valid smartcard and a correct personal identification number to be granted network access.

SACS incorporates advanced features, including cryptographic authentication, secure data storage, support for automatic key distribution, and the capability to generate message authentication codes. SACS is the first smartcard application designed through use of formal verification, a mathematical technique which demonstrates conformance to a specification with much greater precision than conventional software verification techniques.

Volume 96, Number 3, May–June 1991
Journal of Research of the National Institute of Standards and Technology

NIST HOSTS WORKSHOP FOR IMPLEMENTORS OF OPEN SYSTEMS INTERCONNECTION (NIST/OSI WORKSHOP)

On March 11-15, NIST and the IEEE Computer Society co-sponsored the NIST/OSI Workshop, part of a continuing series to develop implementation specifications from international standard design specifications for computer network protocols. The workshop attracted around 300 participants.

Highlights of the workshop included a proposal for a new special interest group on conformance testing and the review of a European document on that subject. Minor changes were made to the text of the stable implementation agreements document in the areas of X.400, message handling systems; file transfer, access, and management; and the lower layers of the OSI Reference Model. In the working document, new text was generated in the areas of transaction processing, network management, and office document architecture.

OMNITAB 80 SOFTWARE RELEASED FOR PUBLIC DISTRIBUTION BY NIST

OMNITAB 80 Version 7.00 maintained by NIST was delivered to the National Technical Information Service for public distribution. The software package consists of procedures written in FORTRAN source language and test problems to assist in implementing OMNITAB on any large mainframe computer. The OMNITAB system is operational at NIST on the Cyber 855 on the SUN and SUN-SPARC workstations and will be available on the new Cray. A users' manual and complete documentation are provided.

OMNITAB 80 Version 7.00 is a highly integrated general-purpose programming language and statistical software computing system. Conceived in the early 1960s by NBS and implemented in collaboration with the NBS Computation Laboratory, OMNITAB was the first data analysis system based on the idea of worksheets (spreadsheets). It has been maintained and moderized by Statistical Engineering since about 1966. The system provides statistical analysis, numerical analysis, matrix/array analysis, and plotting capabilities. The user can also perform simple and complex arithmetic and trigonometric calculations, data manipulation, and special function calculations.

Standard Reference Materials

NEW STANDARD AVAILABLE FOR METAL PRODUCERS

Standard Reference Material (SRM) 347, Magnesium Ferrosilicon, is a new standard in a group of ferroalloy SRMs developed by NIST for quality assurance in the steel industry. The ferroalloy is primarily used in the production of ductile iron. The new standard is for evaluating chemical and instrumental methods of analyses and for calibrating instruments. SRM 347, in the form of a powder (0.1 to 0.2 mm) with certified values for 15 elements, including rare earths, is available for $112 per 100 g unit from the Standard Reference Materials Program, Room 204, Building 202, NIST, Gaithersburg, MD 20899, 301/975-6776, fax: 301/948-3730.

U.S. COMPANY TO MARKET NIST SRM

A leading manufacturer of analytical laboratory instrumentation, has recently agreed to market Standard Reference Material (SRM) 869, Column Selectivity Test Mixture for Liquid Chromatography (Polycyclic Aromatic Hydrocarbons), as part of its product line.

SRM 869 is a result of research carried out over the past 5 years at NIST. This SRM contains three chemical probes that provide a sensitive measure of column selectivity. These probes belong to a class of compounds known as polycyclic aromatic hydrocarbons, and their retention behavior correlates strongly with subtle changes in bonding chemistry used in column preparation.

SRM 869 provides researchers with a much needed tool for assessing differences in generically identical liquid chromatographic columns. Such inconsistencies in column performance often make it difficult to reproduce published results. Chromatographic tests utilizing SRM 869 permit columns to be classified into groups with similar retention behavior, thus facilitating column selection. Because SRM 869 provides an indication of overall column selectivity, a single experiment can indicate the suitability of a specific column for the analysis of various complex mixtures, thus increasing laboratory productivity by reducing methods development time.

NEW SRMs FOR THE MARINE ENVIRONMENT DEVELOPED

Two new Standard Reference Materials (SRMs) have been issued to support the measurement of pollutants in the marine environment. SRM 1941, Organics in Marine Sediment, and SRM 1974, Organics in Mussel Tissue (*Mytilus edulis*), were developed at the request of both state and federal agencies involved in marine environmental monitoring to provide an accuracy base for their national marine monitoring programs. Sediment collected in the Baltimore harbor was air dried, pulverized, and sieved to provide a homogeneous material for distribution as SRM 1941. For the preparation of SRM 1974, approximately 2,400 mussels collected from the Boston harbor were cryogenically pulverized and homogenized. Since SRM 1974 was prepared as a fresh frozen homogenate, it is representative of the tissue matrix generally used in analytical procedures for marine samples and represents the first fresh frozen biological SRM issued by NIST.

Both SRMs were analyzed using gas chromatography-mass spectrometry and liquid chromatography with fluorescence detection to provide certified values for the concentrations of selected polycyclic aromatic hydrocarbons (PAHs). Additional measurements by gas chromatography provided noncertified values for polychlorinated biphenyl (PCB) congeners and chlorinated pesticides such as the DDTs.

Even though the primary use of both SRMs is for the measurement of organic contaminants, inorganic analysts will also find them useful as natural matrix trace element reference materials. Noncertified concentrations for over 30 trace elements were determined at NIST primarily using neutron activation analysis. SRMs 1941 and 1974 are the first environmental matrix SRMs that have NIST-assigned concentrations for PAHs, PCB congeners, chlorinated pesticides, and inorganic constituents on the same material.

NEW PHOTOMASK LINEWIDTH STANDARD

NIST has issued a new bright-chrome on glass photomask linewidth standard SRM 476. This new standard is a companion to the older antireflecting-chrome SRM 475, which has been available for almost a decade. It consists of a series of calibrated lines and spaces ranging from a nominal 0.9 to 10.8 μm as well as a series of pitch (i.e., center-to-center) patterns designed for the microelectronics industry in calibrating their optical instruments that measure the critical dimensions on the photomasks used to pattern integrated circuits and devices. A submicrometer antireflecting-chrome photomask standard, designated SRM 473, is also being developed that will extend this range down to 0.5 μm and up to 30 μm. All of these photomask standards are calibrated on a new automated system that has improved the precision and accuracy of the calibration and reduced the calibration time for each photomask from 2 weeks (for SRM 475 on the old manual system) to less than one 24-hour day (on the new automated system).

CERTIFICATION OF A NEW NIST TOTAL DIET STANDARD REFERENCE MATERIAL

As a result of growing public demands for food quality surveillance programs, there is an increased need for reference materials to improve the quality of chemical analysis for foods. A recent example of this is the recent legislation in the United States on nutrition labelling. In addition, nutritional considerations such as bioavailability are resulting in new design of reference materials requiring specialized kinds of foods. To meet some of these needs, NIST has just released Standard Reference Material (SRM) 1548, Total Diet. This SRM was prepared with foods obtained from collections of the U.S. Food and Drug Administration's Total Diet Study, which represent that consumed by a 25-year-old American male. The foods were weighed, composited, freeze-dried, blended, radiation sterilized, and packaged in bottles. This material has now been certified for fat, protein, (Kjeldahl) nitrogen, ash, cholesterol, fiber and caloric content (bomb calorimetry), sulfur, phosphorous, chlorine, sodium, potassium, calcium, magnesium, iron, zinc, manganese, copper, selenium, and cadmium. Information values are available for additional trace elements, including tin, aluminum, boron, lithium, nickel, molybdenum, lead, and rubidium. Additional reference materials to address future needs in the field are under consideration.

Standard Reference Data

NEW MOLTEN SALTS DATABASE AVAILABLE FOR PC USERS

The new molten salts database, designed for personal computers (PCs), provides rapid access to approximately 750 compilations on the properties of 320 inorganic salts in the molten state, including

density, surface tension, electrical conductance, and viscosity. This information will be helpful for researchers engaged in the development of new high-temperature advanced materials for aerospace products and for scientists performing high-temperature and high-pressure physical property measurements. Information can be obtained by chemical formula or by browsing through the collection of carefully evaluated data to find the system of choice. NIST Standard Reference Database 27, Thermodynamic and Transport Properties of Molten Salts, is available for $190 and is designed for use on any AT- or XT-Class PC with a color monitor. The information occupies 250 kilobytes, which can be stored on a hard disk. To order, contact Standard Reference Data Program, A320 Physics Building, NIST, Gaithersburg, MD 20899, 301/975-2208, fax: 301/926-0416.

STANDARD REFERENCE DATA PRODUCTS CATALOG UPDATED

The NIST Standard Reference Data Products 1991 Catalog (NIST SP 782, 1991 Ed.) provides scientists and design engineers with the latest information on the data computations, publications, and computerized databases available from the NIST Standard Reference Data Program and other sources. Critically evaluated data compilations are available in the following areas: analytical chemistry, atomic physics, chemical kinetics, materials properties, molecular structure and spectroscopy, thermodynamics and thermochemistry, and the thermophysical properties of fluids. Since 1968, the NIST Standard Reference Data Program has been responsible, under an act of Congress, for coordinating on a national basis the evaluation of numerical data in the physical sciences. The evaluation of chemical and physical properties of substances and materials is carried out in the National Standard Reference Data network of data centers. To obtain a copy of SP 782, 1991 Ed., send a self-addressed mailing label to Standard Reference Data Program, A320 Physics Building, NIST, Gaithersburg, MD 20899, 301/975-2208.

Volume 96, Number 3, May–June 1991
Journal of Research of the National Institute of Standards and Technology

Calendar

July 15–18, 1991
FIFTH INTERNATIONAL CONFERENCE ON LIQUID ATOMIZATION AND SPRAY SYSTEMS (ICLASS '91)

Location: National Institute of Standards and Technology
Gaithersburg, MD

Purpose: To present the state of the art in experimental and computational techniques related to all aspects of atomization processes and sprays used in industrial, transportation, and agricultural systems.
Topics: Spray formation; instrumentation techniques; modeling of sprays and spray flames; agricultural sprays; spray coatings; metal atomization; and atomization processes in gas turbines, internal combustion engines, and furnaces/boilers.
Format: Plenary lecture, invited and contributed papers, poster session, and instrumentation and equipment exhibits.
Audience: National and international technical community involved in the application of atomization and sprays to industrial, transportation, and agricultural systems.
Sponsors: NIST and ILASS-Americas (Institute of Liquid Atomization and Spray Systems).
Contact: Hratch Semerjian, B312 Physics Building, NIST, Gaithersburg, MD 20899, 301/975-2609.

July 30–31, 1991
JOINT U.S./GERMAN CONFERENCE ON NEW DIRECTIONS FOR OPERATIONS RESEARCH IN MANUFACTURING

Location: National Institute of Standards and Technology
Gaithersburg, MD

Purpose: To provide a forum for participants to discuss new directions for operations research in manufacturing.
Topics: Distributed real-time scheduling, hierarchical and heterarchical control systems, integrated algorithms for design, process planning, and equipment level programming.
Format: Open forum.
Audience: Universities and manufacturers.
Sponsors: NIST, Operations Research Society of America Special Interest Group on Manufacturing, German Operations Research Society, and National Science Foundation.
Contact: Al Jones, A319 Metrology Building, NIST, Gaithersburg, MD 20899, 301/975-3554.

October 1–4, 1991
14th NATIONAL COMPUTER SECURITY CONFERENCE

Location: Omni Shoreham Hotel
Washington, DC

Purpose: To share information with government and the private sector on present and future technologies on "Information Systems Security: Requirements and Practices."
Topics: Systems application; criteria, evaluation and certification; management and administration; international computer security activities; innovations and new products; awareness, training, and education; disaster prevention and recovery; and privacy and ethical issues.
Format: General session and workshops.
Audience: Federal agencies, industry, and academia.
Sponsors: NIST and the National Computer Security Center, NSA.
Contact: Irene Gilbert, A216 Technology Building, NIST, Gaithersburg, MD 20899, 301/975-3360.

October 14–18, 1991
THIRD INTERNATIONAL SYMPOSIUM ON ESR DOSIMETRY AND APPLICATIONS

Location: National Institute of Standards and Technology
Gaithersburg, MD

Purpose: To focus on current applications of electron spin resonance (ESR) spectroscopy.
Topics: Ionizing radiation dosimetry, including reference and transfer dosimetry, archeological dating, geology, solid-state effects, instrumentation, imaging, medical applications.
Format: Symposium.
Audience: ESR spectroscopists, health physicists, representatives from industry, government, and academia.
Sponsors: NIST, Department of Energy, China University of Science and Technology, and International Atomic Energy Agency.
Contact: Marc F. Desrosiers, C214 Radiation Physics Building, NIST, Gaithersburg, MD 20899, 301/975-5639.

October 22–24, 1991
GAGE BLOCK CALIBRATION

Location: National Institute of Standards and Technology
Gaithersburg, MD

Purpose: To emphasize the concepts, techniques, and apparatus used in gage block calibration.
Topics: Statistics, process control, gage block comparators, interferometry.
Format: Short instructional course.
Audience: Metrology laboratory managers and technicians.
Sponsor: NIST.
Contact: John Stoup, A107 Metrology Building, NIST, Gaithersburg, MD 20899, 301/975-3471.