THEORETICAL STUDIES OF HIGH-POWER ULTRAVIOLET AND INFRARED MATERIALS

Marshall S. Sparks, et al

Xonics, Incorporated

Prepared for:

Defense Supply Service
Advanced Research Projects Agency

6 December 1974

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author)
Xonics, Incorporated
6837 Hayvenhurst Avenue
Van Nuys, California 91406

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

AD/A009256

3 REPORT TITLE
THEORETICAL STUDIES OF HIGH-POWER ULTRAVIOLET AND INFRARED MATERIALS

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)
Fourth Technical Report, 30 June 1974 through 6 December 1974

5 AUTHOR(S) (First name, middle initial, last name)
Marshall S. Sparks, Carl J. Duthler

6 REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
6 December 1974	22/	107

8a. CONTRACT OR GRANT NO	9a. ORIGINATOR'S REPORT NUMBER(S)
DAHC15-73-C-0127, P00002	
b. PROJECT NO	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	ARPA Order No. 1969, Amendment No. 3

10 DISTRIBUTION STATEMENT
This document may be further distributed by any holder only with specific prior approval of Defense Supply Service - Washington, D.C.

11 SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
Sponsored by Advanced Research Projects Agency	Defense Supply Service - Washington Room 1D245, The Pentagon Washington, D.C. 20310

13 ABSTRACT

The previous parametric-instability explanation of the jump in the Stokes intensity I_S as a function of laser intensity I_L in stimulated Raman scattering experiments is valid for many solids, liquids, and gases (with high optical dispersion $dn/d\lambda$ and high Raman frequency ω_f), while a few materials (low $dn/d\lambda$ and ω_f) should show an enhanced gain, $I_S \sim \exp I_L^2$, which is greater than the usual stimulated-Raman gain but not as great as the jump result. Previously anomalous experimental results, which show both types of behavior, are explained. For high dispersion the steady state is not reached until a time much greater than the Raman phonon relaxation time. A phonon parametric instability studied previously in another context can reduce the Stokes intensity. Experiments to detect the phonon instability are suggested. In the earlier golden-rule analyses (perturbation theory treatment of occupation numbers) the enhancement was lost by neglecting the increase of the vibrational amplitudes above their thermal equilibrium values. Even though the probability of an individual ion or molecule being excited is small, the occupation number of the phonon in the Raman process is large. In the previous mode-amplitude analyses, the enhancement was lost in the method of linearizing the nonlinear differential equations. By solving these same mode-amplitude equations without using the previous linearization scheme, the enhancement is obtained and the equivalence of the mode-amplitude and golden-rule Boson-occupation-number analyses is demonstrated explicitly. The loss of phase information in using the occupation numbers is unimportant. Analysis indicates that Raman active crystals fail at intensities I_f which are greater than the value I_R of I_L at the jump by an amount I_T that is generally of the order of or less than I_R. A typical value of I_R at the ruby frequency is a few GW/cm^2, which is less than other intrinsic-mechanism thresholds. The temperature rise from the phonons generated near the sample surface is sufficient to cause material failure when $I > I_f$. Expressions are derived for the rate at which radiation is scattered

DD FORM 1 NOV 65 1473

Unclassified

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
ultraviolet absorption ultraviolet windows stimulated-Raman-scattering instability optical materials theoretical study molecular-ion-impurity absorption Raman gain enhancement laser-materials damage laser heating of metals surface plasmon excitation aluminum absorptance aluminum oxide films infrared absorption ultraviolet reflectors						

ABSTRACT (Continued)

and absorbed because of surface roughness on a semi-infinite material in the presence of a dielectric overlayer, which can greatly enhance the radiation to surface-plasmon coupling. A formalism developed previously is utilized. A series of numerical calculations which explore the roughness-induced scattering and absorption of uv radiation for aluminum overcoated by aluminum oxide is presented. The reflectivity dip produced by roughness-induced coupling to the surface plasmon shifts toward the visible as the thickness of the oxide layer increases. The size of the dip is controlled strongly by the degree of correlation between the roughness on the vacuum-oxide interface, and that on the oxide-substrate surface. The temperature rises T of laser-irradiated metals have drastically different magnitudes and dependences on material parameters such as thermal conductivity K and heat capacity C, on the sample thickness ℓ and lateral dimension L, and on the operating conditions such as the laser-beam diameter D and pulse duration t. The behavior depends on the relative magnitudes of D, ℓ, L, and the thermal diffusion distance d. The highest of the recently measured copper damage thresholds of 125 to 750 J/cm^2 for 0.6 μsec pulses at 10.6 μm are likely to be quite near the intrinsic limit set by the simple process of melting from the intrinsic absorption. The intensity at which the cavity mirrors of recently developed xenon uv lasers fail is explained. The theoretical value of T for metals irradiated for 20 seconds with 10.6 μm radiation is $\sim$100 too small to explain recent experimental damage results, the discrepancy probably being related to plasma ignition at the sample surface. The steady-state value of T for cooled metals is not reduced substantially by increasing the cooling efficiency past a certain point. For $t < \tau$, where the characteristic time τ depends on both the cooling and ℓ, cooling the metal is not effective in preventing the temperature rise.

THEORETICAL STUDIES OF HIGH-POWER ULTRAVIOLET AND INFRARED MATERIALS

M. Sparks, Principal Investigator, 213/787-7380
C. J. Duthler, Principal Scientist, 213/787-7380

Xonics, Incorporated
Van Nuys, California 91406

Fourth Technical Report
6 December 1974

Contract No. DAHC15-73-C-0127
Effective Date of Contract: 7 December 1972
Contract Expiration Date: 6 December 1974

Prepared for
Defense Supply Service - Washington, D.C.

Sponsored by Advanced Research Projects Agency
ARPA Order No. 1969, Amendment No. 3; Program Code No. 4D10

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Defense Supply Service - Washington, D.C. under Contract No. DAHC15-73-C-0127. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U.S. Government.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont'd)

LIST OF ILLUSTRATIONS (Cont'd)

PREFACE

This Fourth Technical Report describes the work performed on Contract Number DAHC15-73-C-0127 on Theoretical Studies of High-Power Ultraviolet and Infrared Window Materials during the period from June 30, 1974 through December 6, 1974. As in the past, previously reported results are not repeated in the present report. The work on the current contract is a continuation of that of the previous Contract Number DAHC15-72-C-1029.

The following investigators contributed to this report:

Dr. C. J. Duthler, principal research scientist

Dr. A. A. Maradudin, consultant, University of California, Irvine, California

Dr. D. L. Mills, consultant, University of California, Irvine, California

Dr. L. J. Sham, consultant, University of California, San Diego, California

Dr. M. Sparks, principal investigator .

In preparation for presentations at the Advanced Research Project Agency's Materials Research Council Studies of High Energy Laser Mirror Materials and of Erosion of Infrared Windows at La Jolla, California in July 1974, the following interesting results were obtained: (a) In some cases, the infrared laser damage to mirrors can be explained by the simple process of ordinary intrinsic absorption, with the temperature dependence of the intrinsic absorptance A determined from the relation $A \sim \rho$, where ρ is the electrical resistivity. Thus, it is not necessary to invoke nonlinear processes as thought when the meeting was arranged. In other cases, the laser ignition of plasmas at the surface obscures the ordinary heating effect. (b) Calculations of the temperature rise in

laser-irradiated metals and transparent materials yield important results that should be useful in mirror, metal target, and window-materials tests. (c) A model of surface absorption with a bulk and surface contribution to the electron relaxation frequency was proposed to explain the anomalous temperature dependence of A observed at China Lake. Subsequent investigations of the model turned up difficulties in the explanation, but indicated that further studies would be promising. The results from the surface model may also be related to the problem of laser ignition of plasmas at solid surfaces.

The mirror-damage results and the results for the temperature rise in laser-irradiated metals are included in Sec. F of the present report. The case of transparent material will be included in a subsequent report, as will a brief report of the surface-absorption model. At the suggestion of Dr. M. Hass, our previous analysis of molecular-ion absorption was applied explicitly to the case of potassium bromide (Sec. II). The emphasis in the remainder of the program was on ultraviolet materials.

A. INTRODUCTION AND SUMMARY

Work on the enhancement in the stimulated Raman scattering process, which was discussed briefly in the previous technical report, has been completed. The final results are contained in the present report. In addition to a more complete description of the former results, three new results are reported: (1) It is shown in Sec. B that materials with low optical dispersion and small Raman frequencies ω_f can exhibit an intermediate Stokes gain that is lower than the strong enhanced gain that gives rise to a jump in Stokes intensity I_S as a function of laser intensity I_L, but greater than the previous linear Raman gain. For this intermediate-gain case, $I_S \sim \exp I_L^2$. Experimental results, which show both types of behavior, are now explained. For high dispersion, the steady state is not reached until a time much greater than the Raman phonon relaxation time. A phonon parametric instability studied before in another context can reduce the Stokes intensity. Experiments to detect the phonon instability and measure the magnitude and temperature dependence of "half-frequency" phonons Q (with frequencies $\omega_Q + \omega_{-Q} = \omega_f$) are suggested.

(2) The reasons for the absence of the enhancement in other treatments of Raman amplification were discovered (see Sec. C). In the earlier golden-rule analyses (perturbation theory treatment of occupation numbers), the enhancement was lost by neglecting the increase of the vibrational amplitudes above their thermal equilibrium values. Even though the probability of an individual ion or molecule being excited is small, the occupation number of the phonon in the Raman process is large. In the previous mode-amplitude analyses, the enhancement was

lost in the method of linearizing the nonlinear differential equations. By solving these same mode-amplitude equations without using the earlier linearization scheme, the enhancement is obtained and the equivalence of the mode-amplitude and golden-rule Boson-occupation-number analyses is demonstrated explicitly. The loss of phase information in using the occupation numbers is unimportant.

(3) In Sec. D an analysis of laser damage to materials by this enhanced stimulated Raman scattering mechanism shows that the damage threshold for Raman active crystals may be lower than the threshold for other mechanisms such as electron avalanche and self focusing.

As mentioned in the Preface, it was discovered in preparing for presentations at the Advanced Research Projects Agency Materials Research Council Studies at La Jolla, California in July, 1974 that in some cases the infrared laser damage to mirrors can be explained by the simple process of ordinary intrinsic absorption. One group was already aware of this result, but apparently it was not generally realized that it was not necessary to invoke nonlinear effects to explain the damage threshold. The increase in absorptance with increasing temperature and consideration of the 0.2 μsec peak in the nominal 10.6 μsec pulse are important in obtaining agreement between the theory and the Hughes experimental results.

At the conference concern was expressed that the experimental results of Saito, Charlton, and Loomis were different from those of the Hughes Malibu group, particularly with respect to differences between copper and molybdenum. We suggested that the differences could be related to the fact that for the 0.6 μsec pulses in the Hughes experiments the thermal diffusion distance d was much less than

the sample thickness ℓ, while $d >> \ell$ was satisfied for the 20 sec pulses in the experiments of Saito, Charlton, and Loomis. The differences in the volumes that were heated means that the temperatures in the two cases will have different dependences on the thermal conductivity K and heat capacity C. It was found that the simple procedure of using an effective heated volume to find the temperature rise requires extreme care in general. Fortunately, an exact simple solution was obtained for the maximum temperature rise in a realistic model that includes the most difficult case of $d >> \ell$. These results were combined with previously unpublished results to form a relatively complete treatment of laser heating of metals in Sec. F.

The theoretical results appear to explain the Hughes experimental results. However, they clearly show that the theoretical value of the temperature in the experiments of Saito and coworkers is two orders of magnitude too small to explain the experimental damage results. It appears that ignition of a plasma at the metal surface is responsible for the difference. Until this plasma ignition result is understood, the differences between results for copper and molybdenum cannot be considered as significant. In particular, the different dependence of T on K and C for the two cases of $d >> \ell$ and $d << \ell$ cannot be used to explain the present differences.

The theoretical intensity at which the cavity mirrors of recently developed xenon uv lasers fail is in good agreement with the experimental values. Further results of the theory are as follows: The steady-state value of T for metals cooled with a surface-heat-transfer coefficient h is not reduced substantially by increasing the cooling efficiency past a certain point ($h > h_\ell \equiv K/\ell$).

For $t < \tau$, where the characteristic time τ depends on both ℓ and h, cooling the metal is not effective in preventing the temperature rise. The temperature rises T of laser-irradiated metals have drastically different magnitudes and dependences on material parameters such as thermal conductivity K and heat capacity C, on the sample thickness ℓ and lateral dimension L, and on the operating conditions such as the laser-beam diameter D and pulse duration t. The behavior depends on the relative magnitudes of D, ℓ, L, and the thermal diffusion distance d.

The long-standing problem of the anomalously large absorption by some oxidized aluminum samples (the absorption by Al_2O_3 is not sufficiently strong to explain the great absorption) motivated the study of the effect of a dielectric overlayer on a rough aluminum sample. See Sec. E. In some samples the presence of a layer of Al_2O_3 on a smooth aluminum surface may be sufficient to cause the measured absorption, as recently realized by H. Ehrenreich (private communication) and by Maradudin and Mills in the previous technical report. In other samples, the absorption is too great to be explained by this result. A preliminary study in which the results of Sec. E are applied directly to this problem suggests that a rough Al_2O_3 layer should give rise to the greater observed absorption, while a smooth Al_2O_3 layer should give rise to the lesser observed absorption. This will be discussed in detail in a future report.

In the previous technical report an estimate of the strength of the two-photon absorption process in transparent materials indicated that this process may have the lowest threshold of the known failure mechanisms. In view of this importance of the two-photon absorption process and the fact that the accuracy of the theoretical estimate may be as low as two orders of magnitude, an improved estimate is

important. Unfortunately, even the one-photon absorption process is not well understood. In fact, even the explanation of absorption edges and the values of the energy gaps in the simple alkali halides are currently controversial. It appears that a relatively simple tight binding calculation may resolve at least some of the difficulties. The final analysis of the two-photon absorption has therefore been delayed for a short time in order to consider some of the fundamental problems whose answers are needed before a completely satisfactory treatment of the two-photon problem can be undertaken. The approach to the one-photon absorption problem is sufficiently simple that it appears that it can be used in conjunction with techniques from our previous treatment of magnetism and phonon problems to obtain information on states in the gap of wide bandgap materials.

Two papers presented at the Advanced Research Projects Agency Fourth Laser Window Conference at Tucson, Arizona in November, 1974 are included in Secs. G and H. Sec. H is an extension of the previous treatment of molecular-ion-impurity infrared absorption (second technical report). It is emphasized that the absorption frequencies result from the internal modes of the molecular-ion impurity which are shifted and broadened only slightly by the particular host crystal. Consequently the previous analysis, which emphasized KCl, is directly applicable to KBr and is approximately valid for other ionic crystals containing halide ions. The absorption frequencies of several molecular-ion impurities are given and it is suggested that the overlapping absorption lines from several polyatomic ions, all of which contain oxygen, could account for the broad previously unidentified absorption peak near 9.5 μm.

B. STIMULATED RAMAN SCATTERING: ENHANCED STOKES GAIN AND EFFECTS OF ANTI-STOKES AND PARAMETRIC PHONON PROCESSES

M. Sparks
Xonics, Incorporated, Van Nuys, California 91406

It is shown that the previous parametric-instability explanation of the jump in the Stokes intensity I_S as a function of laser intensity I_L in stimulated Raman scattering experiments is valid for many solids, liquids, and gases (with high optical dispersion $dn/d\lambda$ and high Raman frequency ω_f), while a few materials (low $dn/d\lambda$ and ω_f) should show an enhanced gain, $I_S \sim \exp I_L^2$, which is greater than the usual stimulated-Raman gain but not as great as the jump result. The previously anomalous experimental results of Grun, McQuillan, and Stoicheff and of others, which show the jump in I_S, and of Hagenlocker, Minck, and Rado, which show both types of behavior, are explained. The instability is expected to be important in laser damage of Raman active crystals and possibly in determining the limiting diameter of self-focused beams. The transient solution for the case of high dispersion indicates that the steady state is not reached until a time much greater than the Raman phonon relaxation time. It is also shown that a phonon parametric instability studied previously in another context can reduce the Stokes intensity. Experiments to detect the phonon instability and measure the magnitude and temperature dependence of "half-frequency" phonons Q (with frequencies $\omega_Q + \omega_{-Q} = \omega_f$) are suggested.

I. INTRODUCTION

Stimulated Raman scattering was first observed[1] and analyzed[2] in 1962. It was recently realized[3] that a parametric instability in the Raman Stokes process causes a Stokes-intensity gain enhancement that has significant consequences including a nearly discontinuous increase, or "jump," in the Stokes intensity I_S as a function of the laser intensity I_L in the absence of self focusing and feedback. Such jumps had been observed, but were not understood. The instability is expected to be important in laser damage in solids that have first-order Raman-active modes, the damage thresholds being well below the 10^{10} W/cm^2 value typical for alkali halides, which are first-order Raman inactive. It is possible that the limiting diameters of self-focused beams may be determined by the stimulated Raman scattering process in some materials.

Another recent study[4] considered the effects of a parametric instability of phonons on nonlinear infrared absorption. In the Raman Stokes process, a laser photon L is annihilated and a Stokes photon S and a fundamental (Reststrahl) phonon f are created. These created phonons f can split into two phonons Q and -Q, and this three-phonon process is subject to the same parametric instability discussed in Ref. 4. In terms of the diagram in Fig. B1, parametric effects are involved at both vertices R and P, corresponding to the stimulated-Raman-scattering enhancement and phonon instabilities, respectively.

The physical interpretation of these instabilities is rather simple. For the Raman process, the balance of the energy put into the f phonons by the Raman process against that removed from the f phonons by relaxation is a key to the explanation. The power out by relaxation (by interaction with impurities or other phonons, for example) increases linearly with the number of phonons n_f, which is

just the condition that a relaxation time exists. On the other hand, the power into the f phonons increases nonlinearly with increasing n_f since the Raman process is a three-Boson process (which results in products of Boson occupation numbers in the expression for the power). Thus, at a critical value of the input power, which is proportional to the laser intensity, the amplitude n_f becomes very large. The value of I_S is also large since a Raman event creates one f phonon and one S photon. The critical value of I_L is denoted I_R.

Any three-Boson splitting process is potentially unstable parametrically. There are analogies between the two instabilities considered here and previous instabilities in ferromagnetic resonance[4a] (premature saturation of the main resonance, subsidiary absorption below the main resonance, and parallel pumping absorption), plasma physics,[4b] and electronic devices.[4c]

In the present paper, two distinct effects are considered. First, it is shown that the jump in I_S is expected to exist in many materials (with high optical dispersion). By contrast, in a few materials with low optical dispersion, such as hydrogen gas for the rotational Raman mode, the enhanced gain is greater than the usual gain[2] of stimulated Raman scattering, but is not sufficiently great to give a jump in I_S as a function of I_L. Second, it is shown that the phonon parametric instability has two important effects on stimulated Raman scattering.

The first effect concerns the previous observation[3] that I_L and I_S were discontinuous at the input surface of the sample when $I_L > I_R$. It was pointed out that this zeroth-order result was related to an infinite (but integrable) value of the fundamental-phonon occupation number n_f at the surface, and that the infinity and resulting discontinuities would be removed by higher-order effects such as

nonlinearities in the Raman oscillators. The parametric instability in the f phonons is one such effect that removes the discontinuities. Two central results of the phonon-parametric instability study[4] are that the fundamental-mode occupation number n_f is limited to values below a critical value n_p and that there is an effective relaxation frequency $\tilde{\Gamma}_f$ of the fundamental phonon that is greater than the usual linear value Γ_f as a result of an increase in the occupation numbers n_Q of the phonon modes into which f decays. Consequently, the discontinuity obviously is removed since n_f is limited to a value below the threshold value n_p, thus removing the infinity in n_f. For $I_L \geq I_R$, the spatial build up of I_S depends on the phonon instability. In gases, the nonlinearities in the oscillations of the individual molecules can limit the amplitudes of oscillation, in analogy with the phonon-instability limit in solids.

The second effect of the phonon-parametric instability is that in the rather rare case in which the phonon instability occurs at $I_L < I_R$, the Stokes intensity is reduced. Since the enhanced Stokes gain is a result of an increase in the occupation numbers n_f of the fundamental phonon, and both the enhanced and ordinary gain depend on $\tilde{\Gamma}_f$, the phonon instability can reduce both the ordinary and enhanced stimulated Raman scattering by reducing the value of n_f and increasing the value of $\tilde{\Gamma}_f$. It is suggested in Sec. VII that experiments in which this reduction of I_S is observed should afford a method of measuring the magnitude and temperature dependence of the lifetimes of the phonons well away from symmetry points in the Brillouin zone.

Next, consider the consequences of the anti-Stokes process, in which a laser photon and a fundamental phonon are annihilated and an anti-Stokes photon is created. The jump in I_S is related to the large number of f phonons created in the

Raman process, as discussed above. The anti-Stokes process in which these created phonons are annihilated can conserve wave vector and energy (within the linewidth for the process) for forward scattering if the optical dispersion $dn/d\lambda$ is low, but not if it is high. Thus, for low optical dispersion, the anti-Stokes process reduces the value of n_f, thereby reducing the Stokes gain to a value ($\sim \exp I_L^2$ roughly) intermediate between the fully enhanced value and the exponential value ($\sim \exp I_L$) of the original theories.

For high dispersion, the forward-scattering anti-Stokes process does not conserve energy and wave vector; thus, the previous enhanced-gain result, which was derived by neglecting the anti-Stokes process, is valid. For scattering at an angle θ_{mxA} from the forward direction, the anti-Stokes process can conserve energy and wave vector. Thus, as θ increases, I_L decreases and I_A increases to a peak value at θ_{mxA}, provided the aspect ratio (length to diameter) of the laser beam is sufficiently small to allow the Stokes and anti-Stokes beams to remain in coincidence with the laser beam, roughly speaking.

The transient solution of Sec. IV for the high-dispersion case shows that the steady state is not attained until a time much greater than the Raman phonon relaxation time. Only the steady state is considered in other sections.

In the analysis, the equations of motion of the occupation numbers, rather than the mode amplitudes, are considered. This greatly simplifies the analysis. A priori it is not expected that the resulting loss of phase information is important since the parametric process amplifies the thermally excited modes that have the proper phase. This is indeed the case in previous calculations.[4-6] The problem

of phases has been settled by Sparks and Wilson,[7] by rederiving the enhanced stimulated Raman scattering results with phase information included, starting from Maxwell's equations and the elastic-wave equation with coupling added in the standard way.

They have also identified the points at which the enhancement was lost in previous analyses, which used either the equations of motion of the Boson occupation numbers obtained from perturbation theory (the golden rule) or the equations of motion of the mode amplitudes (creation and annihilation operators or Fourier components and their complex conjugates of the electric field, for example). In the previous golden-rule type analyses, the enhancement was lost by neglecting the deviation of the vibrational amplitude from the thermal equilibrium value. Although n_f becomes large as mentioned above, the probability of an individual ion or molecule being excited is small, roughly speaking. Specifically, $n_f/N \ll 1$ is usually satisfied, where N is the number of unit cells or molecules. It was this fact that the individual ions or molecules are not highly excited that led to the assumption that the thermal equilibrium values were maintained in the previous analyses. There are similar results for other three-Boson processes. For example, in ferromagnetism, magnon occupation numbers are large at the threshold, while the probability of an individual electron spin being excited to the spin-down state is small.

In the previous mode-amplitude analyses, reducing the nonlinear equations to parametric linear equations (that is, linear equations with time-dependent coefficients) by assuming that the laser mode amplitude $a_L = b_L \exp(-i\omega_L t)$, where b_L is a constant, resulted in the loss of the enhancement. The same linearization scheme applied to well known magnon parametric instabilities results in the loss of

the steady-state solution, even though damping is included and a steady-state solution is expected on the basis of simple physical arguments. By solving the same equations without using this linearization scheme, the difficulties were removed and the equivalence of the mode-amplitude and golden-rule results was demonstrated explicitly for the stimulated-Raman-scattering and magnon problems.

Finally, the standard practice of considering only three waves is followed here since the analysis is greatly simplified. It can be shown that including all Stokes waves that are amplified does not change the central results.

II. HIGH-OPTICAL-DISPERSION CASE OF COUPLED STOKES PHOTONS AND FUNDAMENTAL PHONONS

The anti-Stokes process is neglected in the present section. It is shown in Sec. V that this is a good approximation in the common case of high-optical dispersion.

The enhanced stimulated Raman effect can be obtained simply from the equations of motion of the Stokes-photon and f-phonon occupation numbers n_S and n_f as follows: Using the standard perturbation-theory result $2\pi\hbar^{-2}|\langle f|\mathcal{H}|i\rangle|^2\delta(\tilde{\omega})$ for the transition rate between initial state $|i\rangle$ and final state $|f\rangle$ to calculate the rate of change of the occupation number n_f of the f phonons resulting from the Stokes process and relaxation gives[3,8]

$$dn_f/dt = \tilde{\Gamma}_f \left\{ [(n_S + n_f + 1)n_L - n_S n_f]\, n_R^{-1} - n_f + \bar{n}_f \right\} \tag{2.1}$$

where $\tilde{\Gamma}_f$ is the relaxation frequency of the f phonons (enhanced by the phonon instability),[4] $\bar{n}_f$ is the thermal-equilibrium value of n_f, and the threshold value n_R of n_L is[3]

$$n_R/V = \tilde{\Gamma}_f/\tilde{C}_R \quad , \qquad \tilde{C}_R = 2\pi|V_R|^2 V \delta(\tilde{\omega}_S)$$

where V is the volume of the sample, V_R is the Raman matrix element divided by $\hbar$, and $\delta(\tilde{\omega}_S)$, with $\tilde{\omega}_S = \omega_L - \omega_f - \omega_S$, is the line-shape factor obtained by replacing the delta function by a Lorentzian of width $\tilde{\Gamma}_f$. On resonance, that is at $\tilde{\omega}_S = 0$, the value of δ is $\delta(\tilde{\omega}_S) = 1/\pi\tilde{\Gamma}_f$. The wave-vector Kronecker delta in V_R was eliminated by the summation over Stokes modes, and S in (2.1) is the single wave-vector-conserving Stokes mode.

On resonance, $n_R \sim \tilde{\Gamma}_f^2$ as seen from the equations below (2.1); thus

$$n_R = n_{R\,nop}\,\gamma^2 \quad , \qquad \gamma \equiv \tilde{\Gamma}_f/\Gamma_f \quad , \tag{2.2}$$

where $n_{R\,nop}$ is the (linear) value of n_R with $\tilde{\Gamma}_f = \Gamma_f$, and γ is the enhancement factor for the relaxation of the fundamental mode.[4] The term $-\tilde{\Gamma}_f n_S n_f / n_R$ in (2.1) is important in the saturation region (where a substantial fraction of the laser intensity is converted to Stokes intensity).[3] This depletion of the laser beam is neglected in the present paper. In this case, the steady-state solution to (2.1) is[9]

$$n_f = \frac{\gamma^2 \bar{n}_f + I_{LR}(n_S + 1)}{\gamma^2 - I_{LR}} \quad , \tag{2.3}$$

where $I_{LR} \equiv I_L / I_R = n_L / n_{R\,nop}$ with $I_L = \hbar\omega_L c_L n_L / V$, $I_R \equiv \hbar\omega_L c_L n_{R\,nop}/V$, and c_L is the velocity of the laser beam. Eq. (2.3) reduces to the previous result[3] in the limit of $\gamma^2 = 1$.

For $I_{LR} = 0$, (2.3) gives $n_f = \bar{n}_f$, of course. The increase in the value of n_f caused by the reduction in the value of the denominator in (2.3) as I_{LR} increases is the key to the enhancement in the gain.

The gain-enhancement result and its relation to the previous results showing no enhancement can be seen explicitly as follows: The equation of motion for the number of Stokes photons, which is obtained by the same method used in deriving (2.1), is[3]

$$dn_S/dt = \beta_{I\gamma}(n_S + n_f + 1) - dn_S/\partial x \tag{2.4}$$

where $\beta_{I\gamma} \equiv \beta_c I_{LR} \gamma^{-1}$ with $\beta_c \equiv \tilde{\Gamma}_f / c_S$.

For the formal case of $n_f = \bar{n}_f$ and $\gamma = 1$, the steady-state solution to (2.4) is

$$n_S = (\bar{n}_f + 1)\,[\exp(\beta_c I_{LR} x) - 1] + n_S(0) \exp \beta_c I_{LR} x \quad . \tag{2.5}$$

For $\beta_c I_{LR} x \ll 1$ and $n_S(0) = 0$, $n_S = (\bar{n}_f + 1)\beta_c I_{LR} x$, which is the spontaneous-emission result. The last term on the right-hand side of (2.5) is the previous expression for stimulated Raman scattering,[2] that is, the exponential-gain result without enhancement. With $n_S(0) = 0$ and $\exp \beta_c I_{LR} x \gg 1$, (2.5) gives "amplification of noise" with exponential gain. In previous treatments,[2] the first term in (2.5) was missing. Thus, special arguments were required to determine the value $n_S(0)$ when there was no incident Stokes wave.

Substituting (2.3), rather than the $n_f = \bar{n}_f$ as above, into (2.4) and solving the resulting equation for the case of $\gamma^2 = 1$ gives

$$n_S = (\bar{n}_f + 1)\,[\exp(\beta_{g\,new} x) - 1] + n_S(0) \exp \beta_{g\,new} x \quad , \tag{2.6}$$

where $\beta_{g\,new} = (\Gamma_f / c_S)\, I_{LR} (1 - I_{LR})^{-1}$, which is a central result of Ref. 3. These two results (2.5) and (2.6) show explicitly that the enhancement in the gain seen in (2.6) results from the increase in n_f above its thermal equilibrium value, as already mentioned.

Notice that although n_f is large, the probability of an individual ion or molecule being excited is small, roughly speaking. Specifically, $n_f / N \ll 1$ is usually satisfied, where N is the number of unit cells or molecules. The fact that the individual ions or molecules are not highly excited led to the assumption that the thermal equilibrium values were maintained in some previous analyses. Thus the enhancement was not obtained. There are similar results for other three-Boson processes.

For example, in ferromagnetism, magnon occupation numbers are large at the threshold, while the probability of an individual electron spin being excited to the spin-down state is small, as already mentioned.

The effect of enhancement of the phonon relaxation frequency, that is, the effect of $\gamma^2 > 1$, is considered in Sec. VI. One effect of the enhancement should be mentioned here. In (2.6), $\beta_{g\,new}$ cannot become negative as a result of $1 - I_{LR}$ becoming negative. The reason is that (2.6) is valid only for the case of $\gamma^2 = 1$, which restricts I_{LR} to values less than unity. For the case of $I_{LR} > 1$, the value of I_S is determined by solving (2.1) and (2.4), with the nonlinear term $-\tilde{\Gamma}_f n_S n_f / n_R$ added, and using the values of γ^2 from Sec. VI.

III. LOW-OPTICAL-DISPERSION CASE OF COUPLING OF S, A, AND f MODES

In this section the effect of including the anti-Stokes contribution $(dn_f/dt)_A$ to the equation of motion of n_f is considered. In Sec. V it is shown that this A-f coupling must be included when the optical dispersion is low. Using the same perturbation-theory method used in deriving (2.1) gives

$$(dn_f/dt)_A = 2\pi |V_R|^2 \delta(\tilde{\omega}_A)[(n_L + n_f + 1) n_A - n_L n_f] \quad , \tag{3.1}$$

where n_A is the occupation number of the anti-Stokes mode and $\tilde{\omega}_A = \omega_L + \omega_f - \omega_A$. This contribution (3.1) must be added to the right-hand side of (2.1), which gives $dn_f/dt = 2|V_R|^2 \tilde{\Gamma}_f^{-1} \tilde{n}_{SA}$, where

$$\begin{aligned} n_{SA} &= (n_f + n_S + 1) n_L + (n_f + n_L + 1) n_A - n_L n_f \\ &\cong (n_S + n_A + 1) n_L \quad . \end{aligned} \tag{3.2}$$

The cancellation of the first term $n_f n_L$ and the last term $-n_f n_L$ removes the formal divergence in n_f in (2.3), thereby reducing the enhancement in the Stokes gain. Specifically, the steady-state solution to this equation for dn_f/dt is

$$n_f = \bar{n}_f + I_{LR} \gamma^{-2} (n_S + n_A + 1) \quad . \tag{3.3}$$

Substitution into (2.4) yields

$$dn_S/dx = \beta_c I_{LR} \gamma^{-1} (n_S + \bar{n}_f + 1) + \beta_c I_{LR}^2 \gamma^{-3} (n_S + n_A + 1) \quad . \tag{3.4}$$

There is another process contributing to dn_S/dx to this order -- the second-order process in Fig. B2, in which two laser photons are annihilated and one Stokes and one anti-Stokes photons are created, with a fundamental phonon as the intermediate state. Both time orderings of the vertices and both directions of the f arrow must be included as usual. For this case of output photons having frequencies ω_A and ω_S, which differ from ω_L by $\pm\omega_f$, the intermediate-state energy denominator $[(\omega_L - \omega_S)^2 - \omega_f^2]^2$ vanishes. The standard procedure of replacing ω_f^2 by $(\omega_f + \frac{1}{2}\tilde{\Gamma}_f)^2 \cong \omega_f^2 + \tilde{\Gamma}_f\omega_f$ then gives $\Gamma_f^2\omega_f^2$ for the energy denominator, and the standard perturbation-theory treatment yields

$$dn_S/dx = \beta_c I_{LR}^2 \gamma^{-3} (n_S + n_A + 1) \quad .$$

Adding this contribution from the two-laser-photon process to the result (3.4) for the Stokes process gives

$$dn_S/dx = \beta_c I_{LR} \gamma^{-1} (n_S + \bar{n}_f + 1) + 2\beta_c I_{LR}^2 \gamma^{-3} (n_S + n_A + 1) \quad . \qquad \underline{(3.5)}$$

Since the result (3.5) for dn_S/dx contains n_A, the equation for n_A must be considered simultaneously. If every f phonon created in the Stokes process were annihilated in an anti-Stokes process, the anti-Stokes photons would be created at the same rate as that of the Stokes photons. However, there are other processes competing for the annihilation of the f phonons, such as the splitting of the f phonons into two other phonons Q and -Q, as an example. Thus, fewer anti-Stokes photons are generated in general, although the anti-Stokes generation rate can approach that of the Stokes photons.

The equation for n_A corresponding to (3.5) is obtained by the same method used above:

$$dn_A/dx = \beta_c I_{LR} \gamma^{-1} (\bar{n}_f - n_A) + 2\beta_c I_{LR}^2 \gamma^{-3}(n_S + n_A + 1) . \qquad (3.6)$$

For $\gamma \cong 1$, these simultaneous equations (3.5) and (3.6) for n_A and n_S are linear and can be solved easily as follows: For $\gamma = 1$, (3.5) and (3.6) can be written as

$$\frac{d|n\rangle}{dx} = \mathcal{L}|n\rangle + |u\rangle \qquad (3.7)$$

where $|n\rangle$ is the vector with components $\langle A|n\rangle = n_A$ and $\langle S|n\rangle = n_S$, that is

$$|n\rangle \to \begin{pmatrix} n_A \\ n_S \end{pmatrix} , \qquad |u\rangle \to \beta_c I_{LR} \begin{pmatrix} \bar{n}_f + 2 I_{LR} \\ \bar{n}_f + 1 + 2 I_{LR} \end{pmatrix}$$

and

$$\mathcal{L} \to 2\beta_c I_{LR}^2 \begin{pmatrix} 1 - \frac{1}{2} I_{LR} & 1 \\ 1 & 1 + \frac{1}{2} I_{LR} \end{pmatrix}^{-1} .$$

The arrow indicates the representation in the $|A\rangle$, $|S\rangle$ basis. The solution to (3.7) with $|n\rangle = 0$ at $x = 0$ is

$$|n\rangle = \mathcal{L}^{-1}(e^{\mathcal{L}x}-1)|u\rangle = \sum_{i=\pm 1} \lambda_i^{-1}(e^{\lambda_i x}-1)|\ell_i\rangle\langle\ell_i|u\rangle \quad , \qquad (3.8)$$

where $|\ell_i\rangle$ are the orthonormal eigenvectors of the Hermitian operator $\mathcal{L}$ and the λ_i are the eigenvalues: $\mathcal{L}|\ell_i\rangle = \lambda_i|\ell_i\rangle$. The eigenvalues of $\mathcal{L}$ are

$$\lambda_\pm = 2\beta_c I_{LR}^2 \left[1 \pm \left(1 + \frac{1}{4} I_{LR}^{-2} \right)^{1/2} \right] , \qquad (3.9)$$

and the eigenvectors are

$$|\ell_\pm\rangle \rightarrow C_\pm \begin{pmatrix} -\frac{1}{2} I_{LR}^{-1} \pm \left(1 + \frac{1}{4} I_{LR}^{-2}\right)^{1/2} \\ 1 \end{pmatrix} , \qquad (3.10)$$

where the normalization constants have the values

$$C_\pm^2 = \left\{ 1 + \left[-\frac{1}{2} I_{LR}^{-1} \pm \left(1 + \frac{1}{4} I_{LR}^{-2}\right)^{1/2} \right]^2 \right\}^{-1} . \qquad (3.11)$$

The term $i = +1$ in (3.8) is much greater than the term $i = -1$; thus, the scalar product of $|n\rangle$ with $|S\rangle$ is

$$n_S = \langle S|n\rangle \cong \lambda_+^{-1} \left(e^{\lambda_+ x} - 1 \right) \langle S|\ell_+\rangle\langle\ell_+|u\rangle \quad . \qquad (3.12)$$

For $I_{LR} << \frac{1}{2}$, $\lambda_+ \cong \beta_c I_{LR}$, $\langle S|\ell_+\rangle = c_+ \cong 1$, and $\langle \ell_+|u\rangle$ $\cong \beta_c I_{LR} (\bar{n}_f + 1)$. Thus, (3.12) yields

$$n_S \cong (\bar{n}_f + 1) \left(e^{\beta_c I_{LR} x} - 1 \right) , \qquad \text{for } I_{LR} << \frac{1}{2} \quad . \qquad (3.13)$$

For $I_{LR} >> \frac{1}{2}$, $\lambda_+ \simeq 4\beta_c I_{LR}^2$, $\langle S | \ell_+ \rangle = c_+ \simeq 2^{-1/2}$, and $\langle \ell_+ | u \rangle \simeq 4(2)^{-1/2} \beta_c I_{LR}^2$, and (3.12) gives

$$n_S \simeq \frac{1}{2} e^{4\beta_c I_{LR}^2 x}, \qquad \text{for } I_{LR} >> \frac{1}{2} . \tag{3.14}$$

The low-dispersion results (3.13) and (3.14) are easy to interpret. For $I_{LR} << \frac{1}{2}$, $n_A << n_S$ and dn_S/dx is given by the first term on the right-hand side of (3.5), which is easily integrated to give (3.13). For $I_{LR} >> \frac{1}{2}$, (3.5) and (3.6) give $dn_A/dx \simeq dn_S/dx$ and $n_A \simeq n_S$ in the gain region. Thus (3.5) with $n_A = n_S$ gives $dn_S/dx \simeq 4\beta_c I_{LR}^2 n_S$, which gives (3.14) on integration.

The result (3.8) also indicates that the anti-Stokes intensity is comparable to the Stokes intensity when $I_{LR} \gtrsim 1$ in the low-dispersion case. In particular, for $I_{LR} < \frac{1}{2}$, $I_A/I_S \simeq I_{LR}$, and for $I_{LR} > 2$, $I_A/I_S \simeq 1 - \frac{1}{2} I_{LR}$. These results follow directly from (3.8): Taking the scalar product of (3.8) with $|A\rangle$ gives (3.12) with S replaced by A. Thus, $n_A/n_S = \langle A | \ell_+ \rangle / \langle A | \ell_+ \rangle$. With (3.11), this gives $n_A/n_S = -\frac{1}{2} I_{LR}^{-1} + (1 + \frac{1}{4} I_{LR}^{-2})^{1/2}$, from which the stated results are evident. It should be mentioned that even when $I_A \simeq I_S$, the values of I_S and I_A can be so small that second Stokes, second anti-Stokes and other higher-order intensities are negligible. The enhanced-relaxation case of $\gamma^2 > 1$ is considered in Sec. VI.

IV. TRANSIENT SOLUTION

For the high-dispersion case, the transient solution to the coupled equations (2.1) and (2.4) is obtained as follows: Kroll's solution[10] for the case of no gain enhancement suggests a trial solution of the form

$$n_S = n_{S0} \exp [(A x t)^{1/2} - B t] \tag{4.1}$$

with $c_S^{-1} \partial n_S / \partial t$ negligible. Eliminating n_f from (2.1) and (2.4) by neglecting the $\bar{n}_f$ and 1 terms, taking the derivative of (2.4) with respect to time, and neglecting[11] $\partial n_S / \partial t$ and $\partial^2 n_S / \partial t^2$ gives

$$\partial^2 n_S / \partial x \partial t + \Gamma_f (1 - I_{LR}) \partial n_S / \partial x - c_S^{-1} \Gamma_f^2 I_{LR} n_S = 0 \quad . \tag{4.2}$$

Substituting (4.1) into (4.2) and solving for A and B gives

$$n_S \simeq n_{S0} \exp [(4 \Gamma_f^2 c_S^{-1} I_{LR} x t)^{1/2} - \Gamma_f (1 - I_{LR}) t] \tag{4.3}$$

for

$$I_{LR} c_S / \Gamma_f^2 \lesssim t \lesssim I_{LR} (1 - I_{LR})^{-2} (x / c_S) \quad . \tag{4.4}$$

In solving (4.1) and (4.2) the term $\frac{1}{4} (A / t x)^{1/2}$ was neglected with respect to the terms that are independent of t. The left-hand inequality in (4.4) is required in order for this approximation to be valid.

Substituting $t = I_{LR} (1 - I_{LR})^{-2} (x / c_S)$ from the right-hand side of (4.4) into (4.3) gives the steady-state solution $n_S = n_{S0} \exp (\beta_{g\,new} x)$. See (2.6). Here n_{S0} is the small value of n_S evaluated at the nonzero value of

$\Delta t \cong I_{LR} c_S / \Gamma_f^2$, corresponding to the left-hand side of (4.4). In other words, in considering the transient solution to (2.1) and (2.4), the very short-time, small-x region is neglected [left-hand inequality in (4.4)]; thus, the steady-state limit of this solution corresponds to the last term in (2.6).

For t greater than this value, that is $t > t_{ss} \equiv I_{LR}(1 - I_{LR})^{-2}(x/c_S)$, the steady-state solution is valid. With the definition of $\beta_{g\,new}$ under (2.6), t_{ss} can be written as

$$t_{ss} = \tau_f \beta_{g\,new} x (1 - I_{LR})^{-1} = \tau_f g I_L x (1 - I_{LR})^{-2} \qquad (4.5)$$

where $\tau_f \equiv 1/\Gamma_f$ and $g = \Gamma_f / c_S I_R$ is often called the Stokes gain. In the gain region, $\beta_{g\,new} x >> 1$ is satisfied; thus $t_{ss} >> \tau_f$. This important result (4.5) shows that the steady state is not reached until times much greater than the Reststrahl phonon lifetime τ_f and that the greater x, the longer it takes to reach the steady state. It is important to realize that t_{ss} is a function of I_L. Thus, in a curve of I_L as a function of I_S, the steady state could be attained at small values of I_S while the transient solution applies at large values of I_S. This would give a shape of the I_S vs I_L curve that is different from both the transient and steady-state shapes.

The left-hand inequality in (4.4) is the same as Kroll's results since the enhancement does not affect the region of very small n_S. In the limit $I_{LR} << 1$ of no gain enhancement, (4.5) reduces to Kroll's result, $t_{ss} = \tau_f g I_L x$, and the result (4.3) for n_S reduces to Kroll's result.

The transient solution for the low-dispersion case has not been considered. It is expected that central features will be the same as those discussed under (4.5).

V. RELATION BETWEEN THE CASES OF HIGH AND LOW DISPERSION

Since energy and $\underset{\sim}{k}$ conservation are involved in the arguments to follow, (2.4) is rewritten with the line-shape factor and wave vectors displayed explicitly

$$dn_S(\underset{\sim}{k}_S)/dx = 2\pi |V_R|^2 n_L \delta(\tilde{\omega}_S)[n_S(\underset{\sim}{k}_S) + n_f(\underset{\sim}{k}_L - \underset{\sim}{k}_S) + 1] . \quad (5.1)$$

According to (5.1), $n_S(\underset{\sim}{k}_S)$ is coupled to $n_f(\underset{\sim}{k}_L - \underset{\sim}{k}_S)$. The equation of motion of $n_f(\underset{\sim}{k}_L - \underset{\sim}{k}_S)$ is given by (2.1) and (3.1) with $n_f \equiv n_f(\underset{\sim}{k}_L - \underset{\sim}{k}_S)$ and $n_A \equiv n_A(2\underset{\sim}{k}_L - \underset{\sim}{k}_S)$, where $\underset{\sim}{k}_A = \underset{\sim}{k}_L + \underset{\sim}{k}_f = 2\underset{\sim}{k}_L - \underset{\sim}{k}_S$.

The line-shape factor $\delta(\tilde{\omega}_S)$ in (5.1) is peaked at $\tilde{\omega}_S = \omega_L - \omega_f - c_S k_S = 0$, where $\omega_S = c_S k_S$. For $\tilde{\omega}_S = 0$, $\omega_f = \omega_L - c_S k_S$; thus, $\tilde{\omega}_A = \omega_L + \omega_f - \omega_A$ becomes

$$\tilde{\omega}_A = 2c_L k_L - c_S k_S - c_A |2\underset{\sim}{k}_L - \underset{\sim}{k}_S| . \quad (5.2)$$

The peak in $\delta(\tilde{\omega}_A)$ is at $\tilde{\omega}_A = 0$, and the width is $\sim\tilde{\Gamma}_f$. First consider forward scattering $\hat{k}_S = \hat{k}_L$. Then $|2\underset{\sim}{k}_L - \underset{\sim}{k}_S| = 2k_L - k_S$, with $c_A - c_S \cong 2(c_L - c_S)$ and $k_L - k_S = \omega_f/c_S + k_L(1 - c_L/c_S)$, which is obtained by solving $\tilde{\omega}_S = 0$ for $-k_S$ and adding k_L, (5.2) gives

$$\tilde{\omega}_A = \delta_c \omega_f , \qquad \delta_c \equiv 2(1 - c_L/c_S)[1 + (1 - c_L/c_S)\omega_L/\omega_f] . \quad (5.3)$$

When the shift $\tilde{\omega}_A$ is less than the width $\tilde{\Gamma}_f$, that is

$$D \equiv \delta_c \omega_f / \tilde{\Gamma}_f << 1 \quad (5.4)$$

then $\delta(\tilde{\omega}_A) \cong \delta(\tilde{\omega}_S)$, while for $D >> 1$, $\delta(\tilde{\omega}_A) << \delta(\tilde{\omega}_S)$. Thus, for forward scattering, both $(dn_f/dt)_A$ and $(dn_f/dt)_{rS}$ defined as the right-hand side of (2.1),

are included in dn_f/dt when $D \ll 1$. Only $(dn_f/dt)_{rS}$ is included when $D \gg 1$.

In the latter case of $D \gg 1$, called the high-dispersion case, the previous results[3] such as (2.6) are valid; while in the former case of low dispersion ($D \ll 1$), the anti-Stokes process must be included, as in the analysis in Sec. III. Specifically, $n_f(\underset{\sim}{k}_L - \underset{\sim}{k}_S)$ is coupled to $n_S(\underset{\sim}{k}_S)$ according to (5.1). For forward scattering with high dispersion, $n_f(\underset{\sim}{k}_L - \underset{\sim}{k}_S)$ is coupled back to $n_S(\underset{\sim}{k}_S)$ only. For low dispersion, $n_A(2\underset{\sim}{k}_L - \underset{\sim}{k}_S)$, $n_f(\underset{\sim}{k}_L - \underset{\sim}{k}_S)$ and $n_S(\underset{\sim}{k}_S)$ are all coupled. For $D \gg 1$, I_S should exhibit the jump; while for $D \ll 1$, the $\exp I_L^2$-gain result (3.12) is expected.

The dispersion factor D in (5.4) can be written in terms of $dn/d\lambda$ as follows: The first two terms in the Taylor series expansion of $c(\lambda)$ give

$$1 - c_L/c_S \cong c_L^{-1}\,[dc(\lambda)/d\lambda\,](\lambda_S - \lambda_L) \quad .$$

With $\lambda_S - \lambda_L = c_S/\omega_S - c_L/\omega_L \cong \lambda_L\omega_f/\omega_L$ for $\omega_f/\omega_L \ll 1$, and $c^{-1}dc/d\lambda = -n^{-1}dn/d\lambda$, this gives

$$1 - c_L/c_S \cong (\omega_f/\omega_L)(\lambda/n)\,dn/d\lambda \quad . \tag{5.5}$$

With this result, the second term in (5.3) is equal to $(\lambda/n)(dn/d\lambda)$, which is negligible with respect to unity; thus $\delta_c \cong 2(1 - c_L/c_S)$, and (5.4) and (5.5) give

$$D \cong (2\omega_f^2/\omega_L\tilde{\Gamma}_f)(\lambda/n)\,dn/d\lambda \quad . \tag{5.6}$$

For the rotational Raman mode in gases, $D \ll 1$ tends to be satisfied, while for solids, liquids, and the vibrational Raman modes in gases, $D \gg 1$

typically is satisfied. For example, for liquid nitrogen[12] with $\omega_f = 2,330\ cm^{-1}$, $\omega_L = 1.44 \times 10^4\ cm^{-1}$, $\Gamma_f = 0.0670\ cm^{-1}$, n = 1.20, and a typical value of $dn/d\lambda = 3 \times 10^{-2}\ (\mu m)^{-1}$, (5.6) gives D = 230; thus, the previous results[3] for high dispersion are valid. For the vibrational $Q_1(1)$ Raman mode in hydrogen gas[13] at 300 K and 100 atmospheres, $\omega_f = 4,150\ cm^{-1}$, $\omega_L = 1.44 \times 10^4\ cm^{-1}$, $\Gamma_f = 1.1 \times 10^2\ cm^{-1}$, n = 1.0, and $dn/d\lambda = 10^{-3}\ (\mu m)^{-1}$, (5.6) gives D = 150. Again, the high-dispersion results apply. Finally, for the rotational $S_0(1)$ Raman mode in hydrogen gas at 300 K and seven atmospheres, $\omega_f = 243\ cm^{-1}$, $\omega_L = 1.44 \times 10^4\ cm^{-1}$, $\Gamma_f = 4.7 \times 10^{-3}\ cm^{-1}$, n = 1.0, and $dn/d\lambda = 7 \times 10^{-5}\ (\mu m)^{-1}$, and (5.6) gives $D = 8 \times 10^{-2}$. Thus the low-dispersion results are applicable. These cases are discussed further in Sec. VII.

Next, consider the case of scattering at angles away from the forward direction. In the high-dispersion case, energy and $\underset{\sim}{k}$ are not simultaneously conserved for the Stokes and anti-Stokes processes for forward scattering. However, energy and $\underset{\sim}{k}$ can be conserved for scattering into a range of angles θ, where $\hat{k}_L \cdot \hat{k}_S = \cos\theta$. The angle for exact energy conservation (line-shape factors infinitely narrow) is easily obtained as follows. Consider the value of $|2\underset{\sim}{k}_L - \underset{\sim}{k}_S|$ in (4.2):

$$|2\underset{\sim}{k}_L - \underset{\sim}{k}_S| = \left[4k_L^2 + k_S^2 - 4k_L k_S \cos\theta\right]^{1/2} .$$

For $\theta << 1$, this reduces to

$$|2\underset{\sim}{k}_L - \underset{\sim}{k}_S| \simeq (2k_L - k_S)(1 + \theta^2) .$$

Substitution into (4.2) gives $\tilde{\omega}_A = \delta_c \omega_f - \omega_L \theta^2$. Setting $\tilde{\omega}_A = 0$ gives

$$\theta_{mxA} \cong (\delta_c \omega_f / \omega_L)^{1/2} \quad . \qquad (5.7)$$

For a typical case of $\delta_c = 5 \times 10^{-3}$ and $\omega_f/\omega_L = 2 \times 10^{-2}$, $\theta_{mxA} = 10^{-2}$ radians.

Physically, these results have the following significance. In the high-dispersion case, (2.6) and the results of Ref. 3 are valid for forward scattering. The Stokes intensity decreases and the anti-Stokes intensity increases as θ increases from 0 to θ_{mxA}. At $\theta = \theta_{mxA}$, wave vector and $\underset{\sim}{k}$ are simultaneously conserved for the Stokes and anti-Stokes processes; thus the low-dispersion results (3.12) to (3.14) are valid at $\theta = \theta_{mxA}$. As θ increases further, both the Stokes and anti-Stokes intensities decrease. The anti-Stokes intensity is therefore strongest on the surface of the cone of half angle θ_{mxA}. In Sec. III it is shown that the anti-Stokes intensity becomes comparable with the Stokes intensity when $I_{LR} > 1$ (for energy and $\underset{\sim}{k}$ conserved).

In considering the angle dependence of the Stokes and anti-Stokes intensities, the beam aspect ratio

$$\theta_{sam} \equiv r_b / L_{sam} \qquad (5.8)$$

where r_b is the radius of the laser beam and L_{sam} is the sample length, must be considered. If $\theta_{mxA} > \theta_{sam}$, the scattered Stokes and anti-Stokes radiation at angles $\pm\theta_{mxA}$ leave the laser beam and I_S is not amplified over the full length L_{sam}. Finally, it should be mentioned that in experiments with multimode lasers, simultaneous conservation of energy and $\underset{\sim}{k}$ for the Stokes and anti-Stokes processes may be possible even in the forward-scattering high-dispersion case.

VI. EFFECTS OF ENHANCED PHONON RELAXATION

First consider the high-dispersion case. The enhanced relaxation of f reduces the value of n_S below that given in (2.6). This can be seen directly from (2.4) with $dn_S/dt = 0$, where reducing the value of n_f and increasing the value of γ in $\beta_{I\gamma}$ reduces the value of dn_S/dx. For $\gamma^2 > 1$, (2.4) with n_f given by (2.3) is a nonlinear first-order differential equation which can be solved by direct integration since dn_S/dx is a function of n_S only.

The behavior of the solution can be determined by considering limiting cases. From Ref. 4,

$$\gamma^2 = (1 - n_f/n_p)^{-1} \tag{6.1}$$

where n_p is an upper bound to n_f defined as

$$n_p/N = p(\bar{n}_Q + \tfrac{1}{2})\,\Gamma_Q^2/S_Q\,\omega_f\,\Gamma_f \quad , \tag{6.2}$$

where Γ_Q is the relaxation frequency of the phonons Q into which f decays, p/S_Q is a constant of order unity typically, and N is the number of unit cells in the crystal. Substituting (6.1) into (2.3) and neglecting $\gamma^2\bar{n}_f + I_{LR}$ in (2.3) gives

$$(\gamma^2 - 1)(\gamma^2 - I_{LR}) = \gamma^2 I_2 \tag{6.3}$$

where $I_2 \equiv I_{LR}\, n_S/n_p$.

Negligible phonon enhanced relaxation corresponds to $I_2 \ll 1$. Then, in the region of interest ($I_{LR} < 1$), (6.3) gives $\gamma^2 \simeq 1$. Thus, the solution (2.6) is valid as expected in the absence of enhanced phonon relaxation. The solution to (6.3) for arbitrary I_2 is

$$\gamma^2 = \frac{1}{2}(1 + I_{LR} + I_2)\left\{1 + \left[1 - \frac{4\, I_{LR}}{(1 + I_{LR} + I_2)^2}\right]^{1/2}\right\} . \qquad (6.4)$$

For the case of a large phonon enhanced-relaxation effect ($I_2 \gg 1$), (6.4) gives

$$\gamma^2 \simeq I_2 .$$

Substituting this result into (2.4) gives

$$dI_S/dx = \beta_c (I_{LR} I_p I_S)^{1/2} , \qquad (6.5)$$

which has the solution $I_S = \frac{1}{4}\beta_c^2 I_{LR} I_p x^2$ plus a constant. This value of I_S is less than the value from (2.6) and I_S is increasing rather slowly as functions of I_{LR} (linearly) and of x (square law). The behavior of I_S as a function of I_{LR} according to the limiting cases of (6.5) and (2.6) is illustrated schematically in Fig. B3. For $n_p = \infty$, the result (2.6) for no enhanced phonon relaxation is valid for all $I_{LR} \lesssim 1$. As n_p decreases, the deviation from (2.6) occurs at smaller values of I_S as shown in the figure.

Next consider the low-optical-dispersion case. We seek the solution of (3.5) and (3.6) for the case of $\gamma^2 > 1$. It is clear that values of n_S and n_A are reduced when $\gamma^2 > 1$ since both dn_S/dx and dn_A/dx are decreased according to (3.5) and (3.6). Limiting cases of these nonlinear, coupled differential equations will be analyzed to determine the effect of the enhanced phonon relaxation.

The value of γ is obtained from (3.3) with $\bar{n}_f/n_p \ll 1$ and $n_S + n_A \gg 1$, which are well satisfied:

$$\gamma^2 \cong 1 + I_{SA2} \qquad (6.6)$$

where $I_{SA2} = I_{LR}(n_S + n_A)/n_p$. Note that $I_2 \leq I_{SA2} \leq 2I_2$ since $I_A \leq I_S$. The f-mode amplitude is, from (6.6) and (6.1),

$$n_f/n_p = I_{SA2}/(I_{SA2} + 1) \quad . \tag{6.7}$$

The behavior of the solution to (3.5) is simple. For no Stokes input, $n_S = 0$ at $x = 0$. First consider the case of $I_{LR} \ll 1$, for which n_A is negligible. For very small x such that $n_S \ll \bar{n}_f + 1$ and $n_S + n_f \ll 1$, n_S increases linearly with x. When n_S reaches values $\gg \bar{n}_f + 1$, it increases approximately exponentially until $\gamma^2 - 1 \ll 1$ is no longer satisfied. For example, for $I_{LR} \ll 1$, $n_S \gg \bar{n}_f + 1$, and $\gamma \cong 1$, (3.5) gives

$$n_S \sim \exp(\beta_c I_{LR} x) \quad .$$

As n_S increases further, γ^2 becomes greater than 1, according to (6.6). For $I_{SA2} \cong I_{LR} n_S/n_p \gg 1$, $\gamma \cong (I_{LR} n_S/n_p)^{1/2}$, and (3.5) gives

$$dn_S/dx \cong \beta_c (I_{LR} n_p n_S)^{1/2} \quad . \tag{6.8}$$

According to (6.8), n_S increases rather slowly as x^2.

Next consider the behavior of $n_S(x)$ for $I_{LR} \gg 1$ formally. For $n_S \ll 1$, n_S increases linearly with x. Notice that the slope is proportional to I_{LR}^2, rather than I_{LR} as in the usual spontaneous-emission region. For $n_S \gg 1$ and $I_{SA2} \ll 1$, (3.5) with $n_A \cong n_S$ gives

$$n_S \sim \exp(4\beta_c I_{LR}^2 x) \quad .$$

As n_S increases further, $I_{SA2} \simeq 2 l_{LR} I_S / l_p$ is no longer negligible, and the γ^{-3} terms in (3.5) reduce the rate of increase of n_S. For $I_{SA2} >> 1$, (3.5) becomes

$$dn_S/dx = \beta_c \left(\tfrac{1}{2} l_{LR} n_p n_S \right)^{1/2} + 2\beta_c \left(l_{LR} n_p^3 / 2 n_S \right)^{1/2}$$

so that n_S again increases as x^2 at very large values of x.

For a given thickness x, in the spontaneous-emission region at very small values of I_{LR}, n_S is proportional to l_{LR}. For greater values of I_{LR}, n_S increases according to (3.8) as long as $I_{SA2} << 1$; i.e., $n_S << \frac{1}{2} n_p / I_{LR}$. For larger values of l_2, γ^2 increases and n_S increases much less rapidly than in the case of $\gamma = 1$. For $n_S >> \frac{1}{2} n_p / I_{LR}$, n_S is proportional to I_{LR} (times x^2). This behavior is similar to that already illustrated schematically in Fig. B3 for the case of high dispersion.

VII. APPLICATION TO EXPERIMENTS

There is evidence for the enhanced Stokes gain throughout the literature in various types of experiments in gases, liquids, and solids.[10-17] However, there is little data for which accurate quantitative comparison with experiment can be made, a notable exception being the experiments on liquid oxygen and nitrogen by Grun, McQuillan, and Stoicheff.[12] As the laser power was increased, the Stokes power exhibited successively regions of spontaneous emission, gain, the jump, and a saturation region of little further increase. See Figs. B4 and 5. There was no self focusing. All known mechanisms for explaining the jump, with the possible exception of oscillations due to feedback by Rayleigh scattering, were ruled out by the authors. A simple calculation indicates that Rayleigh scattering is orders of magnitude too small to cause feedback oscillations. For liquid nitrogen, the value of $D = 230$ obtained in Sec. V indicates that the high-dispersion results of Ref. 3 and (2.5) are applicable. The excellent agreement with the experimental results over the full 11-decade change in I_S is seen in Fig. B4. The theoretical curve was plotted using the theoretical value of $\beta_c = \Gamma_f / c_S = 0.51\ \text{cm}^{-1}$ and the observed values of the spontaneous-emission and saturation Stokes intensities and adjusting the value of I_R slightly to position the jump properly. Thus only the scale factors for the axes, and not the shape of the curve, were adjusted to fit the experimental points. Adjusting β_c to smaller values improves the fit, but is unnecessary in view of the already satisfactory agreement and the scatter in the experimental points. The jump, the saturation at large I_L, the transition from unamplified to amplified spontaneous emission, and the previously unexplained magnitude of g and deviation from exponential gain in the gain region

$0.3 < I_L / I_R < 0.9$ (resulting from $\beta_{g\,new} > \beta_g \equiv g I_L$), are apparent in the theoretical curve. Second-order Stokes radiation was observed at high intensities, as expected. In this experiment and those below with one input beam, $\tilde{\Gamma}_f = \Gamma_f$ should be satisfied.

The agreement between the theoretical result (2.5) and the experimental points[12] for liquid oxygen is slightly better than the agreement for liquid nitrogen as seen in Fig. B5. Also, the experimental and theoretical values of the ratio of the values of I_R for O_2 and N_2 are in surprisingly good agreement: In the spontaneous-emission region ($\beta_{g\,new} << 1$), (2.5) gives $I_S \sim (\Gamma_f / c_S I_R) I_L$. Thus, $I_R \sim \Gamma_f (c_S S)^{-1}$, where S is the slope of the I_S vs I_L curve. Clements and Stoicheff[18] have shown that the ratio (O_2 to N_2 always) of the Γ_f's is 0.117/0.067 = 1.75. The ratio of the c_S's is 0.99, and from Ref. 12 the ratio of the slopes is 1.3. Thus, the ratio of the I_R's is 1.75/(0.99)(1.3) = 1.36, which is in much better agreement with the experimental value of 1.23 than expected in view of uncertainties in the experimental values of the parameters.

Hagenlocker and coworkers[13] observed that the vibrational Raman $Q_1(1)$ Stokes power generated by a laser beam focused in hydrogen gas at 100 atmospheres and room temperature, where self focusing is not expected, increased by 10 to 11 orders of magnitude as the laser power was increased by a factor of two. See Fig. 11 of Ref. 13. Such a sharp increase cannot be explained by normal stimulated Raman gain,[13] but agrees well with the present theory. For the vibrational Raman $Q_1(1)$ mode in hydrogen gas at 100 atmospheres and 300 K, for which the steady state was obtained in the experiments, the value of D = 150 obtained in Sec. V indicates that the high-dispersion results are applicable; thus the observed jump is expected.

The laser intensity at the threshold was estimated[19] to be of the order of 250 MW/cm^2. From the values of $\Gamma_f = 2 \times 10^{10}$ sec^{-1} and $\beta_c I_{LR} / I_L = 1.5 \times 10^{-9}$ cm/W, the theoretical value of I_R is $\Gamma_f I_L / c_S \beta_c I_{LR} = 450$ MW/cm^2. The agreement is satisfactory since the experimental value of 250 MW/cm^2 is only an order-of-magnitude estimate and the values of the parameter in the theoretical result are uncertain.

In the same experiments, Hagenlocker and coworkers noted that for the rotational Raman $S_0(1)$ mode[20] in H_2 gas at seven atmospheres and 300 K, $\mathcal{L}n\, I_S$ varied as I_{LR}^2, rather than I_{LR} as expected at that time. As illustrated in Fig. B6, the agreement with the low-dispersion theory is excellent over the full range of 13 decades change in I_S. From Sec. V, $D = 8 \times 10^{-2}$, indicating that the low-dispersion, $\exp(I_{LR}^2)$ results of Sec. II are indeed applicable.

In spite of the good agreement between experiment and theory in both of these cases of H_2 gas, difficulties in explaining the results remain, particularly concerning the pressure P dependence of the threshold. The theory predicts that $I_R = \text{const.}\, \Gamma_f^2 / P$. With[21] $\Gamma_f = 9.4 \times 10^9$ sec^{-1} at 10 atmospheres and $\Gamma_f = 3.8 \times 10^{10}$ sec^{-1} at 100 atmospheres, this result gives

$$I_R\,(100 \text{ atm}) / I_R\,(10 \text{ atm}) = 1.6 \quad .$$

The experimental value[13] of this ratio is 0.25. Even though the steady state was not reached at 10 atmospheres, this decrease is not expected. A previous explanation[22] of the decrease in terms of a transient analysis seems to be in error. The predicted ratio was the same as that predicted here, but the experimental ratio is 1/4, rather than 4 as stated in Ref. 22. Enhanced gain from feedback is always a possibility in high-gain systems, and this effect could set in at values of $I_L < I_R$, thus rendering the parametric-instability enhancement unobservable.

McQuillan, Clements, and Stoicheff[16] observed a very sharp jump in I_S as a function of I_L in diamond. No evidence for self focusing was found. Their explanation of the jump as a feedback oscillation is plausible and may well be correct. The reflectance is 0.17 per surface. However, it is also possible that the instability is involved. The theoretical value of I_R is 4.4 GW/cm^2, and the experimental value is "at least $\sim$1.1 GW/cm^2." It should be possible to determine the source of the jump by, for example, tilting the sample and optical components at an angle with respect to the axis of the beam in order to avoid amplification of the specularly reflected Stokes beam. It is possible that a combination of feedback and the parametric instability are involved in the jump.

One fact favoring the instability mechanism in diamond is that the jump occurs in the spontaneous-emission region rather than the gain region. The Stokes power is linear in I_L up to the threshold, as seen in Fig. 5 of Ref. 16. The oscillation criterion is that $0.17 \exp(\beta_c I_{LR} x) \equiv 0.17 \exp(g\ell I_L) = 1$, which gives $\exp g\ell I_L = \exp(1.8)$. From (2.5), in the spontaneous-emission region, $\exp(g\ell I_L) - 1 \simeq g\ell I_L$. Since $\exp(1.8) - 1 = 5.05 \gg 1.8$, a significant deviation from linearity should have been observed if the jump were caused by feedback. Physically, gain is required for feedback oscillations, but the jump occurred in the spontaneous region, not the gain region. Another factor favoring the instability mechanism is that the observed damage threshold is consistent with the instability mechanism, but not with the feedback mechanism: Since the absorption (by the inelastic-scattering process) is not strongly peaked at any volume element in the crystal in the feedback case, the temperature rise is $(10^9 \text{ W/cm}^2)(10^{-8} \text{ sec})(0.218 \text{ cm})^{-1}(1.56 \text{ J/cm}^3 \text{ K})^{-1}(1332 \text{ cm}^{-1} / 14,403 \text{ cm}^{-1}) = 3 \text{ K}$ for complete conversion of I_L to I_S. For the instability case, the absorption length

becomes extremely short as I_{LR} approaches unity. Thus, the heated volume is assumed to be limited by thermal diffusion. With a diffusion length of 1 μm for 10 nsec in diamond, the temperature rise is $(3\,K)(0.218/10^{-4}) = 6,000\,K$, which is greater than the fracture temperature and melting temperature of diamond.

This calculation showing that diamond damages at $I_L \cong I_R$ illustrates a result that should be common to many Raman-active crystals. This is an important result since the typical values of $I_R \cong 10^9\,W/cm^2$ are much smaller than typical values of $\sim 10^{10}\,W/cm^2$ for breakdown in crystals, such as rocksalt-structure crystals, that are not Raman active. Monitoring the Stokes intensity in damage experiments would be interesting.

Since the instability may be involved in limiting the diameter of self-focused beams, measuring the Stokes intensity in self-focusing experiments also would be desirable. Indeed, Loy and Shen[23] found that the laser light at the forward-moving focus was converted into Raman radiation at such a high rate that little was left to be detected over the background. Additional experiments to determine if self-focusing diameters are limited by the stimulated Raman scattering process would be useful. A rough estimate indicates that the mechanism is feasible. With $I_{LR} \cong 10$ at the focus in toluene,[3] the transient result (4.3) gives $n_S \cong n_{S0} \exp[4(3.7\times 10^{11}\,sec^{-1})^2 10 \times (0.2\,cm)(10^{-11}\,sec)/2\times 10^{10}\,cm/sec]^{1/2} = 10^{10} n_{S0}$, which is sufficient to give complete conversion to Stokes radiation. It was assumed that the Stokes radiation remains in coincidence with the forward-moving focus for 10^{-11} sec (corresponding to a focal region of length 60 μm moving at velocity = 1.03 c_S, for example).

In distinguishing between the parametric-instability source of the jump and other sources, it is useful to realize that the jump is in I_S as a function of I_L, but there is no jump as a function of sample length x. Thus, the jump in I_S vs x

observed[24] in nitrobenzene at an unspecified constant value of I_L is expected to be the result of self focusing. Indeed, this was verified in subsequent experiments.[25]

Jumps similar to those discussed above have been observed in other investigations in gases, liquids, and solids,[13-17] but focusing and/or self-focusing make accurate comparison with the theory difficult. Nevertheless, a number of observed jumps may result from an increase in intensity by self focusing and the parametric instability in the region of high intensity. Amplifiers have not reached their expected gains, as the jump could not be suppressed no matter how much the feedback was reduced. The present interpretation of the jump as the parametric instability at $I_L = I_R$ explains the insensitivity to and lack of need for feedback.

Next consider the experiments[26-28] in which both laser and Stokes beams are incident on the sample, with both intensities I_L and I_S less than I_R. Using the results of the present paper it is easy to show that in the experiments of Colles and Giordmaine[23] the value of n_f approaches the upper bound n_p closely ($n_f = 0.994\ n_p$). In the absence of the enhancement of the f phonon relaxation frequency, the value of n_f would have been a factor of $\gamma^2 \cong 175$ greater. The estimates in Ref. 29 correspond to this greater value, since the enhanced-relaxation effect was unknown at the time of the experiment.

The detection of the phonon instability in these experiments has been difficult. A simple method of observing the phonon instability would be to observe the Stokes output from a sample with the incident Stokes intensity zero or very small. The values of thickness x and intensity I_{LR} are chosen to make $I_{SA2} << 1$ at room temperature and $I_{SA2} >> 1$ at some low temperature such as 77 K. Estimates

indicate that these temperatures should be appropriate for diamond. The experiment affords a method of determining the magnitude and temperature dependence of the lifetime Γ_Q^{-1} of the phonons with $\omega_Q + \omega_{-Q} = \omega_f$. Non-normal incidence should be used to avoid feedback-type oscillations resulting from reflections at the surfaces, and care to avoid other reflections should be exercised. It should be mentioned that for ultrashort lifetimes of the Q phonons,[30] the steady state may not be reached during the laser pulse duration.

Discussions with Dr. R. Orbach on the size of Γ_Q are gratefully acknowledged.

REFERENCES

*This section will be published in The Physical Review (in press).

1. E. J. Woodbury and W. K. Ng, Proc. IRE 50, 2367 (1962).
2. R. W. Hellwarth, Phys. Rev. 130, 1850 (1963). Also see: Y. R. Shen and N. Bloembergen, Phys. Rev. 137, A1787 (1965); N. Bloembergen, Am. J. Phys. 35, 989 (1967); F. De Martini, Phys. Rev. B4, 4556 (1971).
3. M. Sparks, Phys. Rev. Lett. 32, 450 (1974). In an erratum to this paper, the effect of optical dispersion was overlooked. Thus, the reservations expressed in the erratum are not appropriate.
4. M. Sparks and H. C. Chow, Phys. Rev. B10, 1699 (1974).

4a. H. Suhl, J. Phys. Chem. Solids 1, 209 (1959); H. B. Callen, Fluctuation, Relaxation and Resonance in Magnetic Systems, edited by D. ter Haas (Oliver and Boyd, Edinburgh, 1962); R. M. White and M. Sparks, Phys. Rev. 130, 623 (1963); M. Sparks, Ferromagnetic Relaxation Theory (McGraw-Hill, New York, 1964).

4b. S. V. Silin, Sov. Phys. JETP 21, 1127 (1965); D. F. DuBois and M. V. Goldman, Phys. Rev. Lett. 14, 544 (1965).

4c. W. H. Louisell, Coupled Modes and Parametric Electronics (Wiley, New York, 1960).

5. M. Sparks, Ferromagnetic Relaxation Theory (McGraw-Hill, New York, 1964).
6. R. M. White, M. Sparks, and I. Ortenburger, Phys. Rev. 139, A450 (1965).
7. M. Sparks and J. Wilson, unpublished.
8. The application of perturbation theory to problems of Boson equations of motion is discussed in detail in Ref. 5.
9. Important results are denoted by underscored equation numbers.

10. N. M. Kroll, J. Appl. Phys. 36, 34 (1965). Also see Ref. 13.

11. For short pulses (pulse length t less than the order of the sample-traverse time x/c_s) $c_s \partial n_s/\partial x >> \partial n_s/\partial t$ is no longer satisfied.

12. J. B. Grun, A. K. McQuillan, and B. P. Stoicheff, Phys. Rev. 180, 61 (1969).

13. E. E. Hagenlocker, R. W. Minck, and W. G. Rado, Phys. Rev. 154, 226 (1967).

14. G. Bret and G. Mayer, Physics of Quantum Electronics, edited by P. L. Kelley, B. Lax, and P. E. Tannenwald (McGraw-Hill, New York, 1966).

15. G. Bisson and G. Mayer, J. Phys. 29, 97 (1968).

16. A. K. McQuillan, W. R. L. Clements, and B. P. Stoicheff, Phys. Rev. A 1, 628 (1970).

17. W. Maier, W. Kaiser, and J. A. Giordmaine, Phys. Rev. 177, 580 (1969).

18. W. R. L. Clements and B. P. Stoicheff, Appl. Phys. Lett. 12, 246 (1968).

19. W. G. Rado, private communication.

20. The fact that the rotational-mode energy levels are not equally spaced should have negligible effect since $n_f/N << 1$.

21. P. Lallemand, P. Simova, and G. Bret, Phys. Rev. Lett. 17, 1239 (1966).

22. C. S. Wang, Phys. Rev. 182, 482 (1969).

23. Michael M. T. Loy and Y. R. Shen, IEEE J. Quantum Electronics QE9, 409 (1973).

24. P. Lallemand and N. Bloembergen, Appl. Phys. Lett. 6, 210 (1965).

25. Y. R. Shen and Y. Shaham, Phys. Rev. Lett. 15, 1008 (1965); P. Lallemand and N. Bloembergen, Phys. Rev. Lett. 15, 1010 (1965).

26. J. A. Giordmaine and W. Kaiser, Phys. Rev. 144, 676 (1966).

27. J. P. Coffinet and F. De Martini, Phys. Rev. Lett. 22, 60 (1969).

28. F. De Martini, Phys. Rev. B 4, 4556 (1971).

29. M. J. Colles and J. A. Giordmaine, Phys. Rev. Lett. 27, 670 (1971).

30. R. Orbach and L. A. Vredevoe, Physics 1, 91 (1964); L. A. Vredevoe, Phys. Rev. 140, A930 (1965).

Figure Captions

Fig. B1. Diagram of the Raman Stokes process R and phonon process P.

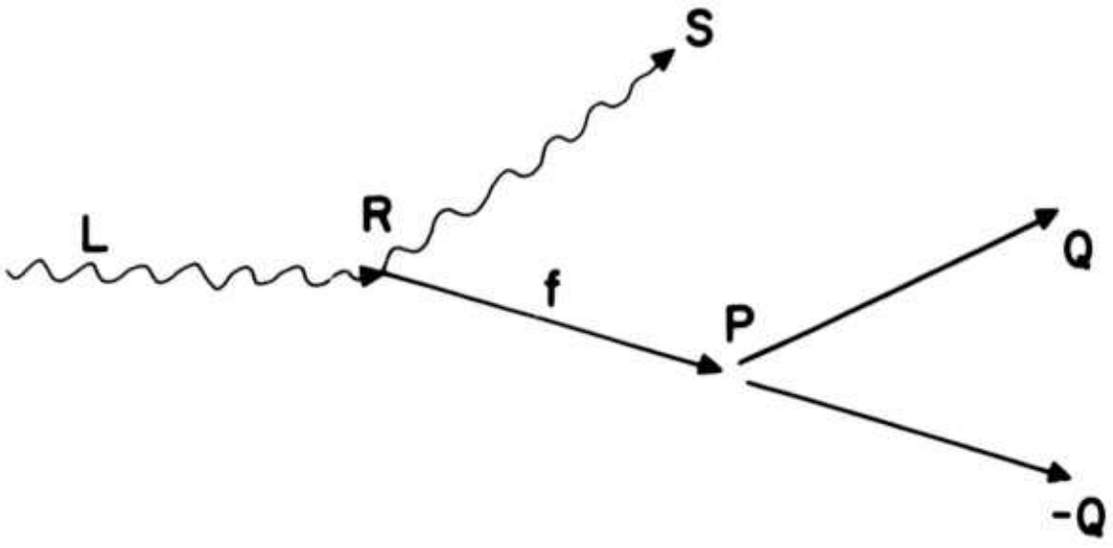
Fig. B2. The higher-order processes contributing to Stokes-photon generation.

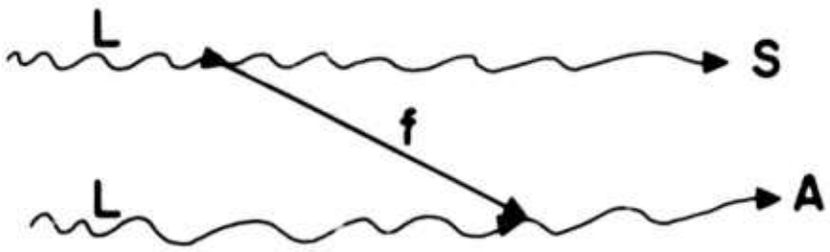
Fig. B3. Schematic illustration of the effect of the phonon instability on the Stokes intensity.

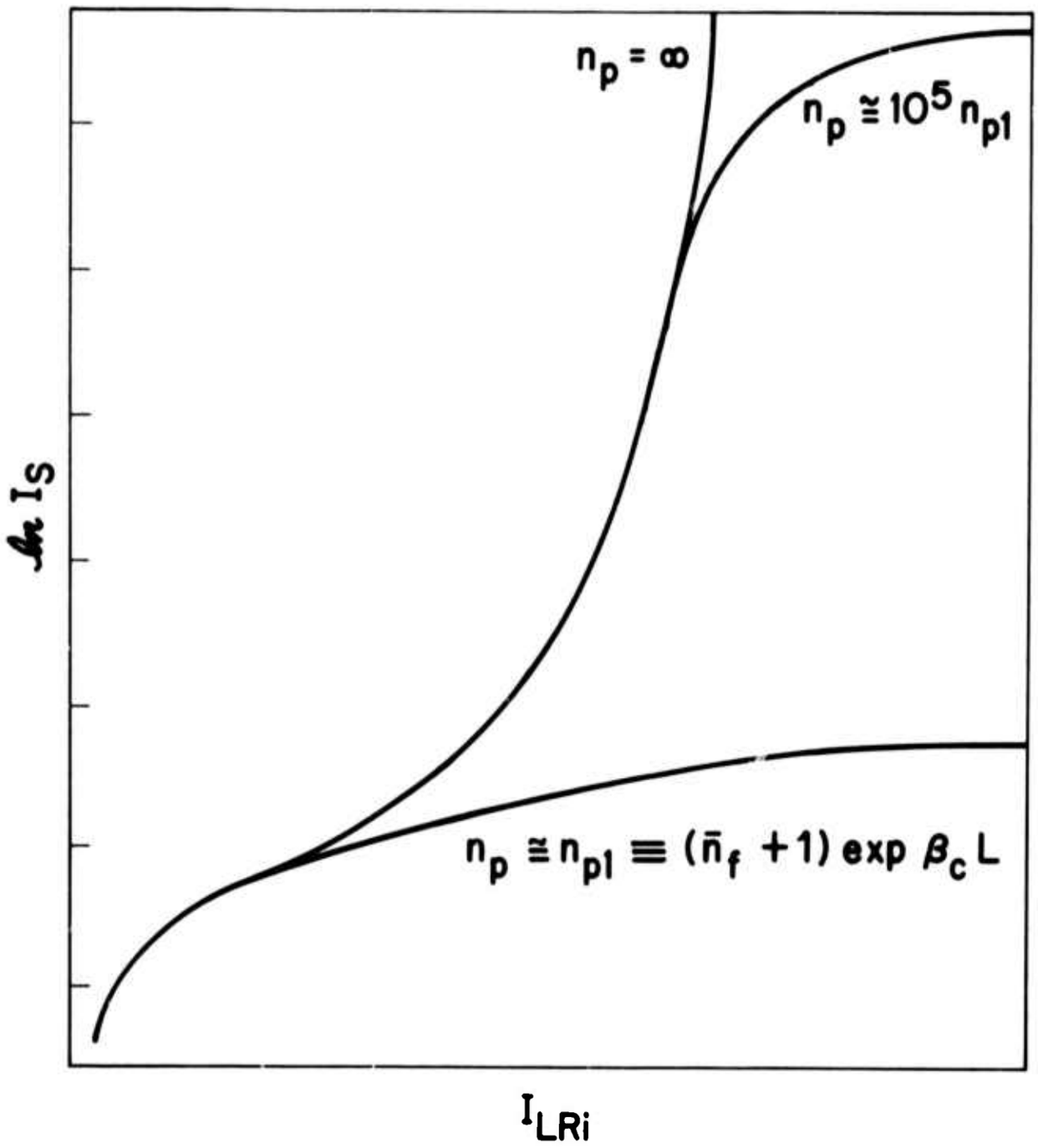
Fig. B4. Stokes-intensity-jump results in liquid nitrogen showing experimental results (crosses) of Grun, McQuillan, and Stoicheff,[12] the previous theoretical results, and the present theoretical results. The portion of the curve in the saturation region $I_L / I_R > 1$ is from Ref. 3.

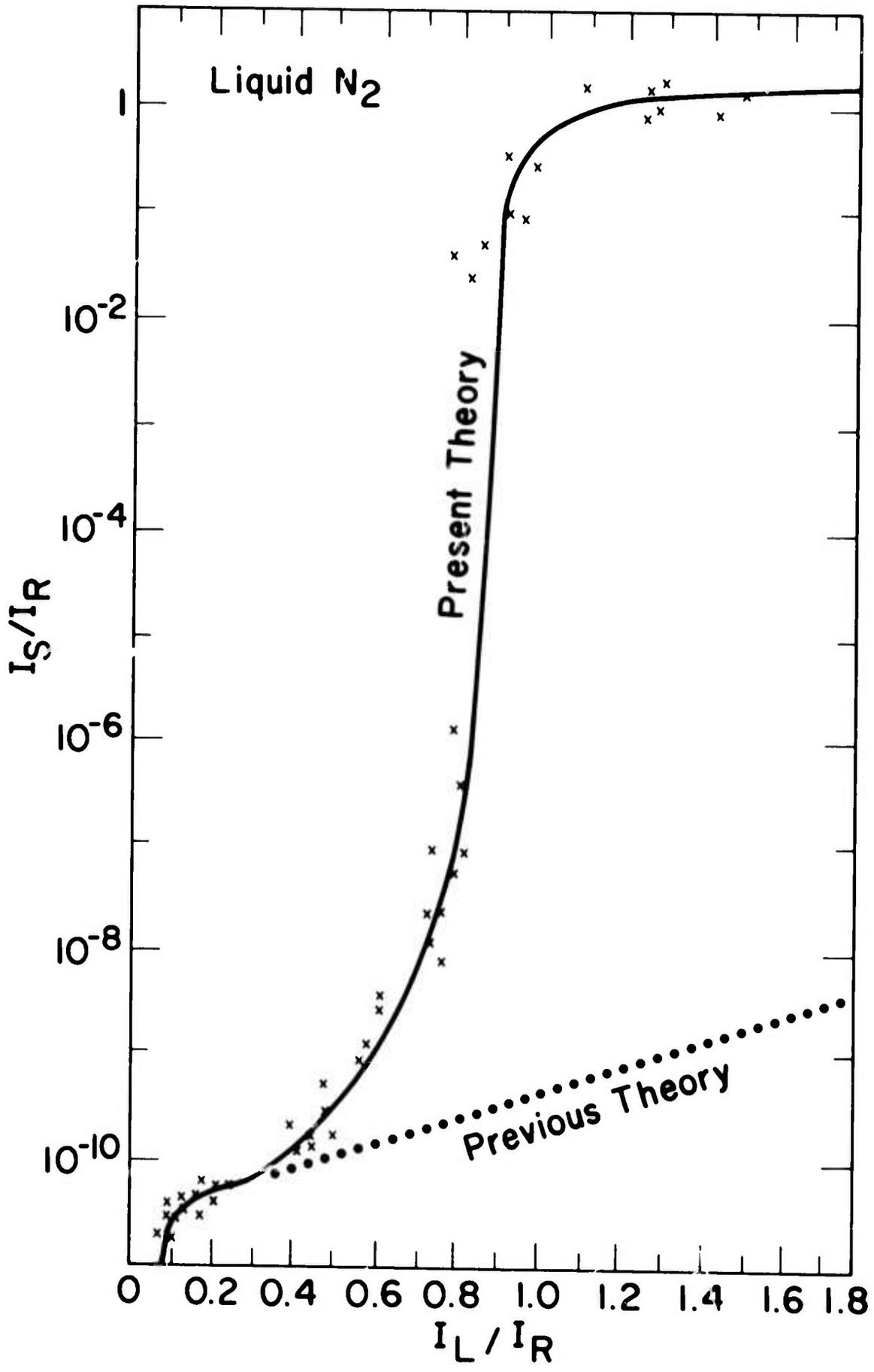
Fig. B5. Stokes-intensity-jump results in liquid oxygen showing experimental results (crosses) of Grun, McQuillan, and Stoicheff,[12] and the theoretical result from (2.6) (solid curve). The dashed curve showing the saturation is sketched, not plotted (in contrast to Fig. B4), since the saturation region is not of primary concern here.

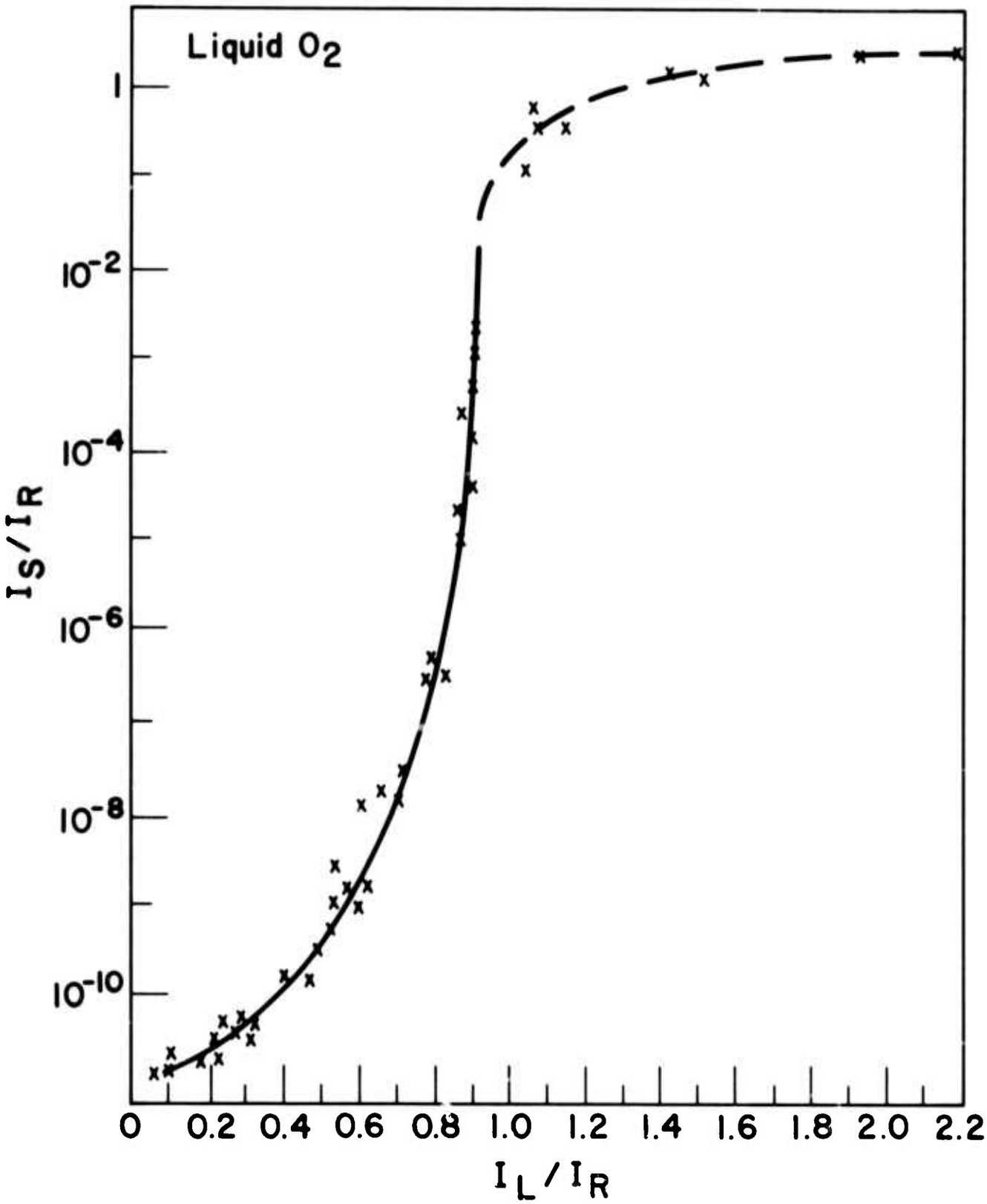
Fig. B6. Comparison of experimental rotational Raman $S_0(1)$ Stokes power in hydrogen gas of Hagenlocker, Minck, and Rado[13] (points) with the $\exp I_L^2$-gain results and original theoretical results. The saturation of I_S for $I_L > 4$ results from the depletion of the laser beam. The ordinate is in relative units and the numbers on the abscissa correspond roughly to power in megawatts, according to Ref. 13.

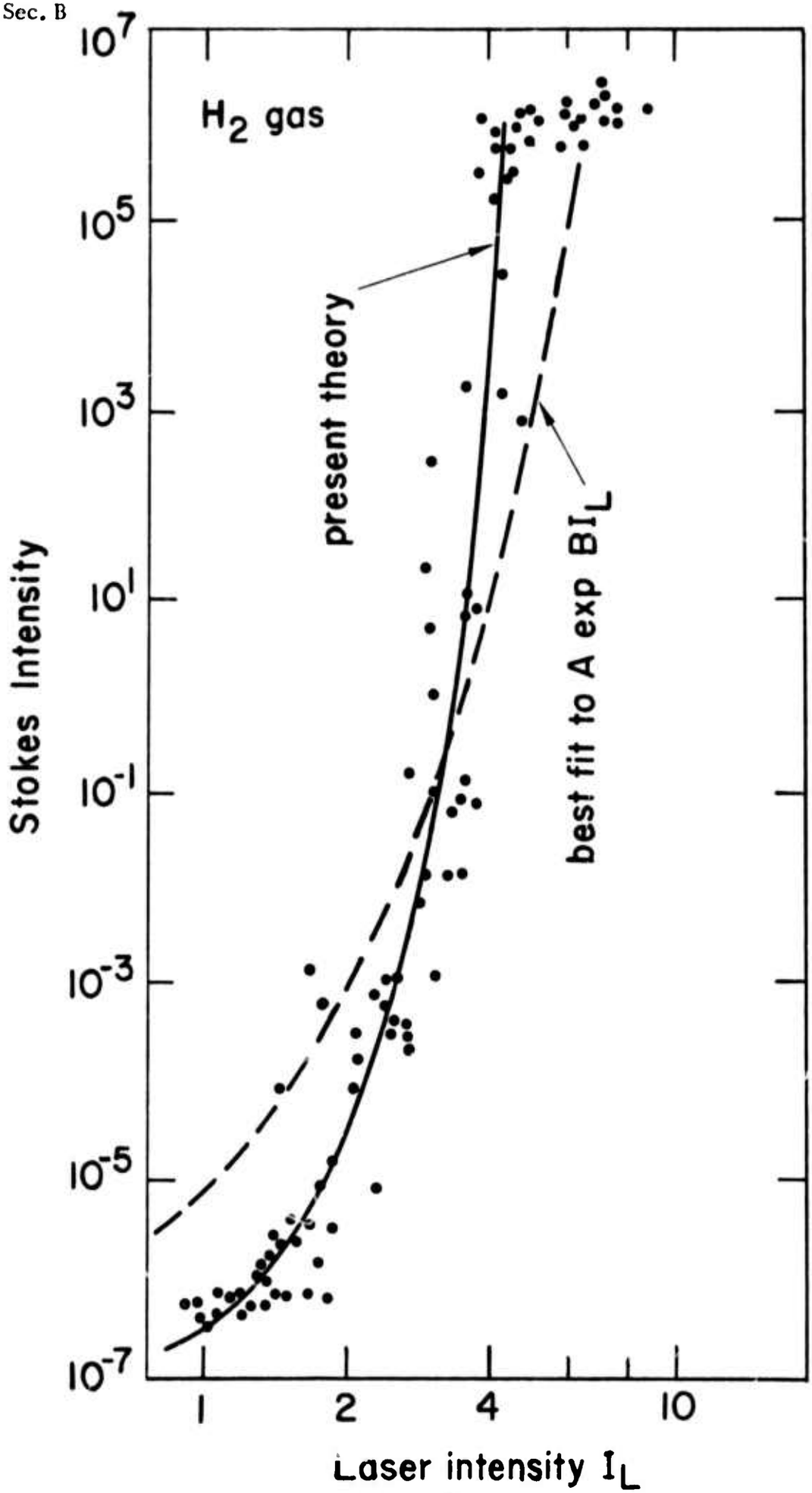

S
L
R
f
P
Q
-Q


L
S
f
L
A


$n_p = \infty$
$n_p \cong 10^5 n_{p1}$
$n_p \cong n_{p1} \equiv (\bar{n}_f + 1) \exp \beta_c L$
$\ln I_S$
I_{LRi}


Liquid N_2
Present Theory
Previous Theory
I_S/I_R
1
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
I_L/I_R


Liquid O_2
1
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
I_S/I_R
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
I_L/I_R


H_2 gas
present theory
best fit to A exp BI_L
Stokes Intensity
10^7
10^5
10^3
10^1
10^{-1}
10^{-3}
10^{-5}
10^{-7}
1
2
4
10
Laser intensity I_L

C. ENHANCED STIMULATED RAMAN SCATTERING AND GENERAL THREE-BOSON PARAMETRIC INSTABILITIES*

M. Sparks and J. H. Wilson
Xonics, Incorporated, Van Nuys, California 91406

A recent theory of stimulated Raman scattering explained a Stokes-intensity enhanced gain that had been observed but that was not predicted by earlier theories. It is shown that in the earlier golden-rule analyses (perturbation theory treatment of occupation numbers) the enhancement was lost by neglecting the increase of the vibrational amplitudes above their thermal equilibrium values. Even though the probability of an individual ion or molecule being excited is small, the occupation number of the phonon in the Raman process is large. In the previous mode-amplitude analyses, the enhancement was lost in the method of linearizing the nonlinear differential equations. By solving these same mode-amplitude equations without using the previous linearization scheme, the enhancement is obtained and the equivalence of the mode-amplitude and golden-rule Boson-occupation-number analyses is demonstrated explicitly. The analysis shows explicitly that the loss of phase information in using the Boson occupation numbers is unimportant. The results are applicable to other three-Boson splitting processes that are important in ferromagnetism, phonon interactions, plasma instabilities, and device physics.

I. INTRODUCTION

Stimulated Raman scattering was first observed[1] and analyzed[2,3] in 1962. It was recently realized[4] that a parametric instability in the Raman Stokes process causes a Stokes-intensity gain enhancement that explains a number of observed anomalies including a nearly discontinuous increase, or "jump," in the Stokes intensity I_S as a function of the laser intensity I_L in the absence of self focusing and feedback. The original purpose of the present investigation was to determine why the early theories[2,3] did not give the gain enhancement, while the later theory[4] did. In addition to identifying the assumptions in both types of the early theories that led to the loss of enhancement, the equivalence of the occupation number (n) rate-equation analysis and the mode amplitude (a and $a^\dagger$) analysis is demonstrated, and the loss of phase information in the occupation-number analysis is shown to be unimportant. The results are of interest in the general three-Boson splitting problem, which arises in a number of fields of physics, as discussed below.

The characteristic feature of a parametric instability is that as the amplitude n_0 of some mode 0 increases, the amplitude n_k of a mode k that is coupled to 0 first increases slowly, then increases rapidly to a great value as n_0 approaches a critical value n_c. For example, in the Raman process a laser photon is annihilated, a Stokes photon is created, and a fundamental (Reststrahl) phonon is created. As the laser-photon occupation number n_L approaches a critical value n_R, the occupation

numbers n_f and n_S of the fundamental-phonon and Stokes-photon modes become very large. This increase in the value of n_S is the gain enhancement and "jump" already mentioned.

Any three-Boson splitting process is potentially unstable parametrically. There are analogies between the instability in the Raman process and previously studied instabilities in ferromagnetic resonance[5-7] (premature saturation of the main resonance subsidiary absorption below the main resonance, and parallel pumping absorption), plasma physics,[8] and electronic devices.[9]

The physical interpretation of these instabilities is rather simple. The balance of energy put into the f phonons by the Raman process against that removed from the f phonons by relaxation is a key to the explanation. The power out by relaxation (by interaction with impurities or other phonons, for example) increases linearly with the number of phonons n_f, which is just the condition that a relaxation time exists. On the other hand, the power into the f phonons increases nonlinearly with increasing n_f since the Raman process is a three-Boson process (which results in products of Boson occupation numbers in the expression for the power). Thus, at a critical value of the laser intensity, the amplitude n_f becomes very large.

Previous analyses of stimulated Raman scattering and other parametric processes used either the equations of motion of the Boson occupation numbers obtained from perturbation theory (the golden rule) or the equations of motion of the mode amplitudes (creation and annihilation operators or Fourier components and their complex conjugates of the electric field, for example). The equivalence of the results has not been demonstrated explicitly in the past. Indeed, in the case of stimulated Raman scattering, the previous results from the mode-amplitude analysis do

not give the enhancement obtained by the recent occupation-number analysis. Furthermore, concern has been expressed that the loss of phase information in using the Boson occupation numbers may be important.[3]

In the present paper the relation between the two approaches is demonstrated explicitly and the points in the previous analyses at which the enhancement was lost are identified. In the previous golden-rule type analyses, the enhancement was lost by neglecting the deviation of the vibrational amplitude from the thermal equilibrium value. Even though n_f becomes large as mentioned above, the probability of an individual ion or molecule being excited is small, roughly speaking. Specifically, $n_f/N \ll 1$ is usually satisfied, where N is the number of unit cells or molecules. It was this fact that the individual ions or molecules are not highly excited that led to the assumption that the thermal equilibrium values were maintained in the previous analyses. There are similar results for other three-Boson processes. For example, in ferromagnetism, magnon occupation numbers are large at the threshold, while the probability of an individual electron spin being in the reversed-spin state is small.

In the previous mode-amplitude analyses, the enhancement was lost in the method of linearizing and decoupling the nonlinear differential equations for the mode amplitudes a_L, a_S, a_f, and their complex conjugates (or Hermitian conjugates in the quantum-mechanical solution). It is shown specifically that reducing the nonlinear equations to parametric linear equations (that is, linear equations with time-dependent coefficients) by assuming that the laser-field amplitude $a_L = b_L \exp(-i\omega_L t)$, where b_L is a constant, results in the loss of the enhancement. The same linearization scheme applied to well known magnon or phonon parametric instabilities results in the loss of the steady-state solution, even though damping is included and a steady-state solution is expected on the basis of simple physical arguments. By solving the same equations

without using this linearization scheme, the difficulties are removed and the equivalence of the mode-amplitude and golden-rule results is demonstrated explicitly for the stimulated-Raman-scattering and magnon problems.

In considering the occupation numbers rather than the mode amplitudes, phase information is lost. A priori it is not expected that the loss of phase information is important since the parametric process amplifies the thermally excited waves that have the proper phase. This was indeed the case in previous analyses. It should be mentioned that wave vector and frequency phase matching are included in the occupation number approach. Wave vector phase matching arises from Kronecker deltas in sums over wave vectors, and frequency phase matching arises from the energy-conserving delta function. The present calculation settles the question by showing explicitly that the phases are unimportant.

The present analysis is concerned only with the steady-state solution. Important results are indicated by underscored equation numbers.

II. FERROMAGNETIC PARAMETRIC INSTABILITIES

In this section the analytical method is applied to the simplest three-Boson parametric instability. The process, illustrated in Fig. C1, is the annihilation of one Boson 0 and the creation of two Bosons having equal frequencies and damping, with negligible propagation of all three Bosons. As specific examples, in the case of ferromagnetic subsidiary-resonance absorption,[5-7] Boson 0 is a uniform precession (wave vector $\underset{\sim}{k} = 0$) magnon, and in parallel pumping, Boson 0 is a photon in the microwave cavity. In both cases, the output Bosons are magnons having wave vectors $\underset{\sim}{k}$ and $-\underset{\sim}{k}$. Propagation effects are negligible since the magnons cannot propagate out of the sample and the sample is small with respect to the electromagnetic wavelength. The process also represents phonon processes[10] and other Boson processes.

It will be demonstrated that the equations of motion of the mode amplitudes can be solved to give the golden rule results (expressions for occupation numbers obtained by use of perturbation theory) directly. The same mode-amplitude equations will be solved by an approximate method of converting nonlinear differential equations into linear differential equations with time-dependent coefficients, or so called parametric equations. This approximate method, which is the same method used in the early treatments of the stimulated Raman scattering, gives incorrect results in the present magnon problem, as it did in the stimulated Raman scattering problem.

The Hamiltonian is

$$\mathcal{H} = \hbar\omega_0 a_0^\dagger a_0 + \hbar\omega_+ a_+^\dagger a_+ + \hbar\omega_- a_-^\dagger a_- - (i\hbar B a_0 a_+^\dagger a_-^\dagger + cc) \qquad (2.1)$$

where $a_i^\dagger$ and a_i with $i = 0, +$, or $-$ are creation and annihilation operators for the 0, $+\underset{\sim}{k}$, and $-\underset{\sim}{k}$ Bosons, respectively, and B is a complex constant that is

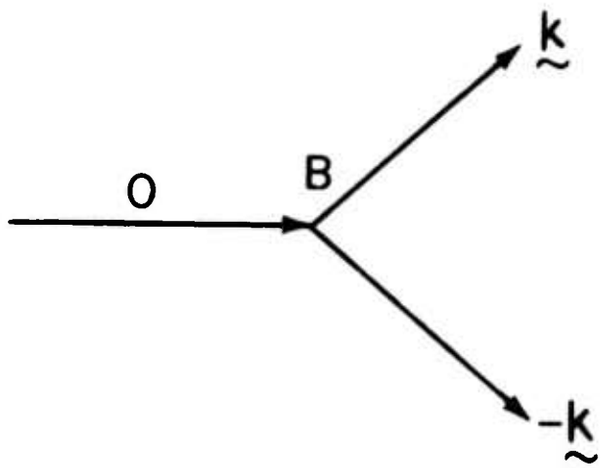


Fig. C1. Three-Boson splitting process that exhibits a parametric instability.

obtained in converting the Hamiltonian from spin operators to magnon creation and annihilation operators.[7] The value of B will not be needed here.

The equations of motion for the a's and $a^\dagger$'s are obtained from the Heisenberg equations of motion:

$$da_+/dt = -i[a_+, \mathcal{H}] \tag{2.2}$$

where $[a_+, \mathcal{H}]$ is the commutator of a_+ and $\mathcal{H}$. Evaluating the commutator in (2.2), with $\mathcal{H}$ given by (2.1), by using standard relations such as $[a_+, a_+^\dagger] = +1$ and $[A, BC] = [A,B]C + B[A,C]$ gives

$$da_+/dt = -i\omega_+ a_+ - Ba_0 a_-^\dagger - \gamma a_+ \quad . \tag{2.3}$$

The term $-\gamma a_+$ was added formally to account for damping. Repeating for $a_-^\dagger$ and a_0 gives

$$da_-^\dagger/dt = i\omega_- a_-^\dagger - B^* a_0^\dagger a_+ - \gamma a_-^\dagger \quad , \tag{2.4}$$

$$da_0/dt = -i\omega_0 a_0 + B^* a_+ a_- \quad . \tag{2.5}$$

This set of equations (2.3)-(2.5) and the Hermitian conjugate equations is a set of six nonlinear differential equations for the six variable a_i and $a_i^\dagger$ with $i = 0, +, -$.

One method of solution is to linearize the equations by formally assuming that $a_0 = b_0 e^{-i\omega_0 t}$ and $a_0^\dagger = b_0^* e^{i\omega_0 t}$ where b_0 and b_0^* are constants. Then (2.3) and (2.4) are a set of two parametric differential equations for a_+ and $a_-^\dagger$. Substituting these expressions for a_0 and $a_0^\dagger$ along with $a_+ = b_+ \exp(-i\omega_+ t)$ and $a_-^\dagger = b_-^* \exp(i\omega_- t)$ into (2.4) and (2.5) and performing the derivatives gives

$$db_+/dt = B\, b_0 b_-^\dagger - \gamma b_+ \tag{2.6}$$

$$db_-^\dagger/dt = B^* b_0^* b_+ - \gamma b_-^\dagger \quad . \tag{2.7}$$

The time-dependent coefficients were eliminated by assuming that the resonance condition $\omega_0 = \omega_+ + \omega_-$ is satisfied. Since these two equations are linear with constant coefficients, their solution is simple. Taking the derivative of (2.6) and using (2.7) gives

$$d^2 b_+/dt^2 + 2\gamma\, db_+/dt + (\gamma^2 - |Bb_0|^2)\, b_+ = 0 \quad . \tag{2.8}$$

Substituting the trial solution $b_+ = b_{+0} \exp(\lambda t)$ into (2.8) and taking the derivatives gives

$$\lambda^2 + 2\gamma\lambda + (\gamma^2 - |Bb_0|^2) = 0$$

which has the solution

$$\lambda = -\gamma \pm |Bb_0| \quad .$$

Thus, the solution to (2.8) is

$$a_+ = e^{-i\omega_+ t}\, e^{-\gamma t} \left[b_+^{(+)} e^{|Bb_0| t} + b_+^{(-)} e^{-|Bb_0| t} \right] \quad . \tag{2.9}$$

There is no non-zero steady-state solution except in the singular case of $|Bb_0| = \gamma$. The common expression "parametric instability" arises from the fact that a_+ becomes infinite as $t \to \infty$ when the amplitude $|b_0|$ of the zero mode is sufficiently great; that is, when

$$|b_0|^2 > \gamma^2/|B|^2 \quad . \tag{2.10}$$

This solution (2.9) clearly is incorrect. Physically, a nonzero steady-state solution is expected for any value of a_0, and this is confirmed experimentally. In the analogous treatment of stimulated Raman scattering in the following section, this parametric linearization method also leads to difficulties. Specifically, the enhancement in the Stokes gain is lost.

To resolve the difficulty, a method of solution other than the parametric linearization used above is needed. One approach would be to linearize (2.3) simply by considering $a_0 a_-^\dagger$ as a single variable. Since $a_0 a_-^\dagger$ is coupled to a_+ according to (2.3), the standard procedure is to consider the equation of motion for $a_0 a_-^\dagger$. If this equation contains only a_+ and $a_0 a_-^\dagger$ (and no product $a_+ a_0 a_-^\dagger$), then the two linear differential equations could be easily solved. Unfortunately the equations do not uncouple at this step. Furthermore, taking the derivatives of the additional variables that appear in the $a_0 a_-^\dagger$ equation couples in still more variables, and the chain of equations becomes large.

A simpler method is to start with the operator $a_+^\dagger a_+ \equiv n_+$, rather than a_+. Then the Heisenberg equations give

$$dn_+/dt = G - \Gamma n_+ \qquad (2.11)$$

where $G \equiv B a_0 a_+^\dagger a_-^\dagger + cc$, and $-\Gamma a_+^\dagger a_+$ is added to account for damping. This same result (2.11) can be obtained from the equation $d a_+^\dagger a_+/dt = a_+^\dagger da_+/dt + cc$ with da_+/dt given by (2.3) and $\Gamma \equiv 2\gamma$. The factor of 2 in $\Gamma = 2\gamma$ arises as usual from the fact that $a \sim \exp(-\gamma t)$ implies $|a|^2 \sim \exp(-2\gamma t)$. This second derivation of (2.11) establishes the consistency of introducing $-\gamma a_+$ into (2.3) and $-\Gamma a_+^\dagger a_+$ into (2.11) and shows the equivalence of the two equations.

From $dG/dt = B(da_0/dt)\,a_+^\dagger a_-^\dagger + Ba_0(da_+^\dagger/dt)\,a_-^\dagger + Ba_0\,a_+^\dagger\,da_-^\dagger/dt + cc$ and (2.3)-(2.5) and their Hermitian conjugate equations, there results

$$dG/dt = 2|B|^2 \left[n_0(n_+ + n_- + 1) - n_+ n_- \right] - \Gamma G \quad . \qquad (2.12)$$

By the same method,

$$dn_-/dt = G - \Gamma n_- \qquad (2.13)$$

$$dn_0/dt = -G \quad . \qquad (2.14)$$

There are several physical situations for which the solutions to (2.11)-(2.14) are of interest. The first is that to which the golden rule is commonly applied. That is, at time $t = 0$ the system is in the state in which $n_0 >> \bar{n}_0$, where $\bar{n}_0$ is the thermal-equilibrium value of n_0, and all other modes are in thermal equilibrium, roughly speaking. The perturbation $-i\hbar Ba_0 a_+^\dagger a_-^\dagger + cc$ is then applied for a time short with respect to the time for n_0 to reach $\bar{n}_0$, but sufficiently long for energy conservation to be well satisfied. The case in which n_0 is maintained at a constant value by the microwave field in the cavity also is of interest. In both of these cases, n_0 is constant, or approximately constant. However, dn_0/dt in (2.14) is not zero because (2.14) is only the contribution to the rate of change of n_0 from the coupling to the $\pm k$ modes. Stated differently, $\hbar\omega_0 dn_0/dt$ from (2.14) gives the power from the zero mode to the pair $\pm\underset{\sim}{k}$, which is not zero in the steady state. The contribution to dn_0/dt from the coupling of the zero mode to the microwave field could be added, but this would carry us too far from the issue at hand.

The steady-state solution to (2.11)-(2.13) is obtained by setting $dG/dt = dn_+/dt = dn_-/dt = 0$ and solving for n_+ and n_-. This gives

$$n_- = n_+ = 2\,|B|^2\,\Gamma^{-2}\left[\, n_0(2n_+ + 1) - n_+^2 \,\right] \tag{2.15}$$

which are the standard results obtained from the golden rule. The complete solution to (2.15) is simple, as discussed in Ref. 10. For the present purpose, it is sufficient to neglect the n_+^2 term, which corresponds to neglecting saturation. Then (2.15) gives directly

$$n_- = n_+ = \frac{1}{2}\,\frac{n_0}{n_c}\,\frac{1}{1 - n_0/n_c}\quad , \qquad n_c \equiv \Gamma^2/4\,|B|^2\quad . \tag{2.16}$$

This same result is obtained simply from the golden rule as follows: The standard expression $(2\pi/\hbar^2)\,|\langle f\,|\,\mathcal{H}\,|\,i\rangle|^2\,\delta(\omega)$ for the transition rate between states $|\,i\rangle$ and $|\,f\rangle$ gives

$$dn_+/dt = (2\pi/\hbar^2)\left[\, |\mathcal{H}|_+^2 - |\mathcal{H}|_-^2 \,\right]\rho(\omega) - \Gamma n_+$$

where $|\,\mathcal{H}\,|_+^2$ is the matrix element for increasing n_+ by one, $|\mathcal{H}|_-^2$ is that for decreasing n_+ by one, and $\rho(\omega)$ is the density of states. For a single transition on resonance ($\omega_0 = \omega_+ + \omega_-$) the appropriate value of $\rho(\omega)$ is[7,6,4] $\rho(\omega) = 1/\pi\Gamma$. Using the usual expressions for the matrix elements of the a's and $a^\dagger$'s gives

$$dn_+/dt = 2\,|B|^2\Gamma^{-1}\left[\, (n_+ + 1)(n_- + 1)\,n_0 - n_+ n_-\,(n_0 + 1) \,\right] - \Gamma n_+\quad . \tag{2.17}$$

Since the bracket factors in (2.17) and (2.12) are equal, the steady-state solution to (2.17) is given by (2.16) as already mentioned.

Even though only the steady-state case is considered here, it should be mentioned that the transient solutions of (2.17) and of (2.11)-(2.14) are different in general. The simplest case of $n_+ n_-$ negligible, $n_+(0)$ = constant, and n_0 = constant can be solved trivially to illustrate this point.

III. MODE AMPLITUDE ANALYSIS OF ENHANCED STIMULATED RAMAN SCATTERING

The Raman scattering process is more complicated than the magnon (or phonon) case just considered because propagation of the laser and Stokes photon must be included. In almost all practical cases the damping of the two photons, but not of the phonon, is negligible; thus the photon attenuation is neglected.

The mode amplitude equations have been obtained classically from Maxwell's equations with terms added to account for the coupling of the electromagnetic and elastic waves.[3,11] Specifically, an interaction Lagrangian was added to the sum of the electromagnetic and elastic Lagrangians and the field-amplitude equations were obtained from the Lagrangian. The resulting second-order partial differential equations were reduced to first-order partial differential equations[11,12] by standard methods.

In order to establish the connection between these first-order partial differential equations for the mode amplitudes and the equations for the occupation numbers, the mode-amplitude equations will be rederived directly from the Hamiltonian

$$\mathcal{H} = \hbar\omega_f a_f^\dagger a_f + \hbar\omega_S a_S^\dagger a_S + \hbar\omega_L a_L^\dagger a_L - (i\hbar V a_L a_f^\dagger a_S^\dagger + cc) \quad . \tag{3.1}$$

Proceeding as in the case of magnons in the previous section, the Heisenberg equations of motion give

$$\frac{\partial a_f}{\partial t} = -i\omega_f a_f - V a_L a_S^\dagger - \frac{1}{2}\Gamma a_f \tag{3.2}$$

$$\frac{\partial a_S}{\partial t} = -i\omega_S a_S - V a_L a_f^\dagger - c_S \frac{\partial a_S}{\partial x} \tag{3.3}$$

$$\frac{\partial a_L}{\partial t} = -i\omega_L a_L + V^* a_f a_S - c_L \frac{\partial a_L}{\partial x} \tag{3.4}$$

where the phonon velocity was neglected in (3.2) as usual[3,4] and the damping term $-\frac{1}{2}\Gamma a_f$ and the propagation terms $-c_S \partial a_S/\partial x$ and $-c_L \partial a_L/\partial x$ were added phenomenologically. (Considering the time derivatives in the Heisenberg equations to be total derivatives, that is $(d/dt)_{c_S} = \partial/\partial t + c_S \partial/\partial x$ for example, would give the c_S and c_L terms in (3.3) and (3.4) directly.)

Converting these equations (3.2)-(3.4), which are identical to the classical equations discussed above, to number-operator equations by using $\partial n_f/\partial t = a_f^\dagger \partial a_f/\partial t + cc$, etc. gives

$$\frac{\partial n_f}{\partial t} = F - \Gamma(n_f - \bar{n}_f) \tag{3.5}$$

$$\frac{\partial n_S}{\partial t} = F - c_S \frac{\partial n_S}{\partial x} \tag{3.6}$$

$$\frac{\partial n_L}{\partial t} = -F - c_L \frac{\partial n_L}{\partial x} \tag{3.7}$$

$$\frac{\partial F}{\partial t} = 2|V|^2 [n_L(n_f + n_S + 1) - n_S n_f] - \frac{1}{2}\Gamma F \tag{3.8}$$

$$F \equiv V a_L a_f^\dagger a_S^\dagger + cc \quad . \tag{3.9}$$

Equation (3.8) was obtained by substituting (3.2)-(3.4) into $\partial F/\partial t = V(\partial a_L/\partial t) a_f^\dagger a_S^\dagger + V a_L(\partial a_f^\dagger/\partial t) a_S^\dagger + V a_L a_f^\dagger \partial a_S^\dagger/\partial t + cc$. The term $Vc_L(\partial_L/\partial x) a_f^\dagger a_S^\dagger + Vc_S a_L a_f^\dagger \partial a_S^\dagger/\partial x + cc$ vanishes since a Stokes photon is created for every laser photon annihilated and the propagation of the two photons

is the same for $c_S \cong c_L$. In (3.5) the term $\Gamma \bar{n}_f$ was added formally to allow n_f to relax to its thermal equilibrium value $\bar{n}_f$, rather than to zero. This is trivial in the occupation-number equation, but incorporating relaxation to a thermal equilibrium value in the amplitude-operator equation (3.2) is quite involved.[5]

Setting the time derivatives of n_f, n_S, and Γ equal to zero gives

$$c_S \frac{\partial n_S}{\partial x} = 4 |V|^2 \Gamma^{-1} [n_L (n_f + n_S + 1) - n_S n_f] \qquad (3.10)$$

and

$$n_R^{-1} [n_L (n_f + n_S + 1) - n_S n_f] - (n_f - \bar{n}_f) = 0 \qquad (3.11)$$

where $n_R \equiv \Gamma^2 / 4 |V|^2$. Neglecting the saturation term $n_S n_f$ in (3.11), as was done in the previous section, and solving for n_f gives

$$n_f = \frac{\bar{n}_f + (n_L / n_R)(n_S + 1)}{1 - n_L / n_R} \quad . \qquad (3.12)$$

The saturation term has been treated elsewhere.[10] These results (3.12) and (3.10) with $n_S n_f$ neglected give the enhanced stimulated Raman scattering result

$$n_S = (\bar{n}_f + 1) [\exp(\beta_{g\,new} x) - 1] + n_S(0) \exp \beta_{g\,new} x \quad , \qquad (3.13)$$

where $\beta_{g\,new} = (\Gamma / c_S) I_{LR} (1 - I_{LR})^{-1}$, with $I_{LR} \equiv n_L / n_R$. The results (3.10)-(3.13) are identical to the results derived previously[4] using the golden rule.

IV. LOSS OF ENHANCEMENT IN PREVIOUS ANALYSES

Previous treatments[2,3] of stimulated Raman scattering did not yield the enhancement obtained in the previous section. These analyses either specifically assumed no increase in the vibrational energy above the thermal equilibrium value or solved equations (3.2) and (3.3) or their equivalents by a method equivalent to that described below. In the former, (3.12) is replaced by

$$n_f = \bar{n}_f \quad .$$

Substituting this expression into (3.10) and solving for n_S gives

$$n_S = (n_f + 1)\,(e^{\beta_{old} x} - 1)$$

where $\beta_{old} = 4|V|^2 n_L/\Gamma c_S = (\Gamma/c_S)(n_L/n_R)$. This is just the Raman gain with no enhancement.

In the latter previous analyses, the nonlinear equations (3.2) and (3.3) were linearized and decoupled from (3.4) by assuming that b_L in the definition

$$a_L \equiv b_L e^{-i\omega_L t} \tag{4.1}$$

is independent of time. Then substituting $a_f = b_f \exp(-i\omega_f t)$ and $a_S^\dagger = b_S^\dagger \exp(i\omega_S t)$ into (3.2) and the Hermitian conjugate of (3.3) gives

$$\frac{\partial b_f}{\partial t} = V b_L b_S^\dagger - \gamma b_f \tag{4.2}$$

$$\frac{\partial b_S^\dagger}{\partial t} = V^* b_L^* b_f - c_S \frac{\partial b_S}{\partial x} \tag{4.3}$$

for the case of resonance, that is $\omega_L = \omega_S + \omega_f$. Setting the time derivatives equal to zero and eliminating b_f from the two equations gives

$$c_S \frac{\partial b_S}{\partial x} = |V|^2 n_L \gamma^{-1} b_S$$

which has the solution

$$b_S(x) = b_S(0) e^{\beta x} \tag{4.4}$$

where $\beta = |V|^2 n_L / c_S \gamma$, which shows no gain enhancement.

In order to further show how the assumption (4.1) causes the loss of enhancement, (4.2) and (4.3) with b_L independent of time will be solved by another method, in direct analogy with the solution of Sec. III where the time dependence of b_L was retained. By the same method used in Sec. III, (4.2) and (4.3) give

$$\frac{\partial n_f}{\partial t} = F - \Gamma n_f \tag{4.5}$$

$$\frac{\partial n_S}{\partial t} = F - c_S \frac{\partial n_S}{\partial x} \tag{4.6}$$

$$\frac{\partial F}{\partial t} = 2|V|^2 n_L (n_f + n_S + 1) - \frac{1}{2}\Gamma F - (c_S V b_L b_f \frac{\partial b_S}{\partial x} + cc) \quad . \tag{4.7}$$

By neglecting the time dependence of b_L, the nonlinear term $-2|V|^2 n_f n_S$ in (3.8) is lost and the last term in (4.7), which did not appear in (3.8) is gained. The former makes the solution incorrect in the saturation region and the latter eliminates the enhancement. Neglecting the time dependence of b_L is equivalent to neglecting the last two terms in Eq. (3.4) for da_L/dt. By neglecting the last

term $c_L \partial a_L / \partial x$ in (3.4), the cancellation of the similar term $c_S \partial a_S / \partial x$ from (3.3) does not occur in the equation for $\partial F / \partial t$; thus, the last term in (4.7) is present. Neglecting the other term $V^* a_f a_S$ in (3.4) corresponds directly to the absence of the term $2 |V|^2 n_S n_f$ in (4.7). This discussion indicates that the physical significance of the linearization by using (4.1) is that the effect of increases in the amplitudes a_f and a_S on the amplitude a_L is neglected and the spatial rate of change of a_L is neglected while a comparable term of a_S is retained.

The loss of enhancement in (4.5)-(4.7) can be seen by setting the time derivatives equal to zero, $\partial n_S / \partial x = \beta n_S$, and $\partial b_S / \partial x = \frac{1}{2} \beta b_S$ in (4.5)-(4.7) and eliminating F. This gives

$$2 |V|^2 n_L (n_f + n_S + 1) - \frac{1}{2} (\Gamma + c_S \beta) c_S \beta n_S = 0$$

$$c_S \beta n_S = \Gamma n_f \quad .$$

Neglecting 1 with respect to $n_f + n_S$, which is well satisfied, and eliminating n_f gives

$$2 |V|^2 \Gamma^{-1} n_L (c_S \beta + \Gamma) - \frac{1}{2} (c_S \beta + \Gamma) c_S \beta = 0 \quad .$$

Dividing by $\frac{1}{2} c_S (c_S \beta + \Gamma)$ gives

$$\beta = 4 |V|^2 / c_S \Gamma$$

in agreement with (4.4).

REFERENCES

*This section has been submitted for publication in The Physical Review.

1. E. J. Woodbury and W. K. Ng, Proc. IRE 50, 2367 (1962).
2. R. W. Hellwarth, Phys. Rev. 130, 1850 (1963).
3. Y. R. Shen and N. Bloembergen, Phys. Rev. 137, A1787 (1965); N. Bloembergen, Am. J. Phys. 35, 989 (1967); F. De Martini, Phys. Rev. B 4, 4556 (1971).
4. M. Sparks, Phys. Rev. Lett. 32, 450 (1974), and Phys. Rev., in press. In an erratum to the letter, the effect of optical dispersion was overlooked. Thus, the reservations expressed in the erratum are not appropriate.
5. H. Suhl, J. Phys. Chem. Solids 1, 209 (1959).
6. H. B. Callen, Fluctuation, Relaxation and Resonance in Magnetic Systems, edited by D. ter Haas (Oliver and Boyd, Edinburgh, 1962); R. M. White and M. Sparks, Phys. Rev. 130, 623 (1963).
7. M. Sparks, Ferromagnetic Relaxation Theory (McGraw-Hill, New York, 1964).
8. S. V. Silin, Sov. Phys. JETP 21, 1127 (1965); D. F. DuBois and M. V. Goldman, Phys. Rev. Lett. 14, 544 (1965).
9. W. H. Louisell, Coupled Modes and Parametric Electronics (Wiley, New York, 1960).
10. M. Sparks and H. C. Chow, Phys. Rev. B 10, 1699 (1974).
11. N. M. Kroll, J. Appl. Phys. 36, 34 (1965).
12. C. S. Wang, Phys. Rev. 182, 482 (1969).

D. THEORY OF LASER-MATERIALS DAMAGE BY ENHANCED STIMULATED RAMAN SCATTERING*

M. Sparks
Xonics, Incorporated, Van Nuys, California 91406

An analysis indicates that Raman active crystals fail at intensities I_f which are greater than the enhanced stimulated Raman-scattering threshold intensity I_R by an amount I_T that is generally of the order of or less than I_R. A typical value of I_R at the ruby frequency is a few gigawatts per square centimeter, which is less than other intrinsic-mechanism thresholds. At intensities $I > I_R$, the excess intensity $I - I_R$ is converted into Stokes radiation and phonons in a distance $\ell << d$, where d is the thermal diffusion distance for a 10 nsec pulse. The temperature rise from the rapidly thermalized phonons in the volume $x < d$ is sufficient to cause material failure when $I > I_f$.

Stimulated Raman scattering was first observed in 1962.[1] Within the last year there has been interest in a parametric instability in the Raman Stokes processes that causes a Stokes-intensity gain enhancement.[2-5] Among the significant consequences are the explanation of the previously anomalous nearly discontinuous increase in the Stokes intensity I_S as a function of the laser intensity I_L in the absence of self focusing and feedback. It was also mentioned that this enhanced stimulated Raman scattering could be an important laser-damage mechanism having a threshold lower than those of other well known damage mechanisms.

In the present paper an analysis of the enhanced stimulated Raman scattering shows that Raman active samples fail at intensities I_f that are greater than the enhanced stimulated Raman scattering threshold intensity I_R by an amount I_T which is usually of the order of or less than I_R. Typical values of I_R for solids are a few gigawatts per square centimeter at the ruby frequency.

Physically the mechanism for failure is simply that at intensities above the threshold the excess intensity above I_R is converted into Stokes radiation in a very thin layer of the sample near its surface. Since one phonon, called an f phonon, is created for each Stokes photon created in the Raman process and the phonons rapidly thermalize, the temperature in the thin surface layer increases.

The general steady-state equations are[3,5]

$$dn_S/dx = (\beta_c/\gamma n_{R0})[(n_S + n_f + 1)n_L - n_S n_f] \tag{1}$$

$$n_f/n_p = (\gamma^2 - 1)/\gamma^2 = A/\gamma^2 - B \tag{2}$$

where n_S, n_L, and n_f are the Bose Einstein occupation numbers for the Stokes photons, laser photons, and fundamental phonons, $\beta_c \equiv \Gamma_f/c_S$, Γ_f is the linear (low intensity) relaxation frequency of the f phonon, c_S is the velocity of the Stokes wave, $n_R = \gamma^2 n_{R0}$, n_{R0} is the threshold value of n_L for enhanced stimulated Raman scattering in the absence of phonon instabilities (which make $\gamma^2 > 1$),[5] n_p is the threshold value of n_f for the phonon instability,[5] $A \equiv n_S n_L/n_p n_{R0}$, and $B = (n_L - n_S)/n_{R0}$. Estimations of the time constants indicate that the assumption that the steady-state results apply is well satisfied.

The solution to (2) is

$$\gamma^2 - 1 = \tfrac{1}{2}(A+B+1) \pm [\tfrac{1}{4}(A+B+1)^2 - B]^{1/2} \quad . \tag{3}$$

The inequalities $n_S << n_p$ and $n_L << n_p$ are well satisfied in all cases of interest except the case of ultralong-lifetime phonons in some materials at low temperatures.[6,7] Since $n_S << n_p$ is extremely well satisfied ($n_p / n_S \sim 10^7$, typically), the approximation

$$n_S << n_p (1 - n_S / n_L) \tag{4}$$

is well satisfied until saturation ($n_S = n_L$) is approached very closely. When (4) is satisfied, then $A << B$, and (3) gives

$$\gamma^2 - 1 \simeq \tfrac{1}{2}(B - 1) + \tfrac{1}{2} A \pm [\tfrac{1}{4}(B - 1)^2 + \tfrac{1}{2} A (B + 1)]^{1/2} \quad . \tag{5}$$

The following limiting cases are of interest: For $1 - B >> [2A(B + 1)]^{1/2}$, (5) and (2) give

$$\gamma^2 - 1 \simeq A/(1 - B) \quad , \qquad n_f \simeq n_S n_L (n_{R0} - n_L + n_S)^{-1} \quad . \tag{6}$$

For $B = 1$, $\gamma^2 - 1 \simeq A^{1/2}$, $n_f \simeq A^{1/2} n_p$, and for $B - 1 >> [2A(B + 1)]^{1/2}$, (4) gives

$$\gamma^2 - 1 \simeq B - 1 \quad , \qquad n_f \simeq n_p \quad . \tag{7}$$

These results show that n_f is small for $B < 1$ and large for $B > 1$, the transition region being very narrow, of order $2A^{1/2}$. See Fig. D1, where in curve (a) the ordinate side is such that A is visible. Then n_p is far off the scale. In (b) the ordinate scale is such that n_p is visible. Then A is indistinguishable from zero. With 1 visible on the abscissa scale, the width of the transition region, in which n_f increases from its small value to n_p, is too small to be observable.

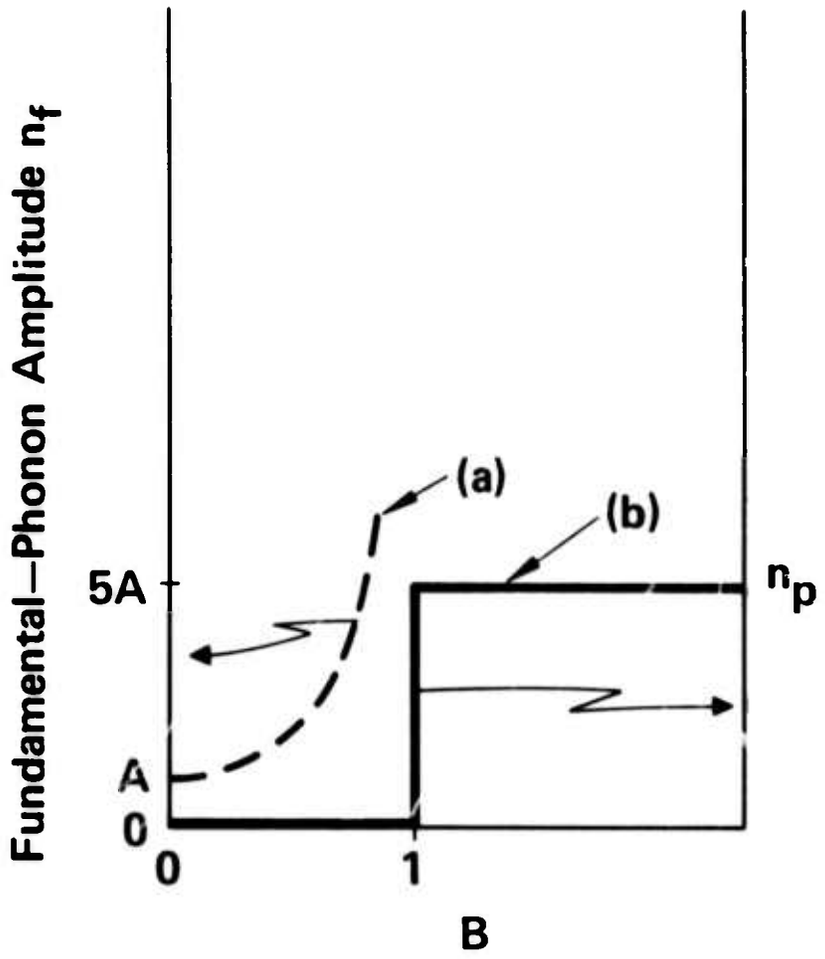


Fig. D1. Increase in f-phonon-mode occupation number n_f with increasing $B \equiv (n_{LO} - 2n_S)/n_{RO}$. The two ordinate scales for the (a) and (b) curves are explained in the text.

The low-B result (6), which is valid when $n_{L0} \lesssim n_{R0}$, along with the differential equation (1), gives the previous enhanced stimulated Raman results.[3,2] In this case n_S increases exponentially with increasing distance x from the sample surface, with the characteristic distance β_c^{-1}. Values of β_c^{-1} range from $\sim 10^{-2}$ to 1 cm for typical solids and liquids. These values usually are not sufficiently small to cause sample damage; thus the detailed results for this case need not be repeated here.

Substituting the high-B results $n_f \cong n_p$ and $\gamma \cong B^{1/2}$ from (7) into (1) and using $n_S + 1 << n_p$ gives

$$dn_S/dx = \beta_c n_p n_{R0}^{-1/2} (n_{L0} - 2n_S)^{1/2} \quad .$$

The solution with the boundary condition $n_S = 0$ at $x = 0$ is

$$n_S = \tfrac{1}{2} n_{L0} \left\{ 1 - [1 - n_p (n_{L0} n_{R0})^{-1/2} \beta_c x]^2 \right\} \quad . \tag{8}$$

According to (8), the value of n_S increases from $n_S = 0$ at $x = 0$ to $n_S = \frac{1}{2} n_{L0}$ at $x = \ell$, where

$$\ell \equiv (n_{L0} n_{R0})^{1/2} / n_p \beta_c \quad . \tag{9}$$

This solution, which is sketched in Fig. D2, is valid for $B - 1 >> [2A(B+1)]^{1/2}$, or $B > 1$ approximately. For example, (8) is valid for n_S up to one half of the asymptotic value of $\frac{1}{2} n_{L0}$ for $n_{L0} = 2 n_{R0}$ or up to 9/10 of $\frac{1}{2} n_{L0}$ for $n_{L0} = 10 n_{R0}$. For greater values of n_S, the increase in n_S is much slower. Consider the case of $n_{L0} = 2 n_{R0}$ illustrated in Fig. D3 as an example. From the definition under (2), $B = 2(1 - n_S/\frac{1}{2} n_{L0})$. For $n_S \lesssim n_{L0}/4$, B satisfies $B \gtrsim 1$, and n_S increases rapidly according to (8). For $n_S \gtrsim n_{L0}/4$, B satisfies $B \lesssim 1$, and n_S increases

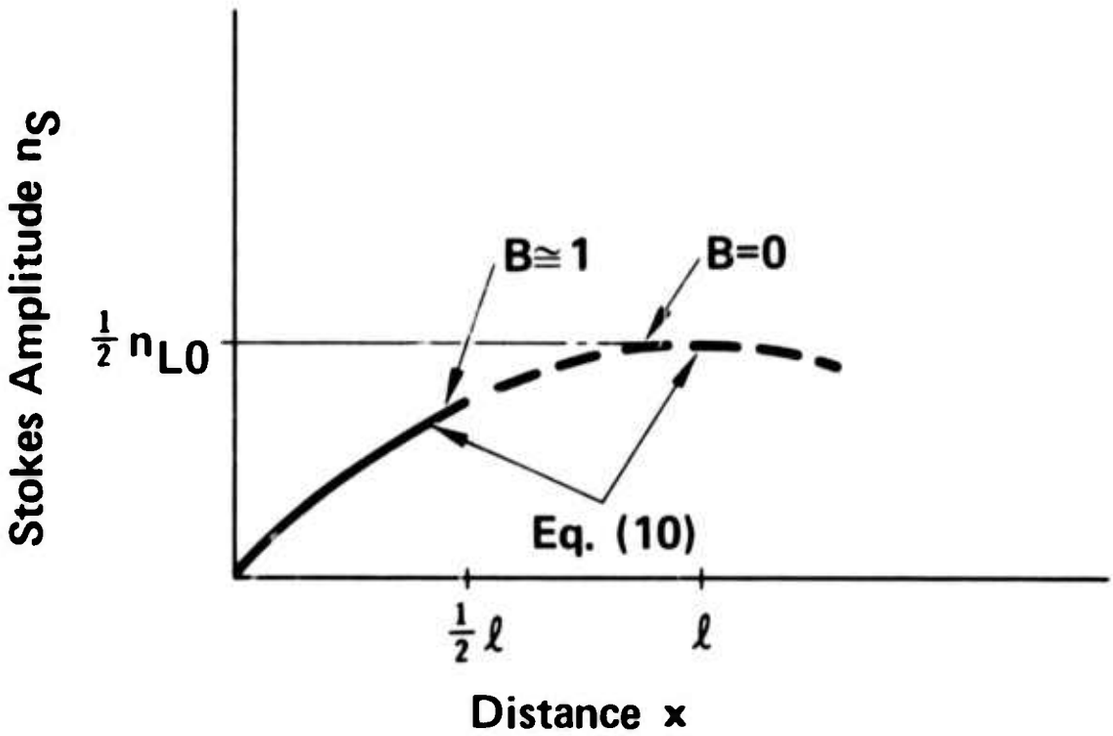


Fig. D2. Distance required for n_S to approach its asymptotic value $\frac{1}{2} n_{L0}$. The solution (10) is valid to the left of B ≅ 1. At $x = \frac{1}{2}\ell \frac{1}{2} \beta_c^{-1} (n_{L0} n_{R0})^{1/2} / n_p$, $n_S = 0.75 (\frac{1}{2} n_{L0})$.

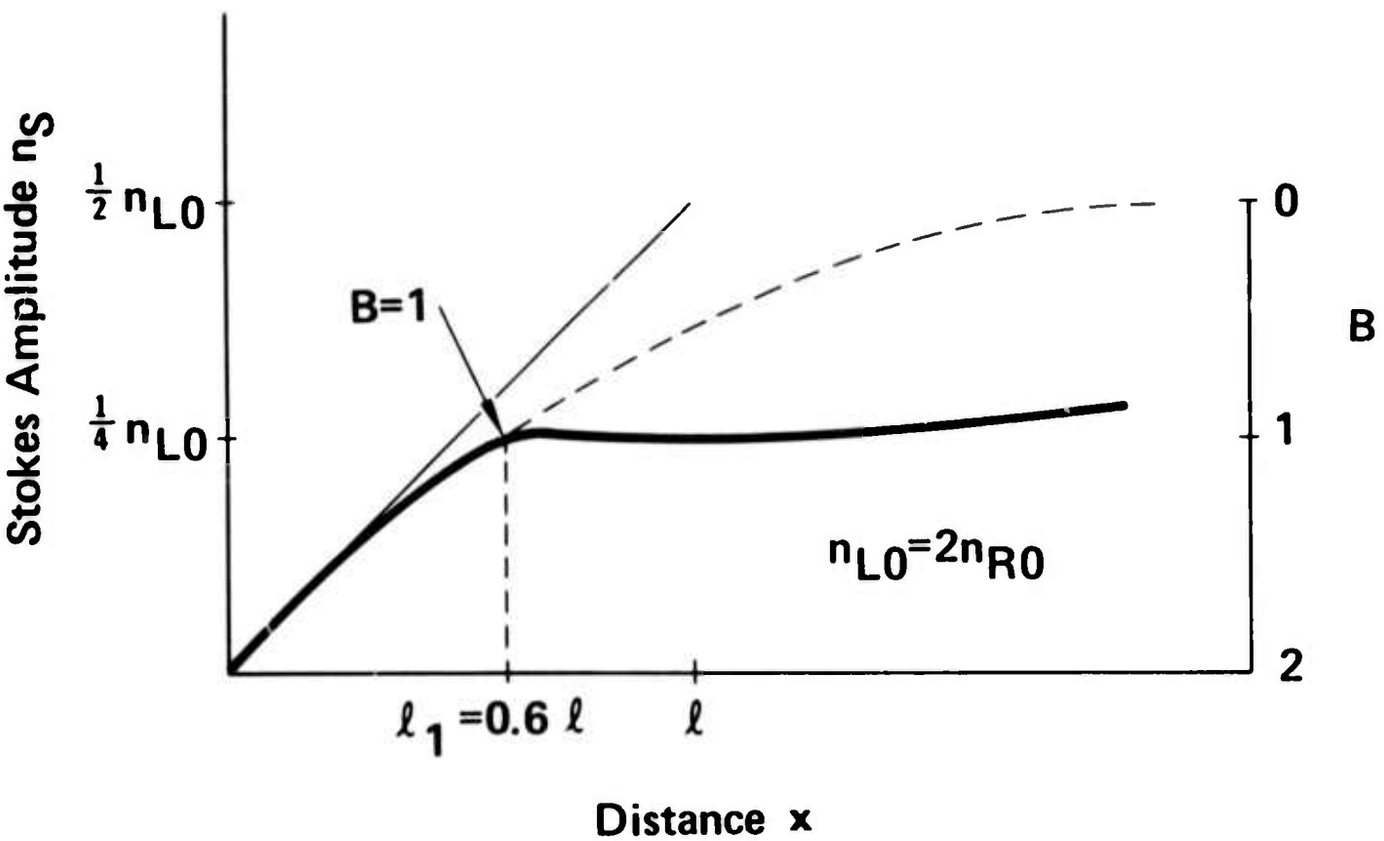


Fig. D3. Rapid increase of Stokes amplitude n_S for $x < \ell_1$ (i.e., $B > 1$) and slow increase for $x > \ell_1$ (i.e., $B < 1$).

much more slowly according to the previous low-B result.[3,2] Thus, n_S increases to the value $n_{L0}/4$ in a distance (0.6ℓ), which is of order ℓ, and thereafter increases very slowly.

For the general case of $n_{R0} = \eta n_{L0}$ with $\eta > 1$, the value of B is $B = \eta(1 - n_S/\frac{1}{2}n_{L0})$ from the definition under (2). Thus $B \gtrsim 1$ for $n_S \lesssim n_{Sx}$, where

$$n_{Sx} \equiv \tfrac{1}{2} n_{L0} (1 - \eta^{-1}) \quad . \tag{10}$$

This result (10) shows that a fraction $\frac{1}{2}(1 - \eta^{-1})$ of the incident laser radiation is converted to Stokes radiation in a distance $\ell_\eta < \ell$. An expression for ℓ_η, which is easily obtained from (8), will not be needed since $\ell_\eta < \ell$ and the thermal diffusion distance $d >> \ell$. The laser amplitude is reduced from n_{L0} to n_{R0} in the region $x < \ell_\eta$.

Since one phonon of energy $\hbar\omega_f$ is created for each laser photon of energy $\hbar\omega_L$ annihilated, the energy

$$\mathcal{E} = I\, \mathcal{A} t_p \tfrac{1}{2}(1 - \eta^{-1})\, \omega_f/\omega \tag{11}$$

is converted to fundamental phonons (and subsequently to heat) in the volume $\mathcal{A}\ell_\eta$, where $\mathcal{A}$ is the surface area.

The temperature rise ΔT is obtained from the heat capacity relation $\mathcal{E} = CV\Delta T$, where C is the heat capacity per unit volume. The volume V is equal to $\mathcal{A}d$, where d is the thermal diffusion distance for the usual case of $\ell_\eta << d$. With (11) and $I = I_R \eta$, where $I_R = \hbar\omega_L c_L n_{R0}/V$, this gives

$$\Delta T = (1/I_R - 1)\, \Delta T_2 \tag{12}$$

where

$$\Delta T_2 \equiv \frac{1}{2} I_R t_p \omega_f / \omega C d \quad . \tag{13}$$

This result is valid for $\eta > 0$; that is, $I > I_R$. For $I < I_R$, ΔT is sufficiently small that $\Delta T = 0$ can be used for the present purpose. See Fig. D4. At $I = 2 I_R$, $\Delta T = \Delta T_2$.

The thermal diffusion distance d is obtained from the standard relation between time and diffusion distance[8]

$$t_{pulse} = 4 C d^2 / \pi K$$

or

$$d = (\pi K t_{pulse} / 4 C)^{1/2} \tag{14}$$

where K is the thermal conductivity. For most materials $C \cong 2\ J/cm^3 K$. Thus, for $t_{pulse} = 10$ nsec, (14) gives $d \cong 1\ \mu m$ for $K = 1$ W/cm K or $d = 0.1\ \mu m$ for $K = 10^{-2}$ W/cm K.

For $\Delta T = \Delta T_f$, where ΔT_f is the value of ΔT at which the crystal fails, (12) gives

$$I_f = I_R + I_T \quad , \qquad I_T \equiv 2 C d \Delta T_f \omega / \omega_f t_p = (\pi K C / t_{pulse})^{1/2} \Delta T_f \omega / \omega_f \quad . \tag{15}$$

This central result (15) shows that the failure intensity I_f is greater than the Raman instability threshold I_R by an amount I_T, whose value is determined by the frequency, pulse length, material parameter (C, K, and ω_f), and the failure temperature ΔT_f.

For CaF_2, $C = 2.71\ J/cm^3 K$, $K = 0.09$ W/cm K, $\omega_f = 260\ cm^{-1}$, and the melting temperature is $T_{melt} = 1400$ C. A complete study of failure would be complicated and the results would vary from sample to sample and application to application.

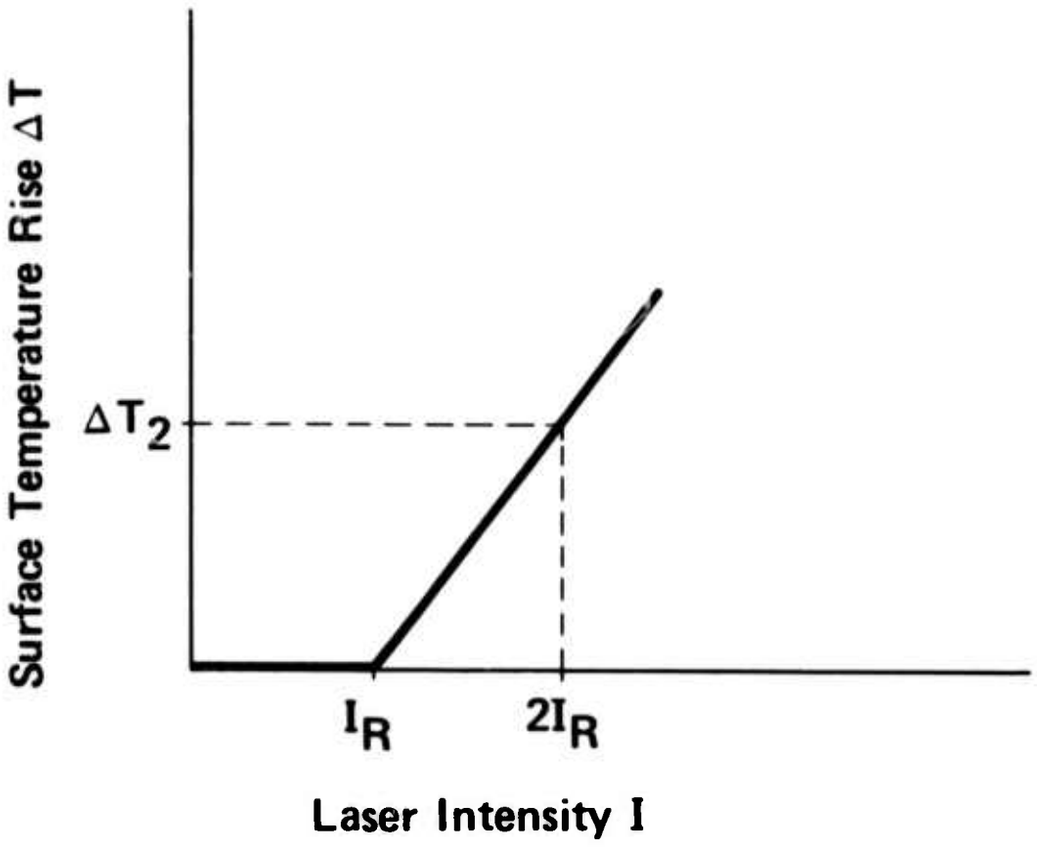


Fig. D4. Increase in surface temperature resulting from the enhanced stimulated Raman scattering.

In order to obtain a rough estimate of the size of I_T, it is assumed that the melting of an infinitely thin (negligible heat of fusion) layer of the surface constitutes failure. Thus $\Delta T_f = T_{melt} - T_{ambient} \cong 1,400\,K$. With these values and $t_{pulse} = 10$ nsec at the ruby frequency $\omega = 1.44 \times 10^4\,cm^{-1}$, (15) gives $d = 0.16\,\mu m$ and (14) gives

$$I_T = 0.7\ GW/cm^2 \quad . \tag{16}$$

Analysis of the stresses in this case of a two-dimensional (radial and axial) temperature distribution is complicated and beyond the scope of the present paper. However, rough estimates indicate that the sample may fracture before melting, thereby lowering the value of I_f. A detailed study is unwarranted at present in view of the uncertainty in the value of I_R. In fact, the value of the critical intensity I_R is not known for CaF_2 and many other materials. Furthermore, $I_T \lesssim I_R$ for most materials.

In order to obtain a rough estimate of I_f for use until measurements of I_R are available, a typical value of $1\,GW/cm^2$ will be used. Then (14) and (15) give for CaF_2

$$I_f \cong 1.7\,GW/cm^2 \quad . \tag{17}$$

For diamond, $C = 1.56\,J/cm^3\,K$, $K = 20\,W/cm\,K$, $\omega_f = 1330\ cm^{-1}$, and $T_{melt} = 3730\,C$. The value of I_R is approximately 1 to $4\,GW/cm^2$. For $t_{pulse} = 10$ nsec and $\omega = 1.44 \times 10^4\,cm^{-1}$, (15) gives $d = 3.2\,\mu m$, and (14) gives

$$I_T = 4\,GW/cm^2 \quad , \qquad I_f \cong 5\text{-}8\ GW/cm^2 \quad . \tag{18}$$

For ZnSe, $C = 2.6\,J/cm^3\,K$, $K = 0.13\,W/cm\,K$, $\omega_f = 207\ cm^{-1}$, and $\Delta T_{melt} = 1500\,K$. Thus $d = 0.2\,\mu m$ and

$$I_T = 1.1\ GW/cm^2 \quad .$$

Then if $I_R \cong 1\,GW/cm^2$, $I_f \cong 2\,GW/cm^2$.

Several additional comments are of interest. First recall that rocksalt-structure crystals such as NaCl are not Raman active (first order). Thus the enhanced stimulated Raman scattering process should not limit the value of I_f in these crystals. Most other crystals of current interest are Raman active. Macroscopic Raman-active impurities in rocksalt-structure crystals could of course lower the damage threshold.

Conversion of the Stokes radiation S to second Stokes radiation S^2, conversion of S^2 to S^3, and so forth, could cause additional heating of the crystal. However, these subsequent conversion processes are not usually important since the inequality $I_S < I_R$ usually is satisfied; i.e., the threshold for conversion from S to S^2 is not reached. This is because the inequality $I_T \lesssim I_R$ usually is satisfied. However, the value of I_T could be great for materials with large values of K and small values of ω_f, as seen in (15).

Thermally induced optical distortion[9] is negligible in general because d is so small. The optical distortion is determined by the average over the sample thickness of the temperature rise on the optical axis of the window. For example, for a 1 cm-thick sample of ZnSe with $d = 0.2\,\mu m$ and $\Delta T = 1500\,K$, the average value of the temperature rise is $1500\,(2 \times 10^{-5}\,cm/1\,cm) = 0.03\,K$, which causes negligible optical distortion.[9] Finally, the relatively small temperature variation of C and K, which were neglected, could be included if greater accuracy is required at a later date.

REFERENCES

This section has been submitted for publication in the Journal of Applied Physics.

1. E. J. Woodbury and W. K. Ng, Proc. IRE 50, 2367 (1962).
2. M. Sparks, Phys. Rev. Lett. 32, 450 (1974).
3. M. Sparks, Phys. Rev. A 10, xxx (1974).
4. M. Sparks and J. H. Wilson (unpublished).
5. M. Sparks and H. C. Chow, Phys. Rev. B 10, 1699 (1974).
6. L. A. Vredevoe, Phys. Rev. 140, A930 (1965).
7. R. Orbach and L. A. Vredevoe, Physics 1, 91 (1964).
8. H. S. Carslaw and J. C. Jaeger, Conduction of Heat in Solids, 2nd Edition (Oxford-Clarendon, 1959). The value of the constant $4/\pi$, gives the correct surface temperature, as shown by M. Sparks, J. Appl. Phys. 44, 4137 (1973).
9. M. Sparks, J. Appl. Phys. 42, 5029 (1971).

E. SURFACE ROUGHNESS AND THE OPTICAL PROPERTIES OF A SEMI INFINITE MATERIAL; THE EFFECT OF A DIELECTRIC OVERLAYER†

D. L. Mills and A. A. Maradudin
Department of Physics
University of California
Irvine, California 92664

and

Xonics Corporation
Van Nuys, California 91406

† Supported in part by Contract No. DAHC15-73-C-0127 of the Advanced Research Projects Agency of the Department of Defense, and in part by Grant No. AFOSR 71-2018 of the Air Force Office of Scientific Research, Office of Aerospace Research, U.S.A.F.

Technical Report No. 74-53

Abstract

We derive expressions for the rate at which radiation is scattered and absorbed because of surface roughness on a semi-infinite material, in the presence of a dielectric overlayer. We confine our attention to the case of normal incidence. A formalism developed in an earlier paper by the present authors is utilized in the discussion. We also present a series of numerical calculations which explore the roughness-induced scattering and absorption of electromagnetic radiation for aluminum overcoated by aluminum oxide, in the ultraviolet region of the spectrum. The position of the reflectivity dip produced by roughness induced coupling to the surface plasmon is found to shift toward the visible as the thickness of the oxide layer increases. The size of the dip is controlled strongly by the degree of correlation between the roughness on the vacuum-oxide interface, and that on the oxide-substrate interface. Under conditions discussed in the text of the paper, the presence of the oxide layer can greatly enhance the coupling between the incident radiation and surface plasmons.

I. Introduction

In the presence of roughness on the surface of a material, light incident on the substance may be scattered away from the specular direction, and roughness-induced absorption can occur. Both effects reduce the reflectivity of the material below the intrinsic value expected for a semi-infinite sample with perfectly smooth surface. The effect is particularly severe for aluminum in the ultra-violet region of the spectrum, since in the presence of surface roughness, the incident light may couple to the surface plasmon with remarkable efficiency.[1]

There has been renewed interest in this problem recently, in part because of the need for highly reflecting mirror materials for use in the construction of cavities for lasers which operate in the ultra-violet region of the spectrum. While aluminum has the highest intrinsic reflectivity of any material in the near ultra-violet, roughness-induced coupling of the incident radiation to surface plasmons can decrease its reflectivity significantly, unless "supersmooth" surfaces are prepared.[2]

Another method that has proved useful in increasing the reflectivity of aluminum films is to overcoat them with a dielectric layer after a very smooth "bare" aluminum surface has been prepared. Such an overlayer will also be present anytime the aluminum has been exposed to an atmosphere that permits oxidation of the surface. One is then led to inquire about the effect of such a dielectric layer on the surface roughness induced coupling to surface plasmons. While this is a topic that has been explored experimentally[2],[3],[4], we know of no theoretical treatment of the effect of a dielectric overlayer on the roughness induced scattering and absorption of light.

There seems a critical need for such a theoreti al analysis, in our view. From simple considerations, one expects that a dielectric overlayer on a metal such as aluminum will shift the reflectivity dip (produced by the roughness induced coupling to surface plasmons) toward the visible. One would like to calculate the magnitude of this shift for an overlayer of given thickness, and a specified configuration of surface roughness. Perhaps more important to understand is the relationship of the magnitude of the dip to the nature of the roughness on the overlayer-substrate and overlayer-vacuum interfaces.

The purpose of this paper is to present such a theory by extending our earlier treatment[5] of the surface roughness induced absorption and scattering of electromagnetic radiation to the case where a dielectric overlayer is present on the surface of the material of interest. In the interest of simplicity, we confine our attention here to the case where the radiation is normally incident on the surface. For this case, we obtain formulas for the angular distribution and polarization of the radiation scattered from the rough surface into the vacuum above the material, for the fraction of the incident radiation flux absorbed within the film, and the fraction of the incident radiation flux absorbed by the substrate material. The treatment is valid in the limit that the amplitude of the surface roughness is very small.

We also present a series of numerical studies of the absorption and scattering of radiation in the near ultraviolet (5-12eV) by an oxidized surface of aluminum. We find here that the magnitude of the reflectivity dip produced by roughness induced coupling of

the incident radiation to surface plasmons depends very dramatically on the manner in which the roughness of the vacuum-oxide overlayer is correlated with that on the oxide-substrate interface.

Before we proceed with the detailed discussion, we elaborate on this remark a bit. Consider a smooth oxide-vacuum interface parallel to the x-y plane located at the position $z = d$, while the smooth oxide-substrate interface is also parallel to the x y plane at $z = o$. Now roughen each interface, where $\zeta_1(x,y)$ measures the position of the oxide-vacuum interface at the point x,y above the plane $z = d$. Similarly, $\zeta_2(x,y)$ denotes the position of a point on the roughened oxide-substrate interface above the plane $z = o$. Then if we denote averages over a given interface by angular brackets, we presume $\langle\zeta_1\rangle = \langle\zeta_2\rangle = o$. In our numerical calculations, we examine the following four situations, illustrated schematically in Figure (1):

(i) $\zeta_1(x,y) \equiv \zeta_2(x,y)$ everywhere. We refer to this as the replicating film model (Figure 1(a)).

(ii) $\zeta_1(x,y) = -\zeta_2(x,y)$ everywhere. We call this the non-uniform film model. (Figure 1(b)). This might be a crude description of a lumpy oxide overlayer.

(iii) $\langle\zeta_1^2\rangle = \langle\zeta_2^2\rangle$, but $\zeta_1(x,y)$ and $\zeta_2(x,y)$ very randomly with respect to each other, so the cross correlation function $\langle\zeta_1\zeta_2\rangle$ vanishes everywhere. We call this the random roughness model (Figure 1(c)).

(iv) $\zeta_2 \equiv o$ but $\zeta_1 \neq o$, i.e. the oxide-substrate interface is perfectly smooth, but the surface of the oxide is rough. We refer to this as the rough oxide layer model (Figure 1d). It serves as a model of a supersmooth aluminum surface overcoated with a non-uniform oxide film.

the incident radiation to surface plasmons depends very dramatically on the manner in which the roughness on the vacuum-oxide overlayer is correlated with that on the oxide-substrate interface.

Before we proceed with the detailed discussion, we elaborate on this remark a bit. Consider a smooth oxide-vacuum interface parallel to the x-y plane located at the position z = d, while the smooth oxide-substrate interface is also parallel to the x y plane at z = o. Now roughen each interface, where $\zeta_1(x,y)$ measures the position of the oxide-vacuum interface at the point x,y above the plane z = d. Similarly, $\zeta_2(x,y)$ denotes the position of a point on the roughened oxide-substrate interface above the plane z = o. Then if we denote averages over a given interface by angular brackets, we presume $\langle\zeta_1\rangle = \langle\zeta_2\rangle = o$. In our numerical calculations, we examine the following four situations, illustrated schematically in Figure (1):

(i) $\zeta_1(x,y) \equiv \zeta_2(x,y)$ everywhere. We refer to this as the replicating film model (Figure 1(a)).

(ii) $\zeta_1(x,y) = -\zeta_2(x,y)$ everywhere. We call this the non-uniform film model. (Figure 1(b)). This might be a crude description of a lumpy oxide overlayer.

(iii) $\langle\zeta_1^2\rangle = \langle\zeta_2^2\rangle$, but $\zeta_i(x,y)$ and $\zeta_2(x,y)$ very randomly with respect to each other, so the cross correlation function $\langle\zeta_1\zeta_2\rangle$ vanishes everywhere. We call this the random roughness model (Figure 1(c)).

(iv) $\zeta_2 \equiv o$ but $\zeta_1 \neq o$, i.e. the oxide-substrate interface is perfectly smooth, but the surface of the oxide is rough. We refer to this as the rough oxide layer model (Figure 1d). It serves as a model of a supersmooth aluminum surface overcoated with a non-uniform oxide film.

When we compare the results of the calculations for the four cases described above, the position of the reflectivity dip is very nearly the same for each case, for an overlayer of given thickness. However, the magnitude of the dip differs markedly in each case. In case (i), the dip moves to lower photon energies as the oxide thickness increases, with no dramatic change in its depth. In case (ii), the dip again moves to lower frequencies, but increases very substantially in depth, i.e. the roughness induced coupling of the incident photon to the surface plasmon is increased markedly by the presence of the overlayer. In case (iii), there is also considerable enhancement of the roughness-induced coupling to the surface plasmon, although the enhancement is smaller than for case (ii). Finally, for case (iv), once the oxide layer becomes sufficiently thick (say greater than 50Å), the coupling between the incident radiation and the surface plasmon is greatly decreased.

The above remarks show that in the presence of an oxide film (or a dielectric overlayer), the strength of the roughness-induced coupling to the surface plasmon depends very sensitively not only on the amplitude of the roughness, but also on the manner in which the roughness on the oxide-vacuum interface is correlated with that on the oxide-substrate interface. This is a principal conclusion of the present paper.

The remainder of the paper is organized as follows. In Section II, we sketch the derivation of expressions for the roughness induced scattering of normally incident light, along with the roughness induced absorption within the film on the substrate. The approach is similar to that employed by us earlier[5], and although the final

formulae are rather cumbersome for the present case, the presentation here is brief. We then present the results of the numerical calculations in Section III.

In an Appendix, we describe certain Green's functions of the electromagnetic field equations, for the present geometry. These Green's functions may be employed in a variety of problems. For example, the limiting form of these Green's functions with retardation ignored have formed the basis of a theory of the inelastic scattering of low energy electrons by electronic excitations in semiconductors.[6]

II. Derivation of the Theoretical Formulas

The geometry which forms the basis of the present paper is illustrated in Figure (2). In the absence of roughness on the two interfaces, the vacuum-overlayer interface is the plane z = d, and the overlayer substrate interface the plane z = o. In the presence of roughness, the function $\zeta_1(x,y)$ measures the elevation of point (x,y) on the vacuum-overlayer interface relative to the plane z = d. Similarly, $\zeta_2(x,y)$ describes the elevation of a point on the overlayer-substrate interface relative to the plane z = o. The overlayer material is presumed to be described by the isotropic, complex, frequency dependent dielectric constant ϵ_1, while the substrate is described by the frequency dependent dielectric constant ϵ_2, again complex and presumed isotropic. To study the reflectivity of the structure, we look for solutions of Maxwell's equations which vary harmonically with time:

$$\vec{E}(\vec{x}t) = \vec{E}(\vec{x},\omega)e^{-i\omega t} \quad , \tag{II-1}$$

where the electric field amplitude $\vec{E}(\vec{x},\omega)$ obeys

$$\vec{\nabla} \times \vec{\nabla} \times \vec{E}(\vec{x},\omega) - \frac{\omega^2}{c^2} \epsilon(\vec{x},\omega)\ \vec{E}(\vec{x},\omega) = o. \tag{II-2}$$

For the geometry of Figure (2), for the spatially varying dielectric constant we have

$$\begin{aligned} \epsilon(\vec{x},\omega) = {} & \theta(z-d-\zeta_1(x,y)) \\ & + \epsilon_1\theta(d + \zeta_1(x,y) - z)\ \theta\ (z-\zeta_2(x,y)) \\ & + \epsilon_2\theta(\zeta_2(x,y) - z) \quad , \end{aligned} \tag{II-3}$$

where in Eq. (II-3), $\theta(x)$ is the Heaviside step function which assumes the value unity when its argument is positive, and vanishes when its argument is negative.

When both $\zeta_1(x,y)$ and $\zeta_2(x,y)$ are small, we expand the right hand side of Eq. (II-3) in a Taylor series by means of the well known expansion

$$\theta(x + a) = \theta(x) + a\ \delta(x) + \ldots \quad , \tag{II-3}$$

where $\delta(x)$ is the Dirac delta function. Then Eq. (II-3) reads

$$\epsilon(\vec{x},\omega) = \epsilon_o(z,\omega) + \Delta\,\epsilon(\vec{x},\omega) \tag{II-5}$$

where

$$\epsilon_o(z,\omega) = \theta(z-d) + \epsilon_1\theta(d-z)\theta(z) + \epsilon_2\theta(-z) \tag{II-6}$$

and

$$\Delta\,\epsilon(x,\omega) = \zeta_1(x,y)(\epsilon_1-1)\delta(z-d) + (\epsilon_2-\epsilon_1)\zeta_2(x,y)\delta(z) \ . \tag{II-7}$$

Then Eq. (II-2) may be arranged to read

$$\vec{\nabla} \times \vec{\nabla} \times \vec{E}(\vec{x},\omega) - \frac{\omega^2}{c^2}\,\epsilon_o(z,\omega)\,\vec{E}(\vec{x},\omega) = \frac{\omega^2}{c^2}\,\Delta\,\epsilon(\vec{x},\omega)\,\vec{E}(\vec{x},\omega) \ . \tag{II-8}$$

To solve Eq. (II-8) in the limit $\zeta_1(x,y)$ and $\zeta_2(x,y)$ are small, we follow the approach used in our preceding paper.[5] We introduce a set of Green's functions $D_{\mu\nu}(\vec{x}\ \vec{x}',\omega)$ which satisfy

$$\sum_\mu \left\{\frac{\omega^2}{c^2}\,\epsilon_o(z,\omega)\delta_{\lambda\mu} - \frac{\partial^2}{\partial x_\lambda \partial x_\mu} + \delta_{\lambda\mu}\,\nabla^2\right\} D_{\mu\nu}(\vec{x}\ \vec{x}';\omega) = 4\pi\,\delta_{\lambda\mu}\,\delta(\vec{x} - \vec{x}') \quad , \tag{II-9}$$

along with boundary conditions appropriate to the present scattering problem.

In terms of these Green's functions, we may rewrite Eq. (II-8) in integral form

$$E_\mu(\vec{x},\omega) = E_\mu^{(o)}(\vec{x},\omega) - \frac{\omega^2}{4\pi c^2}\sum_\nu \int d^3x'\, D_{\mu\nu}(\vec{x},\vec{x}';\omega)\Delta\epsilon(\vec{x}',\omega)\, E_\nu(\vec{x}',\omega) \tag{II-10}$$

In Eq. (II-10), $E_\nu^{(o)}(\vec{x},\omega)$ is a solution of Eq. (II-8) with $\Delta\epsilon(\vec{x},\omega) = o$. The formal structure of Eq. (II-10) is identical to Schrödinger's equation of quantum mechanics, when it is written in integral form.[7] For small $\Delta\epsilon(\vec{x},\omega)$, we may generate an approximation analogous to the first Born approximation of quantum mechanics by iterating Eq. (II-10), and approximating the amplitude of the scattered wave $E_\mu^{(s)}(\vec{x},\omega)$ by retaining the first term. This gives

$$E_\mu^{(s)}(\vec{x},\omega) = \frac{\omega^2}{4\pi c^2}\sum_\nu \int d^3x'\, D_{\mu\nu}(\vec{x},\vec{x}';\omega)\Delta\epsilon(\vec{x}',\omega)E_\mu^{(o)}(\vec{x}',\omega). \tag{II-11}$$

The electric field amplitude $E_\mu^{(o)}(\vec{x},\omega)$ which appears in the right hand side of Eq. (II-11) is the electric field associated with the incident field, in the absence of surface roughness. The Green's functions $D_{\mu\nu}(\vec{x}\,\vec{x}',\omega)$ are constructed in the Appendix of the present paper. Thus, it is a straightforward, (but algebraically complex) matter to evaluate the scattered fields in the vacuum, within the overlayer, or within the substrate. We call the reader's attention to the rather extensive discussions in Reference (5), which explore a number of issues we do not examine here.

As before, since the dielectric function $\epsilon_o(z,\omega)$ which appears on the left hand side of Eq. (II-9) depends on z only, and not on

x and y, one may represent the Green's function by the partial Fourier decomposition

$$D_{\mu\nu}(\vec{x}\,\vec{x}\,',\omega) = \int \frac{d^2k_\parallel}{(2\pi)^2} e^{i\,\vec{k}_\parallel\cdot(\vec{x}_\parallel-\vec{x}_\parallel')} d_{\mu\nu}(\vec{k}_\parallel\omega|zz'). \quad \text{(II-12)}$$

We also write (where i = 1 or 2)

$$\zeta_i(x,y) = \int \frac{d^2k_\parallel}{(2\pi)^2} e^{i\,\vec{k}_\parallel\cdot\vec{x}_\parallel}\,\hat{\zeta}_i(\vec{k}_\parallel) \quad . \quad \text{(II-13).}$$

We presume here that the incident electric field is normally incident on the structure, with electric field parallel to the $\hat{x}$ axis. Then we have

$$E_\nu^{(o)}(\vec{x}'\omega) = \delta_{\nu x}\, E^{(o)}(\omega,z') \quad . \quad \text{(II-14)}$$

After these forms are substituted into Eq. (II-11), the scattered field assumes the form[8]

$$E_\mu^{(s)}(\vec{x},\omega) = -\frac{\omega^2(\epsilon_1-1)}{2(2\pi)^3c^2} E^{(o)}(\omega,d)\int d^2k_\parallel e^{i\,\vec{k}_\parallel\cdot\vec{x}_\parallel}\,\hat{\zeta}_1(\vec{k}_\parallel)$$

$$\times\, d_{\mu x}(\vec{k}_\parallel\omega|zd) \quad \text{(II-15)}$$

$$-\frac{\omega^2(\epsilon_2-\epsilon_1)}{2(2\pi)^3c^2} E^{(o)}(\omega,o)\int d^2k_\parallel\, e^{i\,\vec{k}_\parallel\cdot\vec{x}_\parallel}\,\hat{\zeta}_2(\vec{k}_\parallel)\;\; d_{\mu x}(\vec{k}_\parallel\omega|zo) \quad .$$

From the discussion in the Appendix, the functions $d_{\mu\nu}(\vec{k}_\parallel\omega|zz')$ are related to a second set of functions $g_{\mu\nu}(k_\parallel\omega|zz')$ via the transformation

$$d_{\mu\nu}(\vec{k}_\parallel\omega|zz') = \sum_{\mu'\nu'} S_{\mu'\nu}(\vec{k}_\parallel)S_{\nu'\nu}(\vec{k}_\parallel)g_{\mu'\nu'}(k_\parallel\omega|zz') \quad , \quad \text{(II-16)}$$

where

$$\vec{K}_{\parallel} = k_x \hat{x} + k_y \hat{y} \tag{II-17}$$

and the matrix $S(\vec{K}_{\parallel})$ is given by

$$\underset{\sim}{S}(\vec{K}_{\parallel}) = \frac{1}{k_{\parallel}} \begin{pmatrix} k_x & k_y & o \\ -k_y & k_x & o \\ o & o & k_{\parallel} \end{pmatrix}. \tag{II-18}$$

One then has the relations

$$d_{xx}(\vec{k}_{\parallel}\omega|zz') = \frac{k_x^2}{k_{\parallel}^2} g_{xx}(k_{\parallel}\omega|zz') + \frac{k_y^2}{k_{\parallel}^2} g_{yy}(k_{\parallel}\omega|zz') \tag{II-19a}$$

$$d_{yx}(\vec{k}_{\parallel}\omega|zz') = \frac{k_x k_y}{k_{\parallel}^2} \left[g_{xx}(k_{\parallel}\omega|zz') - g_{yy}(k_{\parallel}\omega|zz') \right] \tag{II-19b}$$

$$d_{zx}(\vec{k}_{\parallel}\omega|zz') = \frac{k_x}{k_{\parallel}} g_{zx}(k_{\parallel}\omega|zz'). \tag{II-19c}$$

To proceed, we now need to evaluate the scattered field in the three distinct regions of interest: in the vacuum above the overlayer, inside the overlayer, and in the substrate. We consider each regime separately.

(a) The scattered fields in the vacuum above the overlayer, and the angular distribution of the scattered radiation.

In this regime, we consider the limit $z \to +\infty$, for fixed z'. Then the Green's functions g_{xx}, g_{yy} and g_{zx} in Eqs. (II-19) have the form

$$g_{yy}(k_{\parallel}\omega|zz') = \frac{4\pi}{W_{\perp}(k_{\parallel},\omega)} E_y^{>}(k_{\parallel}\omega|z) E_y^{<}(k_{\parallel}\omega|z') \tag{II-20a}$$

$$g_{xx}(k_{\parallel}\omega|zz') = \frac{4\pi}{W_{\parallel}(k_{\parallel},\omega)} E_x^{>}(k_{\parallel}\omega|z) E_x^{<}(k_{\parallel}\omega|z') \tag{II-20b}$$

$$g_{zx}(k_{\parallel}\omega|zz') = \frac{4\pi}{W_{\parallel}(k_{\parallel},\omega)} E_z^{>}(k_{\parallel}\omega|z) E_x^{<}(k_{\parallel}\omega|z') \tag{II-20c}$$

where for $z > d$, one has

$$E_y^{>}(k_\parallel \omega | z) = E_z^{>}(k_\parallel \omega | z) = e^{ik_o z} \tag{II-21a}$$

and

$$E_x^{<}(k_\parallel \omega | z) = -\frac{k_o}{k_\parallel} e^{ik_o z} \tag{II-21b}$$

with

$$k_o = \left(\frac{\omega^2}{c^2} - k_\parallel^2\right)^{\frac{1}{2}} \quad . \tag{II-22}$$

As discussed in the Appendix, the positive square root is to be chosen in Eq. (II-22), and if $k_\parallel > \omega/c$, we choose

$$\mathrm{Im}(k_o) > o \quad .$$

The remaining quantities in Eqs. (II-20) and Eqs. (II-21) are defined in the Appendix.

The scattered electric field has the form

$$E_\mu^{(s)}(\vec{x}, \omega) = \int d^2 k_\parallel \; \varepsilon_\mu(\vec{k}_\parallel, \omega) e^{i\vec{k}\cdot\vec{x}} \quad , \tag{II-23}$$

where in Eq. (II-23),

$$\vec{k} = \vec{k}_\parallel + \hat{z}\, k_o \quad . \tag{II-24}$$

It is a short exercise to show that the time average of the Poynting vector, $\langle \vec{S} \rangle$, may be cast into the form

$$\langle \vec{S} \rangle = \frac{c^2}{8\pi\omega} \mathrm{Re} \int d^2 k_\parallel d^2 k_\parallel' \; e^{i(\vec{k} - \vec{k}'^{*})\cdot\vec{x}}$$

$$\times \left[\vec{k} \left\{ \vec{\varepsilon}^{*}(\vec{k}_\parallel' \omega) \cdot \vec{\varepsilon}(\vec{k}_\parallel \omega) \right\} \right. \tag{II-25}$$

$$\left. - \vec{\varepsilon}(\vec{k}_\parallel \omega) \left\{ \vec{k} \cdot \vec{\varepsilon}^{*}(\vec{k}_\parallel' \omega) \right\} \right]$$

We are interested here in the energy radiated into the vacuum. Thus, we confine our attention to the contributions to the integral from the regions $k_{\parallel} < \omega/c, k'_{\parallel} < \omega/c$. As explained earlier,[5] the regions with $k_{\parallel} > \omega/c$ describe scattered energy which is confined to the near vicinity of the surface, and which propagates parallel to it (i.e. stored in surface plasmons excited by the incident radiation, for example). The regions $k_{\parallel} > \omega/c$, $k'_{\parallel} > \omega/c$ give contributions to the energy flux which are small,[5] unless the surface plasmon in mean free path is comparable to the linear dimensions of the region illuminated by the incident beam.

We may calculate the Poynting vector by inserting the amplitudes of the scattered fields into $\langle \vec{S} \rangle$, and then averaging over the distribution of surface roughness, as we did before. The calculation proceeds along very similar lines to our earlier work.

We comment on one point, however. When one averages over the distribution of surface roughness, one encounters averages of the form $\langle \hat{\zeta}_i(\vec{k}_{\parallel})^* \, \zeta_j(\vec{k}'_{\parallel}) \rangle$. The two functions $\langle \hat{\zeta}_1(\vec{k}_{\parallel})^* \, \hat{\zeta}_1(\vec{k}'_{\parallel}) \rangle$ and $\langle \hat{\zeta}_2(\vec{k}_{\parallel})^* \, \hat{\zeta}_2(\vec{k}'_{\parallel}) \rangle$ describe the nature of the roughness on the vacuum-overlayer and overlayer-substrate interfaces, respectively. In general, the "off diagonal" averages $\langle \hat{\zeta}_1(\vec{k}_{\parallel})^* \, \hat{\zeta}_2(\vec{k}'_{\parallel}) \rangle$ and $\langle \hat{\zeta}_2(\vec{k}_{\parallel})^* \, \hat{\zeta}_1(\vec{k}'_{\parallel}) \rangle$ will be also non-zero. These functions contain information about the manner in which the roughness on the vacuum-overlayer interface is <u>correlated</u> with that on the overlayer-substrate interface. These functions will vanish only if the roughness on the outermost interface is distributed randomly relative to that on the innermost interface, a possibility that seems unlikely for a thin overlayer.

By a straightforward generalization of our earlier definitions, we write

$$\langle \hat{\zeta}_i^*(\vec{k}_\parallel)\zeta_j(\vec{k}_\parallel')\rangle = (2\pi)^2\delta(\vec{k}_\parallel-\vec{k}_\parallel')\delta_i\delta_j\, g_{ij}(\vec{k}_\parallel) \quad , \tag{II-26}$$

where

$$g_{ij}(\vec{k}_\parallel) = \frac{1}{\delta_i\delta_j}\int d^2r_\parallel e^{-i\vec{k}_\parallel\cdot\vec{r}_\parallel}\,\langle\zeta_i(o)\zeta_j(\vec{r}_\parallel)\rangle \ . \tag{II-27}$$

In Eq. (II-26) and Eq. (II-27), the quantities δ_1 and δ_2 are the root mean square roughness amplitudes for the vacuum-overlayer interface, and the overlayer-substrate interface, respectively, i.e.

$$\delta_1 = \langle\zeta_1^2\rangle^{\frac{1}{2}} \tag{II-28a}$$

and

$$\delta_2 = \langle\zeta_2^2\rangle^{\frac{1}{2}} \tag{II-28b}$$

It follows from this definition that

$$\int\frac{d^2k_\parallel}{4\pi^2}\,g_\parallel(\vec{k}_\parallel) = \int\frac{d^2k_\parallel}{4\pi^2}\,g_{22}(\vec{k}_\parallel) = 1, \tag{II-29}$$

while no simple normalization requirement exists for $g_{12}(\vec{k}_\parallel)$ or $g_{21}(\vec{k}_\parallel)$, although necessarily $g_{12}(\vec{k}_\parallel) = g_{21}(\vec{k}_\parallel)^*$.

The overlayer roughness configurations illustrated in Fig. (1) can be seen to correspond to particular choices of $g_{12}(\vec{k}_\parallel)$. For example, the replicating film model of Figure (1a) corresponds to the choice $g_\parallel(\vec{k}_\parallel) = g_{22}(\vec{k}_\parallel) = g_{12}(\vec{k}_\parallel)$, and the non-uniform film model of Figure (1b) to the choice $g_\parallel(\vec{k}_\parallel) = g_{22}(\vec{k}_\parallel) = -\,g_{12}(\vec{k}_\parallel)$.

With the above remarks and our preceding discussion in hand, one may construct expressions for the angular distribution of the scattered energy flux. We shall simply quote the results here, since the algebrai manipulations are lengthy and offer no enlightenment.

We let $(df_s/d\Omega)d\Omega$ be the fraction of the incident radiation (recall we consider only normal incidence here) scattered into final states with s polarization, directed toward the solid angle $d\Omega$. In a similar fashion, $(df_p/d\Omega)$ describes the angular distribution of radiation of p polarization. The direction of the outgoing radiation is described by the spherical angles θ_s and φ_s, where φ_s is measured from the x axis.

Before we write down the final expressions, we define the following quantities:

$$\kappa_1 = (\epsilon_1 - \sin^2\theta_s)^{\frac{1}{2}} \quad , \quad \mathrm{Im}(\kappa_1) > 0 \tag{II-30a}$$

$$\kappa_2 = (\epsilon_2 - \sin^2\theta_s)^{\frac{1}{2}} \quad , \quad \mathrm{Im}(\kappa_2) > 0 \tag{II-30b}$$

$$d_s(\theta_s,\omega) = (\kappa_2 + \cos\theta_s)\cos\left(\frac{\omega}{c}\kappa_1 d\right) - \frac{i}{\kappa_1}(\kappa_1^2 + \kappa_2\cos\theta_s)\sin\left(\frac{\omega}{c}\kappa_1 d\right) \tag{II-30c}$$

$$e_s(\theta_s,\omega) = \cos\left(\frac{\omega}{c}\kappa_1 d\right) - i\frac{\kappa_2}{\kappa_1}\sin\left(\frac{\omega}{c}\kappa_1 d\right) \tag{II-30d}$$

$$d_p(\theta_s,\omega) = -(\epsilon_2\cos\theta_s + \kappa_2)\cos\left(\frac{\omega}{c}\kappa_1 d\right) - i\left(\cos\theta_s\,\epsilon_1\frac{\kappa_2}{\kappa_1} + \frac{\epsilon_2}{\epsilon_1}\kappa_1\right)\sin\left(d\frac{\omega}{c}\kappa_1\right) \tag{II-30e}$$

$$e_p(\theta_s,\omega) = \cos\left(\frac{\omega}{c}\kappa_1 d\right) - i\frac{\epsilon_2}{\epsilon_1}\frac{\kappa_1}{\kappa_2}\sin\left(\frac{\kappa}{c}\kappa_1 d\right) \quad . \tag{II-30f}$$

We then have[9]

$$\frac{df_s(\theta,\varphi_s)}{d\Omega} = \frac{\omega^4}{\pi^2 c^4} \frac{\cos^2\theta_s \sin^2\varphi_s}{|d_s(\theta_s,\omega)|^2 |d_s(o,\omega)|^2}$$

$$\times \left[\delta_1^2 |\epsilon_1 - 1|^2 |\varepsilon_s(\theta_s,\omega)|^2 |\varepsilon_s(o,\omega)|^2 g_{11}(\vec{k}_\parallel)\right. \qquad \text{(II-31)}$$

$$+ \delta_2^2 |\epsilon_2 - \epsilon_1|^2 g_{22}(\vec{k}_\parallel)$$

$$\left. + 2\delta_1\delta_2 \text{ Re } \left\{(\epsilon_1^* - 1)(\epsilon_2 - \epsilon_1)\varepsilon_s^*(o,\omega)\varepsilon_s(\theta_s,\omega) g_{12}(\vec{k}_\parallel)\right\}\right]$$

and

$$\frac{df_p(\theta_s,\varphi_s)}{d\Omega_p} = \frac{\omega^4}{\pi^2 c^4} \frac{\cos^2\theta_s \cos^2\varphi_s |\kappa_2|^2}{|d_p(\theta_s,\omega)|^2 |d_s(o,\omega)|^2}$$

$$\times \left[\delta_1^2 |\epsilon_1 - 1|^2 |\varepsilon_p(\theta_s,\omega)|^2 |\varepsilon_s(o,\omega)|^2 g_{11}(\vec{k}_\parallel)\right. \qquad \text{(II-32)}$$

$$+ \delta_2^2 |\epsilon_2 - \epsilon_1|^2 g_{22}(\vec{k}_\parallel)$$

$$\left. + 2\delta_1\delta_2 \text{ Re } \left\{(\epsilon_1^* - 1)(\epsilon_2 - \epsilon_1)\varepsilon_s^*(o,\omega)\varepsilon_p(\theta_s,\omega) g_{12}(\vec{k}_\parallel)\right\}\right] \quad .$$

In Eq. (II-31) and Eq. (II-32), $\vec{k}_\parallel$ is the projection of the wave vector of the scattered wave vector on a plane parallel to the surface. Thus, the magnitude of $\vec{k}_\parallel$ is given by

$$|\vec{k}_\parallel| = \frac{\omega}{c} \sin\theta_s \quad . \qquad \text{(II-33)}$$

(b) The scattered electric fields within the overlayer, and the fraction of the incident energy absorbed within the overlayer.

To compute the scattered electric field within the film, we may use Eq. (II-15), with the Green's functions $d_{\mu x}(\vec{k}_\parallel \omega | zd)$ and $d_{\mu x}(\vec{k}_\parallel \omega | zo)$ given by Eq. (II-19).

If we define (as in the Appendix)

$$k_1 = \left(\frac{\omega^2}{c^2}\epsilon_1 - k_\parallel^2\right)^{\frac{1}{2}} \quad , \quad \text{Im}\,(k_1) < o \quad , \tag{II-34}$$

then the electric field within the film has the form

$$E_\mu^{(s)}(\vec{x},\omega) = \int d^2k_\parallel e^{i\vec{k}_\parallel \cdot \vec{x}_\parallel}\left\{\mathcal{E}_\mu^{(+)}(\vec{k}_\parallel,\omega)e^{+ik_1 z} + \mathcal{E}_\mu^{(-)}(\vec{k}_\parallel,\omega)e^{-ik_1 z}\right\}. \tag{II-35}$$

where after some algebra, one finds

$$\begin{aligned}
\mathcal{E}_x^{(\sigma)}(\vec{k}_\parallel,\omega) = &- \frac{\omega^2(\epsilon_1 - 1)}{(2\pi)^2 c^2} E^{(o)}(\omega,d)e^{ik_o d}\hat{\zeta}_1(\vec{k}_\parallel) \\
&\times\left\{\sigma\frac{k_o k_1 k_x^2}{W_\parallel(k_\parallel,\omega)k_\parallel^4}C_\sigma^{(\parallel)} + \frac{k_y^2}{k_\parallel^2 W_\perp(k_\parallel,\omega)}C_\sigma^{(\perp)}\right\} \\
&- \frac{\omega^2(\epsilon_2-\epsilon_1)}{(2\pi)^2 c^2} E^{(o)}(\omega,o)\hat{\zeta}_2(\vec{k}_\parallel)\left\{\sigma\frac{k_1 k_2 k_x^2}{k_\parallel^2 W_\parallel(k_\parallel,\omega)}A_\sigma^{(\parallel)}\right. \\
&\left. + \frac{k_y^2}{k_\parallel^2 W_\perp(k_\parallel,\omega)}A_\sigma^{(\perp)}\right\} \quad ,
\end{aligned} \tag{II-36}$$

$$\begin{aligned}
\mathcal{E}_y^{(\sigma)}(\vec{k}_\parallel,\omega) = &- \frac{\omega^2(\epsilon_1-1)}{(2\pi)^2 c^2} E^{(o)}(\omega,d)e^{ik_o d}\hat{\zeta}_1(\vec{k}_\parallel)\frac{k_x k_y}{k_\parallel^2} \\
&\times\left\{\sigma\frac{k_o k_1}{k_\parallel^2 W_\parallel(k_\parallel,\omega)}C_\sigma^{(\parallel)} - \frac{1}{W_\perp(k_\parallel,\omega)}C_\sigma^{(\perp)}\right\}
\end{aligned}$$

(II-37)

$$- \frac{\omega^2(\epsilon_2-\epsilon_1)}{(2\pi)^2 c^2} E^{(o)}(\omega,o)\hat{\zeta}_2(\vec{k}_\parallel) \frac{k_x k_y}{k_\parallel^2}$$

$$\times \left\{ \sigma \frac{k_1 k_2}{k_\parallel^2 W_\parallel(k_\parallel,\omega)} A_\sigma^{(\parallel)} - \frac{1}{W_\perp(k_\parallel,\omega)} A_\sigma^{(\perp)} \right\}$$

$$\mathcal{E}_z^{(\sigma)}(\vec{k}_\parallel,\omega) = + \frac{\omega^2(\epsilon_1-1)}{(2\pi)^2 c^2} E^{(o)}(\omega,d) e^{ik_o d} \hat{\zeta}_1(\vec{k}_\parallel) \frac{k_o k_x}{k_\parallel^2 W_\parallel(k_\parallel,\omega)} C_\sigma^{(\parallel)}$$

(II-38)

$$+ \frac{\omega^2(\epsilon_2-\epsilon_1)}{(2\pi)^2 c^2} E^{(o)}(\omega,o) \hat{\zeta}_2(\vec{k}_\parallel) \frac{k_x k_2}{k_\parallel^2 W_\parallel(k_\parallel,\omega)} A_\sigma^{(\parallel)}$$

In Eq. (II-36), Eq. (II-37) and Eq. (II-38), we have as in the Appendix,

$$k_2 = \left(\frac{\omega^2}{c^2}\epsilon_2 - k_\parallel^2\right)^{\frac{1}{2}} \quad \mathrm{Im}(k_2) < o \tag{II-39}$$

$$k_o = \left(\frac{\omega^2}{c^2} - k_\parallel^2\right)^{\frac{1}{2}} \quad \mathrm{Im}(k_o) \geq o \quad . \tag{II-40}$$

The coefficients $C_\pm^{(\parallel)}$, $C_\pm^{(\perp)}$, etc. are given by

$$C_\sigma^{(\perp)} = \tfrac{1}{2}\left[1 + \sigma \frac{k_2}{k_1}\right] \tag{II-41a}$$

$$C_\sigma^{(\parallel)} = \tfrac{1}{2}\left[\frac{\epsilon_2}{\epsilon_1} + \sigma \frac{k_2}{k_1}\right] \tag{II-41b}$$

$$A_\sigma^{(\perp)} = \tfrac{1}{2}\left[1 + \sigma \frac{k_o}{k_1}\right] e^{ik_o d} e^{-i\sigma k_1 d} \tag{II-41c}$$

$$A_\sigma^{(\parallel)} = \tfrac{1}{2}\left[\frac{1}{\epsilon_1} + \sigma \frac{k_o}{k_1}\right] e^{ik_o d} e^{-i\sigma k_1 d} \quad . \tag{II-41d}$$

It is now a straightforward, but tedious matter to evaluate the fraction of the incident radiation absorbed by the overlayer. By symmetry, the only component of the Poynting vector which has a non-vanishing value is the z component, $\langle S_z \rangle$. The rate at which energy is dissipated in the film is then $L_x L_y \left\{ \langle S_z \rangle \Big|_{z=o} - \langle S_z \rangle \Big|_{z=d} \right\}$, where $L_x L_y$ is the area illuminated by the incident beam. We calculate this quantity, and divide by rate at which incident energy strikes the surface to form an expression for the fraction $f^{(1)}$ of the incident energy absorbed by the overlayer. The quantity $f^{(1)}$ has the form

$$f^{(1)} = \sum_{\sigma=\pm 1} \sum_{\sigma'=\pm 1} f^{(1)}_{\sigma\sigma'} , \qquad \text{(II-42)}$$

where

$$f^{(1)}_{\sigma\sigma'} = \frac{\omega}{4\pi^2 c} \frac{1}{|d_s(o,\omega)|^2} \mathrm{Re}\left[\int d^2k_\parallel \left(1 - e^{id[\sigma k_1 - \sigma' k_1^*]}\right) \right.$$

$$\times \left\{ \delta_1^2 |\epsilon_1 - 1|^2 |\varepsilon_s(o,\omega)|^2 g_\parallel(\vec{k}_\parallel) \right.$$

$$\times \left(\sigma' \cos^2\varphi_s \frac{c^2}{\omega^2} \frac{|k_o|^2 \left(k_1 |k_1|^2 + k_\parallel^2 k_1^*\right)}{|d_p(k_\parallel,\omega)|^2} \left[\frac{\epsilon_2^*}{\epsilon_1^*} + \sigma' \frac{k_2^*}{k_1^*} \right] \left[\frac{\epsilon_2}{\epsilon_1} + \sigma \frac{k_o}{k_1} \right] \right.$$

$$\left. \left. + \sigma \sin^2\varphi_s \frac{\omega^2}{c^2} \frac{k_1}{|d_s(k_\parallel,\omega)|^2} \left[1 + \sigma' \frac{k_2^*}{k_1^*} \right] \left[1 + \sigma \frac{k_2}{k_1} \right] \right) \right.$$

$$\left. + \delta_2^2 |\epsilon_2 - \epsilon_1|^2 g_{22}(\vec{k}_\parallel) e^{-i[\sigma k_1 - \sigma k_1^*]d} \right.$$

$$\times \left(\sigma' \cos^2\varphi_s \frac{c^2}{\omega^2} \frac{|k_2|^2 \left(k_1 |k_1|^2 + k_\parallel^2 k_1^*\right)}{|d_p(k_\parallel,\omega)|^2} \left[\frac{1}{\epsilon_1^*} + \sigma' \frac{k_o^*}{k_1^*} \right] \left[\frac{1}{\epsilon_1} + \sigma \frac{k_o}{k_1} \right] \right.$$

$$+ \sigma \sin^2\varphi_s \frac{\omega^2}{c^2} \frac{k_1}{|d_s(k_\parallel,\omega)|^2} \left[1 + \sigma' \frac{k_o^*}{k_1^*}\right]\left[1 + \sigma \frac{k_o}{k_1}\right]\Bigg)$$

$$+ \delta_1\delta_2(\epsilon_2-\epsilon_1)(\epsilon_1^*-1)\, \mathcal{E}_s^*(o,\omega)\, e^{-i\sigma k_1 d}\, g_{12}(k_\parallel)$$

$$\times\Bigg(\sigma' \cos^2\varphi_s \frac{c^2}{\omega^2} \frac{k_o^*(k_1|k_1|^2 + k_\parallel^2 k_1^*)k_2}{|d_p(k_\parallel,\omega)|^2}\left[\frac{\epsilon_2^*}{\epsilon_1^*} + \sigma' \frac{k_2^*}{k_1^*}\right]\left[\frac{1}{\epsilon_1} + \sigma \frac{k_o}{k_1}\right]$$

$$+ \sigma \sin^2\varphi_s \frac{\omega^2}{c^2} \frac{k_1}{|d_s(k_\parallel,\omega)|^2}\left[1 + \sigma' \frac{k_2^*}{k_1^*}\right]\left[1 + \sigma \frac{k_o}{k_1}\right]\Bigg)$$

$$+ \delta_1\delta_2(\epsilon_2^*-\epsilon_1^*)(\epsilon_1-1)\, \mathcal{E}_s(o,\omega)\, e^{+i\sigma' k_1^* d}\, g_{12}^*(\vec{k}_\parallel)$$

$$\times\Bigg(\sigma' \cos^2\varphi_s \frac{c^2}{\omega^2} \frac{k_o(k_1|k_1|^2 + k_\parallel^2 k_1^*)k_2^*}{|d_p(k_\parallel,\omega)|^2}\left[\frac{1}{\epsilon_1^*} + \sigma' \frac{k_o^*}{k_1^*}\right]\left[\frac{\epsilon_2}{\epsilon_1} + \sigma \frac{k_2}{k_1}\right]$$

$$+ \sigma \sin^2\varphi_s \frac{\omega^2}{c^2} \frac{k_1}{|d_s(k_\parallel,\omega)|^2}\left[1 + \sigma' \frac{k_o^*}{k_1^*}\right]\left[1 + \sigma \frac{k_2}{k_1}\right)\Bigg)\Bigg]$$

In Eq. (II-42), we have

$$d_s(k_\parallel,\omega) = (k_o - k_2)\cos(k_1 d) - i\left(k_1 - \frac{k_o k_2}{k_1}\right)\sin(k_1 d), \quad \text{(II-43)}$$

and

$$d_p(k_\parallel,\omega) = (\epsilon_2 k_o - k_2)\cos(k_1 d) - i\left(\frac{\epsilon_2 k_1}{\epsilon_1} - \epsilon_1 \frac{k_2 k_o}{k_1}\right)\sin(k_1 d). \quad \text{(II-44)}$$

The definitions of the remaining quantities may be found earlier in the present section.

(c) The scattered electric fields within the substrate, and the fraction of the incident energy absorbed within the substrate.

We evaluate the scattered electric field within the substrate through the use once again of Eq. (II-15) in concert with Eqs. (II-19). Now we require the Green's functions for $z < o$, $z < z'$. In this regime we have

$$g_{xx}(k_\parallel \omega | z z') = \frac{4\pi}{W_\parallel (k_\parallel, \omega)} E_x^< (k_\parallel \omega | z) E_x^> (k_\parallel \omega | z') \tag{II-45a}$$

$$g_{yy}(k_\parallel \omega | z z') = \frac{4\pi}{W_\perp (k_\parallel, \omega)} E_y^< (k_\parallel \omega | z) E_y^> (k_\parallel \omega | z') \tag{II-45b}$$

and

$$g_{zx}(k_\parallel \omega | z z') = \frac{4\pi}{W_\parallel (k_\parallel, \omega)} E_z^< (k_\parallel \omega | z) E_x^> (k_\parallel \omega | z') \quad , \tag{II-45c}$$

where

$$E_x^< (k_\parallel \omega | z) = - \frac{k_2}{k_\parallel} e^{ik_2 z} \tag{II-46a}$$

$$E_y^< (k_\parallel \omega | z) = e^{ik_2 z} \tag{II-46b}$$

and

$$E_z^< (k_\parallel \omega | z) = e^{ik_2 z} \quad . \tag{II-46c}$$

The scattered field in the substrate then assumes the form

$$E_\mu^{(s)} (\vec{x}, \omega) = \int d^2 k_\parallel e^{i\vec{k}_\parallel \cdot \vec{x}_\parallel} e^{ik_2 z} \varepsilon_\mu^{(2)} (\vec{k}, \omega) \tag{II-47}$$

where the explicit form of the quantities $\varepsilon_\mu^{(2)} (\vec{k}_\parallel, \omega)$ is

$$\varepsilon_x^{(2)} (\vec{k}_\parallel, \omega) = + \frac{\omega^2 (\epsilon_1 - 1)}{(2\pi)^2 c^2} E^{(o)} (\omega, d) \hat{\zeta}_1 (\vec{k}_\parallel) \tag{II-48a}$$

$$\times \left\{ \frac{k_x^2 k_2}{k_\parallel^3 W_\parallel (k_\parallel, \omega)} E_x^> (k_\parallel \omega | d) - \frac{k_y^2}{W_\perp (k_\parallel, \omega) k_\parallel^2} E_y^> (k_\parallel \omega | d) \right\}$$

$$+ \frac{\omega^2(\epsilon_2-\epsilon_1)}{(2\pi)^2 c^2} E^{(o)}(\omega,o)\hat{\zeta}_2(\vec{k}_\parallel) \left\{ \frac{k_x^2 k_2 E_x^>(k_\parallel\omega|o)}{k_\parallel^3 W_\parallel(k_\parallel,\omega)} - \frac{k_y^2 E_y^>(k_\parallel\omega|o)}{k_\parallel^2 W_\perp(k_\parallel,\omega)} \right\}$$

$$e_y^{(2)}(\vec{k}_\parallel\omega) = \frac{\omega^2(\epsilon_1-1)}{(2\pi)^2 c^2} E^{(o)}(\omega,d)\hat{\zeta}_1(k_\parallel) \frac{k_x k_y}{k_\parallel^2} \qquad \text{(II-48b)}$$

$$\times \left\{ \frac{k_2 E_x^>(k_\parallel\omega|d)}{k_\parallel W_\parallel(k_\parallel,\omega)} - \frac{E_y^>(k_\parallel\omega|d)}{W_\perp(k_\parallel,\omega)} \right\}$$

$$+ \frac{\omega^2(\epsilon_2-\epsilon_1)}{(2\pi)^2 c^2} E^{(o)}(\omega,o)\hat{\zeta}_2(k_\parallel) \frac{k_x k_y}{k_\parallel^2}$$

$$\times \left\{ \frac{k_2 E_x^>(k_\parallel\omega|o)}{k_\parallel W_\parallel(k_\parallel,\omega)} + \frac{E_y^>(k_\parallel\omega|o)}{W_\perp(k_\parallel,\omega)} \right\}$$

$$e_z^{(2)}(\vec{k}_\parallel,\omega) = -\frac{\omega^2(\epsilon_1-1)}{(2\pi)^2 c^2} E^{(o)}(\omega,d)\hat{\zeta}_1(\vec{k}_\parallel) \frac{k_x E_x^>(k_\parallel\omega|d)}{k_\parallel W_\parallel(k_\parallel,\omega)} \qquad \text{(II-48c)}$$

$$- \frac{\omega^2(\epsilon_2-\epsilon_1)}{(2\pi)^2 c^2} E^{(o)}(\omega,o)\hat{\zeta}_2(\vec{k}_\parallel) \frac{k_x E_x^>(k_\parallel\omega|o)}{k_\parallel W_\parallel(k_\parallel,\omega)} \quad .$$

In reference (5), it was argued that in the limit that the mean free path of the surface plasmon is short compared to the linear size of the region of the surface illuminated by the incident beam, then the dominant contribution to the energy absorption by the substrate comes from the energy flow in the direction normal to the surface. This rate is equal to $L_x L_y \langle S_z \rangle\big|_{z=o-}$, where again

$L_x L_y$ is the area of the surface illuminated by the incident beam. It is straightforward to compute this quantity, and divide it by the energy/unit time that strikes the surface to obtain the fraction $f^{(2)}$ of the incident energy absorbed within the substrate. When this is done, we find the following expression:

$$f^{(2)} = \frac{\omega}{\pi^2 c} \frac{1}{|d_s(o,\omega)|^2} \int d^2k_\parallel |\mathrm{Re}(k_2)|$$

$$\times \left\{ \cos^2\varphi_s \frac{c^2}{\omega^2} \frac{|k_o|^2(k_\parallel^2 + |k_2|^2)}{|d_p(k_\parallel,\omega)|^2} \left[\delta_1^2 |\epsilon_1 - 1|^2 |\mathcal{E}_s(o,\omega)|^2 g_{11}(\vec{k}_\parallel) \right.\right.$$

$$+ 2\delta_1\delta_2 \mathrm{Re}\left(\left[\epsilon_1^* - 1\right]\left[\epsilon_2 - \epsilon_1\right] \mathcal{E}_s^*(o,\omega)\epsilon_p(k_\parallel,\omega) g_{12}(k_\parallel) \right)$$

$$\left. + \delta_2^2 |\epsilon_2 - \epsilon_1|^2 |\epsilon_p(k_\parallel,\omega)|^2 g_{22}(\vec{k}_\parallel) \right] \qquad \text{(II-49)}$$

$$+ \sin^2\varphi_s \frac{\omega^2}{c^2} \frac{1}{|d_s(k_\parallel,\omega)|^2} \left[\delta_1^2 |\epsilon_1 - 1|^2 |\mathcal{E}_s(o,\omega)|^2 g_{11}(\vec{k}) \right.$$

$$+ 2\delta_1\delta_2 \mathrm{Re}\left(\left[\epsilon_1^* - 1\right]\left[\epsilon_2 - \epsilon_1\right] \mathcal{E}_s^*(o,\omega)\epsilon_s(k_\parallel,\omega) g_{12}(\vec{k}_\parallel) \right)$$

$$\left.\left. + \delta_2^2 |\epsilon_2 - \epsilon_1|^2 |\epsilon_s(k_\parallel,\omega)|^2 g_{22}(\vec{k}_\parallel) \right] \right\} .$$

In Eq. (II-49), we have introduced the quantities

$$\epsilon_s(k_\parallel \omega) = \cos(k,d) - i \frac{k_o}{k_1} \sin(k,d) \qquad \text{(II-50)}$$

$$\epsilon_p(k_\parallel \omega) = \cos(k,d) - i \frac{k_1}{\epsilon_1 k_o} \sin(k,d) \quad . \qquad \text{(II-51)}$$

III. Numerical Calculations

In this section, we present the results of a set of numerical calculations of the change in reflectivity of the structure, in the presence of surface roughness.

To carry out these calculations, we require values of the complex dielectric constant ϵ_1 of the overlayer, and the complex dielectric constant ϵ_2 of the substrate. We have chosen to carry out the calculations for aluminum metal overcoated with an oxide film. For the dielectric constant of the overlayer, we have employed the dielectric constant of Al_2O_3 films reported by Arakawa and Williams.[10] This data shows that the dielectric constant of Al_2O_3 is real below photon energies of ≈ 8 eV, and absorption sets in for photon energies higher than this value. For the dielectric constant ϵ_2 of the substrate, we have employed the values for aluminum reported by Ehrenreich, Philipp and Segall.[11]

We also require values for the correlation functions $g_{11}(\vec{k})$, $g_{22}(\vec{k}_\parallel)$ and $g_{12}(\vec{k})$. We shall restrict our attention to the four model situations depicted in Figure (1). In each case, a simple relation exists between the three correlation functions, so we only need specify one of them to proceed. The relations are as follows:

(i) The replicating film model (Figure 1(a)).

Here we have $\zeta_1(x,y) = \zeta_2(x,y)$ everywhere so that

$$g_{11}(\vec{k}_\parallel) = g_{22}(\vec{k}_\parallel) = g_{12}(\vec{k}_\parallel) \qquad \text{(III-1)}$$

and also

$$\delta_1 = \delta_2 \quad . \qquad \text{(III-2)}$$

(ii) The non-uniform film model (Figure 1(b)).

Here we have $\zeta_1(x,y) = -\zeta_1(x,y)$ everywhere. Then

$$g_{11}(\vec{k}_\parallel) = g_{22}(\vec{k}_\parallel) = -g_{12}(\vec{k}_\parallel) \qquad \text{(III-3)}$$

and again

$$\delta_1 = \delta_2 \quad . \qquad \text{(III-4)}$$

(iii) The random roughness model (Figure 1(c)).

In this model, we presume that the roughness on the vacuum-oxide interface is uncorrelated with that on the oxide-substrate interface. This means that

$$g_{12}(\vec{k}_\parallel) = o \quad , \qquad \text{(III-5)}$$

while $g_{11}(\vec{k}_\parallel)$ and $g_{22}(\vec{k}_\parallel)$ bear no simple relation to each other, in general. For simplicity, however, we shall choose

$$g_{11}(\vec{k}_\parallel) = g_{22}(\vec{k}_\parallel) \text{ and} \qquad \text{(III-6)}$$

$$\delta_1 = \delta_2 \qquad \text{(III-7)}$$

for this model, while the condition in Eq. (IV-5) holds also.

(iv) The rough oxide layer model. We presume the oxide-substrate interface is perfectly smooth, while the interface between the oxide film and the vacuum is rough. This means that $\zeta_2(x,y) = o$ everywhere, so we have the conditions

$$g_{22}(\vec{k}_\parallel) = g_{12}(\vec{k}_\parallel) = o \qquad \text{(III-8)}$$

$$g_{11}(\vec{k}_\parallel) \neq o. \qquad \text{(III-9)}$$

Each of the four models described above requires knowledge of one correlation function $g_{11}(\vec{k}_{\parallel})$, and the remaining correlation functions may be obtained from it. As in our earlier calculations, we choose a gaussian for $g_{11}(\vec{k}_{\parallel})$:

$$g_{11}(\vec{k}_{\parallel}) = \pi a^2 \exp\left[-\tfrac{1}{4} a^2 k_{\parallel}^2\right] \quad . \qquad \text{(III-10)}$$

The parameter a is the transverse correlation length. It is a measure of the average distance between neighboring peaks on the rough surface.

To begin, we calculate the change in reflectivity for a rough surface of pure aluminum, with no oxide overlayer present. While we presented similar calculations in our earlier work, in the present calculation we have chosen a value for the transverse correlation length which provides a rough fit to the data reported by Endriz and Spicer.[12] These authors have completed an extensive series of experimental studies of the effect of roughness on the reflectivity of aluminum in the ultra violet.

In their paper, Endriz and Spicer have also provided detailed fits to their data. However, in their fitting procedure, they employed theoretical expressions which have appeared in the literature,[13] but which are in error.[5],[14] In our present calculations, we have not attempted to obtain the kind of detailed quantitative fit to the data attempred by Endriz and Spicer. Our interest here is in a calculation which provides a reasonable qualitative fit.

We find that if we choose the transverse correlation length a = 200Å, we obtain results rather similar to the experimental data. In Figure (3), we present our results, for the case where the root

mean square height of the roughness (the parameter δ) is chosen to be 12Å. The dominant contribution to the roughness-induced change in reflectivity comes from roughness induced coupling to the surface plasmon. In aluminum, the surface plasmon energy is 10.6 eV, and one sees that the minimum in the dip in the reflectivity occurs near, but below this energy.

We would like to comment on one feature of our calculation, for pure aluminum. In the literature, it is frequently presumed[12],[15] that for frequencies above the surface plasmon energy, there is no roughness-induced absorption by the substrate, and as a consequence the roughness-induced change in reflectivity has its origin entirely in the scattering of the incident light away from the specular direction. As we pointed out earlier,[5] since the imaginary part of the dielectric constant of the substrate is non-zero, there is roughness-induced absorption present at all frequencies, even above the surface plasmon frequency. For the parameters chosen to describe pure aluminum, even at 12eV we find the dominant contribution to the roughness induced change in reflectivity comes not from roughness induced-scattering away from the specular direction, but rather from absorption in the substrate. In the calculations reported in the paper by Endriz and Spicer, the roughness induced scattering rate was found to be considerably larger than that we calculate here. These authors used a considerably larger value of the correlation length ($\approx$ 1000Å) than we have. We find that for larger values of the correlation length, our calculated scattering rate increases appreciably, but we can no longer obtain a reasonable fit to the reflectivity change produced by roughness at lower energies where the surface plasmon-induced dip occurs.

In Figure (4), we present our calculations of the roughness-induced change in reflectivity for the replicating film model described above. One sees that as the thickness of the oxide layer increases, the reflectivity dip shifts toward the visible. The reason for the shift is that the presence of the oxide layer modifies the dispersion relation of the surface plasmon.[16] In particular, for a metal with bulk plasma frequency ω_p, in the limit that the wave vector $k_\parallel \to \infty$, the surface plasmon frequency for a metallic substrate overcoated with a dielectric layer with dielectric constant ϵ approaches the value $\omega_p/(1+\epsilon)^{\frac{1}{2}}$ rather than the value $\omega_p/\sqrt{2}$ associated with the metal-vacuum interface. We would then expect that for large values of the overlayer thickness d, the reflectivity dip to shift downward in frequency to lie just below $\omega_p/(1+\epsilon)^{\frac{1}{2}}$. If we choose $\epsilon \cong 4$ as a typical value for Al_2O_3 in the frequency range of interest, then $\omega_p/(1+\epsilon)^{\frac{1}{2}} = 6.7$eV. Thus, by the time the thickness of the oxide layer reaches 100Å, the calculations show that the reflectivity dip lies near this asymptotic value.

Note that for the replicating film model, the magnitude of the dip is not affected by the presence of the overlayer in any dramatic manner.

In Figure (5), we present calculations of the roughness-induced change in reflectivity for the non-uniform film model described earlier, and illustrated in Figure 1(b). While the position of the minimum in the reflectivity for each value of the oxide thickness coincides quite closely with the minima displayed in Figure (4) for the replicating film model, the most striking feature of the results in Figure (5) is the very substantial enhancement of the

strength of the coupling between the light and the surface plasmon. Note that in the calculations illustrated in Figure (5), we have reduced the rms height of the roughness on each interface from the value 12Å used in Figure (3) and Figure (4) to the smaller value of 6Å. Also, note the difference in the scale used on the ordinate in Figure (4) and Figure (5).

At this point, we may appreciate that the position of the reflectivity minimum is controlled simply by the film thickness, but the strength of the interaction between the incident wave and the surface plasmon is a very sensitive function of the nature of the correlation between the surface roughness on the oxide-vacuum interface, and that on the oxide-substrate interface. The reason for this is the following, if we compare the results in Figure (4) and Figure (5). When $\zeta_1(x,y) = -\zeta_2(x,y)$, as in the non-uniform film model, the scattered electromagnetic wave from the oxide-vacuum interface interferes constructively within the oxide film with that which comes from the oxide-substrate interface. This greatly enhances the coupling between the incident radiation and the surface plasmon. Note that in Figure (5), coupling to the surface plasmon is strongest when d = 20Å. On the same curve, one sees an appreciable change in reflectivity above 9eV, well above the surface plasmon-induced reflectivity dip. The large roughness induced change in the reflectivity above 9eV comes from energy dissipation within the oxide layer; recall that one is past the absorption edge of the oxide film in this energy range. The constructive interference which produces strong coupling to the surface plasmon thus also leads to considerable absorption within the oxide film in the energy region above its absorption edge. We shall see

that in the random roughness model, where there is no correlation between the roughness on each interface (and hence no constructive interference of the type just described), in the presence of the oxide overlayer, the roughness induced coupling of the incident radiation to the surface plasmon is still considerably enhanced over the value for the pure aluminum surface, although the magnitude of the enhancement is smaller than for the non-uniform film model. This means that in the replicating film model, the two scattered fields evidently interfere destructively, and the enhancement effect provided by the oxide film is suppressed as a consequence.

In Figure (6), we present the results of our calculations for the random roughness model (Figure 1(c)). Again the position of the minimum in the change in reflectivity occurs at the same photon energy as for the replicating film model. The strength of coupling between the incident radiation and the surface plasmon is significantly larger than is the case for the pure aluminum surface, although the enhancement factor is considerably smaller in each case than for the non-uniform film model in Figure (5).

In Figure (7), we display the results of the calculations for the rough oxide layer model. As remarked earlier, we assume here that the oxide-substrate interface is perfectly smooth, but roughness is present on the oxide-vacuum interface. For small values of the oxide layer thickness, the interaction between the incident radiation and the surface plasmon is enhanced, as in the other two examples where destructive interference does not occur between the scattered fields generated by the pair of rough interfaces. However, as the oxide layer thickness increases, the strength of the effective

coupling eventually begins to decrease. Quite clearly this occurs because the fields associated with the surface plasmon are localized to the inner interface, and as the oxide layer thickness increases, the rough surface responsible for the reflectivity dip moves sufficiently far from the inner interface to cut off the coupling between the surface plasmons and the incident radiation.

With the results of the above four cases in hand, we make some remarks about the experimental data.

Feuerbacher and Steinman[15] have studied roughness induced reflectivity dips for aluminum films, and also for roughened films overcoated with 50Å of LiF. The position of the reflectivity minimum of the roughened aluminum film overcoated with 50Å of LiF agrees quite well with the calculations presented above. (Of course, our calculations were carried out for aluminum overcoated with aluminum oxide, but in the spectral regime of interest, both LiF and Al_2O_3 are transparent, and their dielectric constants do not differ greatly.) If one examines the magnitude of the reflectivity dip they observe, then for the roughened film $\Delta R_{MAX} \approx 0.25$, while for the overcoated film, $\Delta R_{MAX} \approx 0.45$. Thus, while the overcoating procedure shifts the reflectivity minimum toward the visible, it does not greatly affect the strength of the roughness-induced coupling of the incident radiation to the surface plasmon. This suggests that the LiF overlayer has roughness on its outer surface which tracks rather closely that on the LiF-substrate interface, as in our replicating film model of Figure 1(a).

Stanford and Bennett[17] have studied the effect of overcoating a roughened Ag surface with films of Al_2O_3 roughly 250Å thick.

They present several measurements in this paper. For a supersmooth uncoated Ag surface, they find a smooth variation of the reflectivity, with no sign of a dip characteristic of roughness induced coupling to surface plasmons. For an uncoated surface they characterize as "slightly rough", the measured reflectivity tracks that of the supersmooth surface, although a clear hint of a surface plasmon dip is present. The surface plasmon dip appears as a clear feature in data on a surface they characterize as "relatively rough." When the slightly rough surface is overcoated with Al_2O_3, a very large pronounced dip appears. The reflectivity change, only barely visible for the uncoated surface, assumes a maximum of ≈ 0.50 for the overcoated. While these measurements are carried out on a rather different substrate-overlayer system than that considered here (and also in a different wavelength regime), this data provides a clear example of the behavior illustrated in Figure (5) and Figure (6), where the overcoating produces an enormous enhancement of the roughness induced coupling of the incident radiation to the surface plasmon.

The calculations in Figure (7) suggest that if a supersmooth aluminum surface is overcoated with a dielectric, then if the dielectric layer is sufficiently thick, the reflectivity of the structure becomes relatively insensitive to the presence of roughness on the outer surface of the overlayer. However, it must be kept in mind that as the thickness of the oxide layer increases, the reflectivity of the structure drops substantially in the ultra violet even if both interfaces are perfectly smooth, as Ehrenrich has pointed out recently.[18] We illustrate this in Figure (8),

where we present the reflectivity for aluminum overcoated with an oxide layer of uniform thickness, for the case where both interfaces are perfectly smooth. The dot-dashed curve is the data of Banning,[11] which shows a decrease in reflectivity at large photon energies of the sort expected for a surface overcoated with an oxide film. Indeed, the data is fit reasonably by the curve for d = 20Å, for photon energies above 9eV. It is tempting to suggest that the measured reflectivity drops below the theoretical curve in the region from 7 to 9 eV because of roughness induced coupling to surface plasmons. However, it is difficult to see how superposition of two distinct mechanisms could produce a curve as smooth and featureless as the data of Banning.

APPENDIX: CONSTRUCTION OF THE GREEN'S FUNCTIONS FOR THE ELECTROMAGNETIC WAVE EQUATION

In Section II of the paper, we introduced a set of Green's functions $D_{\mu\nu}(\vec{x},\vec{x}';\omega)$ that satisfy the differential equations

$$\sum_{\mu}\left\{\frac{\omega^2}{c^2}\,\epsilon_o(z,\omega)\,\delta_{\lambda,\mu} - \frac{\partial^2}{\partial x_\lambda \partial x_\mu} + \delta_{\lambda\mu}\nabla^2\right\} D_{\mu\nu}(\vec{x},\vec{x}';\omega) = -4\pi\,\delta_{\lambda\mu}\delta(\vec{x}-\vec{x}') \tag{A-1}$$

along with the outgoing wave boundary conditions appropriate to the present scattering problem. In Eq. (A-1), the dielectric function $\epsilon_o(z,\omega)$ is given by Eq. (II-6).

In an Appendix of our preceding paper[5], we derived the form of these Green's functions for the semi-infinite dielectric, which corresponds to the limit $d \to o$ in the present geometry. In our preceding paper, we constructed the Green's functions by directly solving the differential equation Eq. (A-1). This procedure becomes most cumbersome for the present geometry. We present here a much more compact method of constructing the Green's functions.

As in the text, we write

$$D_{\mu\nu}(\vec{x},\vec{x}';\omega) = \int \frac{d^2k_\parallel}{(2\pi)^2}\, e^{i\vec{k}_\parallel\cdot(\vec{x}_\parallel-\vec{x}_\parallel')}\; d_{\mu\nu}(\vec{k}_\parallel\omega|zz') \tag{A-2}$$

and we note that one may write

$$\delta(\vec{x}-\vec{x}') = \delta(z-z')\int \frac{d^2k_\parallel}{(2\pi)^2}\; e^{i\vec{k}_\parallel\cdot(\vec{x}_\parallel-\vec{x}_\parallel')} \tag{A-3}$$

With these expressions, one may readily derive a set of one dimensional, coupled differential equations for the functions $d_{\mu\nu}(\vec{k}_\parallel \omega | zz')$. These equations simplify considerably if we perform a coordinate rotation which aligns the $\hat{x}$ axis with the direction of $\vec{k}_\parallel$. This is achieved by the action of the matrix

$$\underset{\sim}{S}(\vec{k}_\parallel) = \frac{1}{k_\parallel} \begin{pmatrix} k_x & k_y & 0 \\ -k_y & k_x & 0 \\ 0 & 0 & k_\parallel \end{pmatrix} . \tag{A-4}$$

We introduce a new set of functions $g_{\mu'\nu'}(k_\parallel \omega | zz'')$ related to $d_\mu(k_\parallel \omega | zz')$ by the rotation just described:

$$d_{\mu\nu}(\vec{k}_\parallel \omega | zz') = \sum_{\mu'\nu'} g_{\mu'\nu'}(k_\parallel \omega | zz') \; S_{\mu'\nu}(\vec{k}_\parallel) \; S_{\nu'\nu}(\vec{k}_\parallel) \tag{A-5}$$

It is a straightforward matter to construct the equations obeyed by the functions $g_{\mu\nu}(k_\parallel \omega | zz')$. These equations read

$$\left[\epsilon_o(z,\omega) \frac{\omega^2}{c^2} - k_\parallel^2 + \frac{d^2}{dz^2}\right] g_{yy}(k_\parallel \omega | zz' = 4\pi \delta(z-z') , \tag{A-6}$$

$$\left[\epsilon_o(z,\omega) \frac{\omega^2}{c^2} + \frac{d^2}{dz^2}\right] g_{xx}(k_\parallel \omega | zz') - i k_\parallel \frac{dg_{zx}}{dz}(k_\parallel \omega | zz') = 4\pi \delta(z-z') \tag{A-7}$$

$$-ik_\parallel \frac{d}{dz} g_{xx}(k_\parallel \omega | zz') + \left[\epsilon_o(z,\omega) \frac{\omega^2}{c^2} - k_\parallel^2\right] g_{zx}(k_\parallel \omega | zz') = 0 , \tag{A-8}$$

$$\left[\epsilon_o(z,\omega) \frac{\omega^2}{c^2} + \frac{d^2}{dz^2}\right] g_{xz}(k_\parallel \omega | zz') - i k \frac{d}{dz} g_{zz}(k_\parallel \omega | zz') = 0 \tag{A-9}$$

$$-ik_{\parallel} \frac{d}{dz} g_{xz} (k_{\parallel}\omega|zz') \tag{A-10}$$

$$+\left[\epsilon_o(z,\omega) \frac{\omega^2}{c^2} - k_{\parallel}^2\right] g_{zz}(k_{\parallel}\omega|zz') = 4\pi\delta(z-z')$$

The remaining functions (g_{xy}, g_{yx}, g_{yx}, g_{zy}) obey homogeneous equations, and thus vanish identically.

We begin with $g_{yy}(k_{\parallel}\omega|zz')$, since Eq. (A-6) is uncoupled with the remaining four equations. We first observe that for a medium characterized by the z dependent dielectric constant $\epsilon_o(z,\omega)$, Maxwell's equations yield solutions of the form

$$\vec{E}(k_{\parallel}\omega|\vec{x}) = \hat{y}\, E_y(k_{\parallel}\omega|z) e^{ik_{\parallel}x} \quad , \tag{A-11}$$

where $E_y(k_{\parallel}\omega|z)$ obeys the homogeneous version of Eq. (A-6):

$$\left[\epsilon_o(z,\omega) \frac{\omega^2}{c^2} - k_{\parallel}^2 + \frac{d^2}{dz^2}\right] E_y(k_{\parallel}\omega|z) = 0 \quad . \tag{A-12}$$

There are two linearly independent solutions of the differential equation Eq. (A-12). We denote the two solutions by $E_y^{>}(k_{\parallel}\omega|z)$, and $E_y^{<}(k_{\parallel}\omega|z)$, where we choose the functions to satisfy the boundary conditions

$$\lim_{z \to +\infty} E_y^{>}(k_{\parallel}\omega|z) = e^{+ik_o z} \tag{A-13}$$

and

$$\lim_{z \to -\infty} E_y^{<}(k_{\parallel}\omega|z) = e^{+ik_2 z} \quad , \tag{A-14}$$

where we define the quantities

$$k_{1,2} = \left(\frac{\omega^2}{c^2} \epsilon_{1,2} - k_{\parallel}^2\right)^{\frac{1}{2}} \quad , \quad \mathrm{Im}(k_{1,2}) < o \tag{A-15}$$

$$k_o = \left(\frac{(\omega + i\eta)^2}{c^2} - k_{\parallel}^2\right)^{\frac{1}{2}} \quad , \quad \mathrm{Im}(k_o) > o \quad . \tag{A-16}$$

In Eq. (A-15), we presume ϵ_1 and ϵ_2 have a positive, non-zero imaginary part, and for the proper square root to be chosen for k_o, we have added a positive imaginary infinitesimal $i\eta$ to the frequency. The limit $\eta \to o$ is always to be taken in Eq. (A-16).

The Green's function $g_{yy}(k_\parallel \omega | zz')$ is to be constructed so that Eq. (II-11) describes a scattered wave which radiates into the vacuum for $z > d$, and one that attenuates in the region $z < o$. This Green's function is simply expressed in terms of $E_y^>(k_\parallel \omega | z)$ and $E_y^<(k_\parallel \omega | z)$ as follows:[20]

$$g_{yy}(k_\parallel \omega | zz') = \frac{4\pi}{W_\perp(k_\parallel, \omega)} \left\{ E_y^>(k_\parallel \omega | z) E_y^<(k_\parallel \omega | z') \; \theta\,(z-z') + E_y^<(k_\parallel \omega | z) E_y^>(k_\parallel \omega | z') \; \theta\,(z'-z) \right\} , \tag{A-15}$$

where

$$W_\perp(k_\parallel, \omega) = \frac{\partial E_y^>}{\partial z}(k_\parallel \omega | z) \, E_y^<(k_\parallel \omega | z) - \frac{\partial E_y^<}{\partial z}(k_\parallel \omega | z) \, E_y^>(k_\parallel \omega | z) , \tag{A-16}$$

is the Wronskian, a quantity independent of z.[20]

The form given in Eq. (A-15) is valid for any function $\epsilon_o(z, \omega)$. For the particular geometry of concern here, where $\epsilon_o(z, \omega)$ is given by Eq. (II-6), it is a straightforward exercise to construct these two functions. One has

$$E_y^>(k_\parallel \omega | z) = \begin{cases} e^{+ik_o z} , & z > d \\ A_+^{(\perp)} e^{ik_1 z} + A_-^{(\perp)} e^{-ik_1 z} , & o < z < d \\ B_+^{(\perp)} e^{ik_2 z} + B_-^{(\perp)} e^{-ik_2 z} , & z < o \end{cases} \tag{A-17}$$

and

$$E_y^{<}(k_\parallel \omega|z) = \begin{cases} D_+^{(\perp)} e^{ik_o z} + D_-^{(\perp)} e^{-ik_o z} , & z > d \\ C_+^{(\perp)} e^{ik_o z} + C_-^{(\perp)} e^{-ik_1 z} , & o < z < d \\ e^{ik_z z} , & z < o \end{cases} \qquad \text{(A-18)}$$

where in these expressions, with σ = + or − ,

$$A_\sigma^{(\perp)} = \frac{e^{ik_o d}}{2} \left(1 + \sigma \frac{k_o}{k_1}\right) e^{-i\sigma k_\perp d} , \qquad \text{(A-19)}$$

$$B_\sigma^{(\perp)} = \frac{e^{ik_o d}}{2} \left[\left(1 + \sigma \frac{k_o}{k_1}\right)\cos(k_1 d) - i \left(\frac{k_o}{k_1} + \sigma \frac{k_1}{k_2}\right) \sin(k_1 d)\right] \qquad \text{(A-20)}$$

$$C_\sigma^{(\perp)} = \tfrac{1}{2} \left(1 + \sigma \frac{k_2}{k_1}\right) \qquad \text{(A-21)}$$

$$D_\sigma^{(\perp)} = \frac{e^{ik_o d}}{2} \left[\left(1 + \sigma \frac{k_2}{k_o}\right)\cos(k_1 d) + i \left(\frac{k_2}{k_1} + \sigma \frac{k_1}{k_o}\right) \sin(k_1 d)\right], \qquad \text{(A-22)}$$

and the Wronskian is given by

$$W_\perp(k_\parallel, \omega) = \frac{e^{ik_o d}}{k_1} \left[(k_1^2 - k_2 k_o) \sin(k_1 d) + i(k_o - k_2) k_1 \cos(k_1 d)\right] \qquad \text{(A-23)}$$

The functions $g_{xx}(k_\parallel \omega|zz')$ and $g_{zx}(k_\parallel \omega|zz')$ obey the coupled equations Eq. (A-7) and Eq. (A-8). These functions may be constructed by generalizing the method used to obtain $g_{yy}(k_\parallel \omega|zz')$.

We begin by noting that if we seek a solution of Maxwell's equations in the form

$$\vec{E}(k_\parallel \omega|\vec{x}) = \left\{\hat{x}\, E_x(k_\parallel \omega|z) + \hat{z}\, E_z(k_\parallel \omega|z)\right\} e^{ik_\parallel x} , \qquad \text{(A-24)}$$

then the functions $E_x(k_\parallel \omega|z)$ and $E_z(k_\parallel \omega|z)$ satisfy the coupled equations

$$\left[\epsilon_o(z,\omega)\,\frac{\omega^2}{c^2} + \frac{d^2}{dz^2}\right] E_x(k_\parallel \omega|z) - i\,k_\parallel \frac{d}{dz}\,E_z(k_\parallel \omega|z) = o \qquad \text{(A-25)}$$

$$-i\,k_\parallel \frac{d}{dz}\,E_x(k_\parallel \omega|z) + \left[\epsilon_o(z,\omega)\,\frac{\omega^2}{c^2} - k_\parallel^2\right] E_z(k_\parallel \omega|z) = o \quad . \qquad \text{(A-26)}$$

For the geometry under consideration here, where $\epsilon_o(z,\omega)$ is piecewise constant, we must have $\nabla\cdot\vec{E} = o$ everywhere except at the singular points $z = o$ and $z = d$. This requires (except at the two points)

$$\frac{d}{dz}\,E_z(k_\parallel \omega|z) + ik_\parallel\,E_x(k_\parallel \omega|z) = o \quad . \qquad \text{(A-27)}$$

Thus, if we are given $E_z(k_\parallel \omega|z)$, then from Eq. (A-27) we may compute $E_x(k_\parallel \omega|z)$ in each regime of interest. We confine our attention to $E_z(k_\parallel \omega|z)$ as a consequence.

There are two linearly independent sets of solutions of the system of equations from Eq. (A-25) through Eq. (A-27), just as when we examined Eq. (A-12). We append the superscript > to the set $E_x^>(k_\parallel \omega|z)$, $E_z^>(k_\parallel \omega|z)$ for which $E_z^>(k_\parallel \omega|z)$ obeys the boundary condition

$$\lim_{z \to +\infty} E_z^>(k_\parallel \omega|z) = e^{+ik_o z} \quad , \qquad \text{(A-28)}$$

and we append the superscript < to the set for which

$$\lim_{z \to -\infty} E_z^<(k_\parallel \omega|z) = e^{+ik_2 z} \quad . \qquad \text{(A-29)}$$

Before we proceed, we display the explicit form of the fields $E_{x,z}^>(k_\parallel \omega|z)$ and $E_{x,z}^<(k_\parallel \omega|z)$. One has

$$E_z^{>}(k_\parallel \omega|z) = \begin{cases} e^{ik_o z}, & z > d \\ A_+^{(\parallel)} e^{ik_1 z} + A_-^{(\parallel)} e^{-ik_1 z}, & o < z < d \\ B_+^{(\parallel)} e^{ik_2 z} + B_-^{(\parallel)} e^{-ik_2 z}, & z < o \end{cases} \tag{A-30}$$

$$E_x^{>}(k_\parallel \omega|z) = \begin{cases} -\frac{k_o}{k_\parallel} e^{ik_o z}, & z > d \\ -\frac{k_1}{k_\parallel}\left[A_+^{(\parallel)} e^{ik_1 z} - A_-^{(\parallel)} e^{-ik_1 z}\right], & o < z < d \\ -\frac{k_2}{k_\parallel}\left[B_+^{(\parallel)} e^{ik_2 z} - B_-^{(\parallel)} e^{-ik_2 z}\right], & z < o \end{cases} \tag{A-31}$$

and

$$E_z^{<}(k_\parallel \omega|z) = \begin{cases} D_+^{(\parallel)} e^{ik_o z} + D_-^{(\parallel)} e^{-ik_o z}, & z > d \\ C_+^{(\parallel)} e^{ik_1 z} + C_-^{(\parallel)} e^{-ik_1 z}, & o < z < d \\ e^{ik_2 z}, & z < o \end{cases} \tag{A-32}$$

$$E_x^{<}(k_\parallel \omega|z) = \begin{cases} -\frac{k_o}{k_\parallel}\left[D_+^{(\parallel)} e^{ik_o z} - D_-^{(\parallel)} e^{-ik_o z}\right], & z > d \\ -\frac{k_1}{k_\parallel}\left[C_+^{(\parallel)} e^{ik_1 z} - C_-^{(\parallel)} e^{-k_1 z}\right], & o < z < d \\ -\frac{k_2}{k_\parallel} e^{ik_2 z}, & z < o \end{cases} \tag{A-33}$$

In these expressions, one has, with σ = + or −,

$$A_\sigma^{(\parallel)} = \tfrac{1}{2}\left(\frac{1}{\epsilon_1} + \sigma \frac{k_o}{k_1}\right) e^{ik_o d} e^{-i\sigma k_1 d} \tag{A-34}$$

$$B_{\sigma}^{(\parallel)} = e^{ik_o d}\left[\left(\frac{1}{\epsilon_2} + \sigma\frac{k_o}{k_2}\right)\cos(k_1 d) - i\left(\frac{k_o\epsilon_1}{k_1\epsilon_2} + \sigma\frac{k_1}{\epsilon_1 k_2}\right)\sin(k_1 d)\right] \quad \text{(A-35)}$$

$$C_{+}^{(\parallel)} = \tfrac{1}{2}\left(\frac{\epsilon_2}{\epsilon_1} + \sigma\frac{k_2}{k_1}\right) \quad \text{(A-36)}$$

$$D_{+}^{(\parallel)} = \frac{e}{2}^{ik_o d}\left[\left(\epsilon_2 + \sigma\frac{k_2}{k_o}\right)\cos(k_1 d) + i\left(\epsilon_1\frac{k_2}{k_1} \sigma\frac{k_1\epsilon_2}{k_o\epsilon_1}\right)\sin(k_1 d)\right] \quad \text{(A-37)}$$

Given the fields defined in Eq. (A-30) - Eq. (A-3 , we seek solutions of Eq. (A-7) and Eq. (A-8) in the form

$$g_{xx}(k_\parallel\omega|zz') = \frac{4\pi}{\omega_\parallel(k_\parallel,\omega)}\left[E_x^{>}(k_\parallel\omega|z)\,E_x^{<}(k_\parallel\omega|z')\,\theta\,(z-z')\right. \quad \text{(A-38)}$$
$$\left. + E_x^{<}(k_\parallel\omega|z)\,E_x^{>}(k_\parallel\omega|z')\,\theta\,(z'-z)\right]$$

and

$$g_{zx}(k_\parallel\omega|zz') = \frac{4\pi}{\omega_\parallel(k_\parallel,\omega)}\left[E_z^{>}(k_\parallel\omega|z)\,E_x^{<}(k_\parallel\omega|z')\,\theta\,(z-z')\right. \quad \text{(A-39)}$$
$$\left. + E_z^{<}(k_\parallel\omega|z)\,E_x^{>}(k_\parallel\omega|z')\,\theta\,(z'-z)\right]$$

Substitution of these forms into Eq. (A-7) and Eq. (A-8) show that the solution indeed has the form of Eq. (A-38) and Eq. (A-39) if we choose

$$W_\parallel(k_\parallel,\omega) = W_{xx}(k_\parallel,\omega) - W_{zz}(k_\parallel,\omega) \quad \text{(A-40)}$$

where $W_{xx}(k_\parallel,\omega)$ and $W_{zz}(k_\parallel,\omega)$ are given by Eq. (A-18), but with y replaced by x or z.

Explicit calculation shows that for our geometry $W_{xx}(k_\parallel,\omega)$ and $W_{zz}(k_\parallel,\omega)$ are only piecewise constant, i.e. these functions are constant everywhere, but experience jump discontinuities at z = o and z = d. However, the function $\omega_\parallel(k_\parallel,\omega)$ is truly

constant, with a value everywhere given by

$$W_{\parallel}(k_{\parallel},\omega) = \frac{\omega^2 k_o}{ic^2 k_{\parallel}^2} e^{ik_o d}\left[\left(\epsilon_2 - \frac{k_2}{k_o}\right)\cos(k_1 d) + i\left(\epsilon_1 \frac{k_2}{k_1} - \frac{\epsilon_2 k_1}{\epsilon_1 k_o}\right)\sin(k_1 d)\right] . \tag{A-41}$$

Thus, we have left only the two functions $g_{zz}(k_{\parallel}\omega|zz')$ and $g_{xz}(k_{\parallel}\omega|zz')$. If we attempt to search for a solution of Eq. (A-9) and Eq. (A-10) by constructing the direct analogues of Eq. (A-38) and Eq. (A-39), we shall find the resulting functions fail to satisfy Eq. (A-9) and Eq. (A-10). We recall that when we explicitly constructed the Green's functions in reference (5), we found that g_{zz} contained a term directly proportional to $\delta(z-z')$. Thus, we look for a solution of the form

$$g_{xz}(k_{\parallel}\omega|zz') = \frac{4\pi}{W_{\parallel}(k_{\parallel},\omega)}\left[E_x^{>}(k_{\parallel}\omega|z)E_z^{<}(k_{\parallel}\omega|z')\ \theta\ (z-z') + E_x^{<}(k_{\parallel}\omega|z)E_z^{>}(k_{\parallel}\omega|z')\ \theta\ (z'-z)\right] \tag{A-42}$$

for g_{xz}, but for g_{zz} we take

$$g_{zz}(k_{\parallel}\omega|zz') = \Gamma(z')\delta(z-z') + \frac{4\pi}{W_{\parallel}(k_{\parallel},\omega)}\left[E_z^{>}(k_{\parallel}\omega|z)E_z^{<}(k\omega|z')\ \theta\ (z-z') + E_z^{<}(k_{\parallel}\omega|z)E_z^{>}(k_{\parallel}\omega|z')\ \theta\ (z'-z)\right] . \tag{A-43}$$

This form indeed solves the differential equation with $W_{\parallel}(k_{\parallel},\omega)$ given by Eq. (A-40) and Eq. (A-41) provided we choose

$$\Gamma(z') = \frac{4\pi c^2}{\omega^2 \epsilon_o(z',\omega)} . \tag{A-44}$$

We now have the explicit form for all the Green's functions required for the calculation of the scattered fields in each region of interest.

REFERENCES

(1) For example, see the experimental data reported by J. Endriz and W. Spicer, Phys. Rev. B4, 4144 (1971) along with the discussion presented by these authors.

(2) See the data in aluminum films deposited on fire polished substrates reported by B. P. Feuerbacher and W. Steinmann, Opt. Commun. 1, 81 (1969).

(3) J. L. Stanford and H. E. Bennett, Applied Optics 8, 2556 (1969).

(4) J. L. Stanford, J. Opt. Soc. Am. 60, 49 (1970).

(5) A. A. Maradudin and D. L. Mills, Phys. Rev. B (to be published).

(6) H. Froitzheim, H. Ibach and D. L. Mills, preprint entitled "Surface Optical Constants of Silicon and Germanium Derived from Electron Energy Loss Spectroscopy" (to be published).

(7) For example, see the discussion in Chapter XIX of A. Messiah, Quantum Mechanics, Vol. II, (North Holland Publishing Company, Amsterdam, 1962).

(8) To obtain Eq. (II-15), one needs to evaluate integrals of the form $\int dz' F(z')\delta(z'-z_o)$ where here z_o is either o or d. For the case of normal incidence considered, all functions $F(z')$ encountered are continuous at the point $z' = z_o$ where the argument of the delta function vanishes, in contrast to the functions encountered in reference (5), when the case where incident radiation that strikes the surface at non-normal incidence was also considered.

(9) The definition of φ_s used here differs from that in reference (5). The two are related by a 90^o rotation, so $\sin^2\varphi_s$ appears here where $\cos^2\varphi_s$ appeared previously, and conversely.

(10) E. T. Arakawa and M. W. Williams, J. Phys. Chem. Solids 29, 735 (1968).

(11) H. Ehrenreich, H. R. Phillipp and B. Segall, Phys. Rev. 132, 1918 (1963).

(12) J. G. Endriz and W. E. Spicer, Phys. Rev. B4, 4144 (1971).

(13) J. M. Elson and R. H. Ritchie, Phys. Rev. B4, 4129 (1971).

(14) J. M. Elson and R. H. Ritchie, Phys. Stat. Sol. (to be published).

(15) B. P. Feuerbacher and W. Steinman, Optics Communications 1, 81 (1969).

(16) For example, see the work of J. J. Cowan and E. T. Arakawa, Phys. Stat. Sol. 1, 695 (1970).

(17) J. L. Stanford and H. E. Bennett, Applied Optics 8, 2556 (1969).

(18) H. Ehrenreich, preprint entitled "UV Optical Properties of Oxide Coated Aluminum Films", to be published.

(19) M. Banning, J. Opt. Soc. Am. 32, 98 (1942).

(20) See the discussions in Chapter 8 of Principles and Techniques of Applied Mathematics, B. Friedman (John Wiley and Sons, New York, 1956).

FIGURE CAPTIONS

Figure (1): Illustration of the roughness on the vacuum-oxide-interface relative to that on the oxide-substrate interface for the four cases

(a) $\zeta_1 = \zeta_2$ (replicating film model)

(b) $\zeta_1 = -\zeta_2$ (non-uniform film model)

(c) $\langle\zeta_1^2\rangle = \langle\zeta_2^2\rangle$, but $\langle\zeta_1\zeta_2\rangle = o$ (random roughness model)

(d) $\zeta_2 \equiv o$, but $\zeta_1 \neq o$ (rough oxide layer model)

Figure (2): The geometry considered in the present paper. The thickness of the oxide layer, $\zeta_1(x,y)$ measures the position of a point on the oxide-vacuum interface from the plane $z = d$, and $\zeta_2(x,y)$ measures the position of a point on the oxide-substrate interface from the plane $z = d$.

Figure (3): The change in reflectivity for a rough aluminum surface, for the case where the transverse correlation length is chosen to be 200Å, and the rms amplitude of the surface roughness is 12Å

Figure (4): The change in reflectivity for an aluminum substrate overcoated with oxide films of various thicknesses. The calculations have been carried out for the replicating film model illustrated in Figure 1(a).

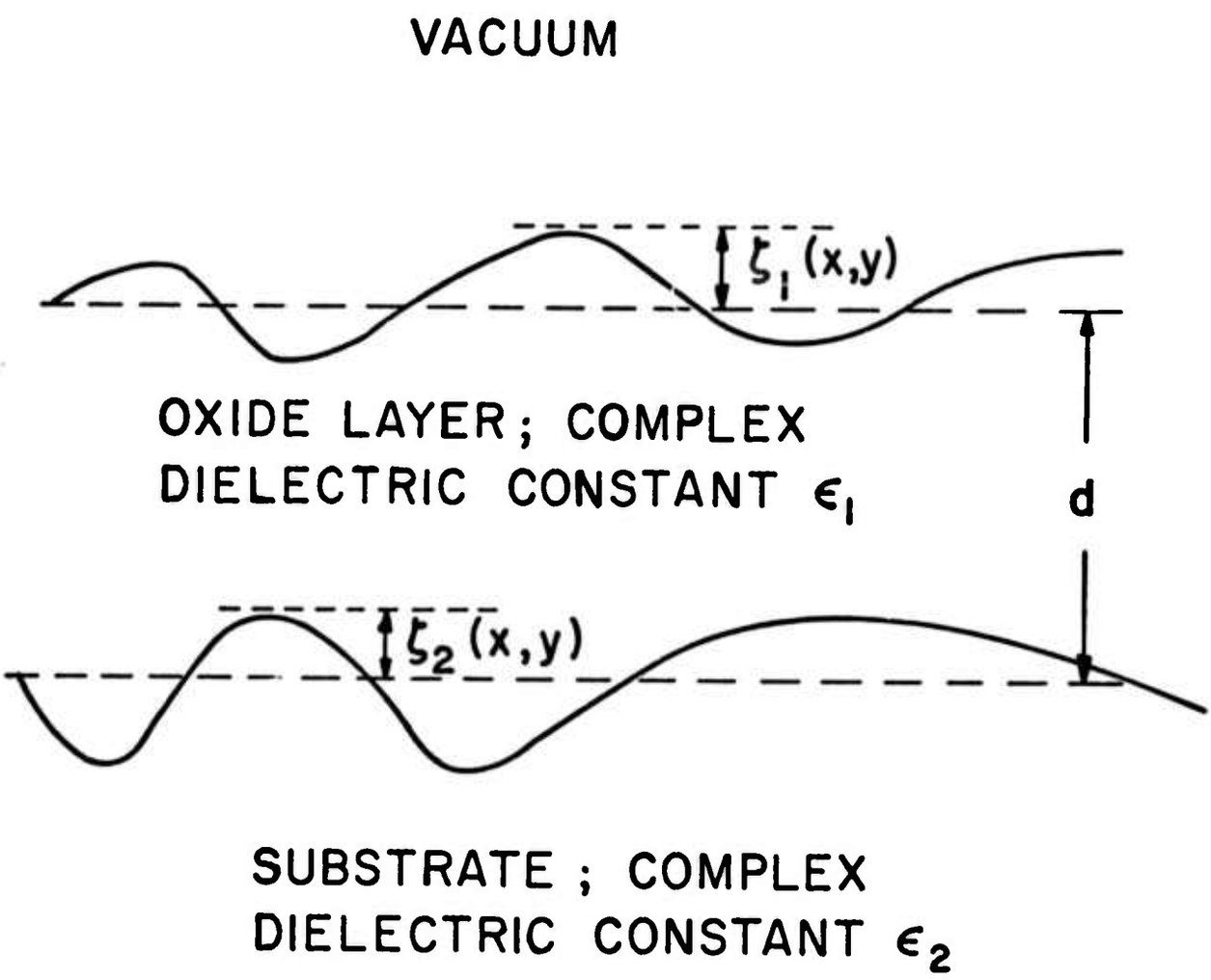
Figure (5): The change in reflectivity produced by surface roughness, for aluminum overcoated with an oxide film. The calculations have been carried out for parameters given in the figure, and for the non-uniform film model illustrated in Figure 1(b).

Figure (6): The change in reflectivity produced by surface roughness, for aluminum overcoated with oxide films of various thickness.

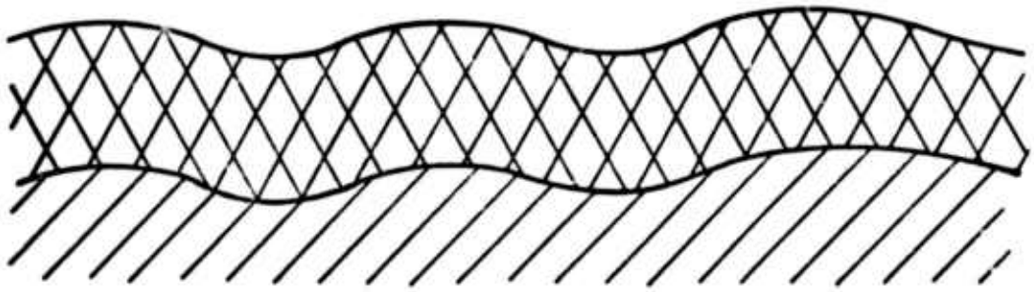
The calculations have been carried out for the parameters given in the figure, and for the random roughness model illustrated in Figure 1(c).

Figure (7): The change in reflectivity produced by surface roughness, for aluminum overcoated with oxide films of various thickness. The calculations have presumed the oxide-substrate interface is perfectly smooth, with roughness on only the oxide-vacuum interface. The parameters used in the calculation are given on the figure.

Figure (8): The reflectivity of an aluminum surface overcoated with oxide in the ultra violet, for various oxide thicknesses. In these calculations, it is presumed that both interfaces are perfectly smooth.

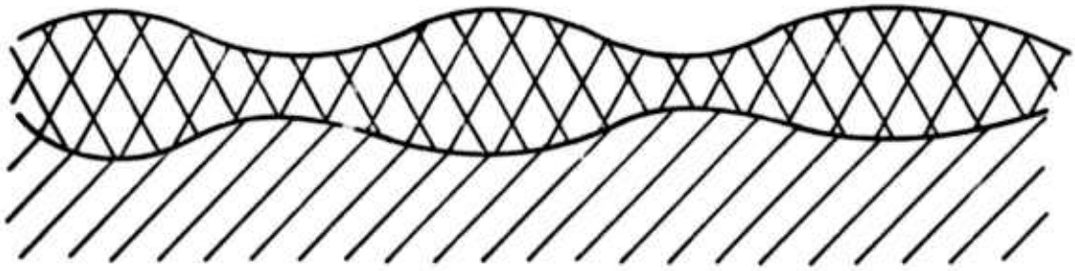

VACUUM
$\zeta_1(x,y)$
OXIDE LAYER; COMPLEX
DIELECTRIC CONSTANT ϵ_1
d
$\zeta_2(x,y)$
SUBSTRATE ; COMPLEX
DIELECTRIC CONSTANT ϵ_2

(a)



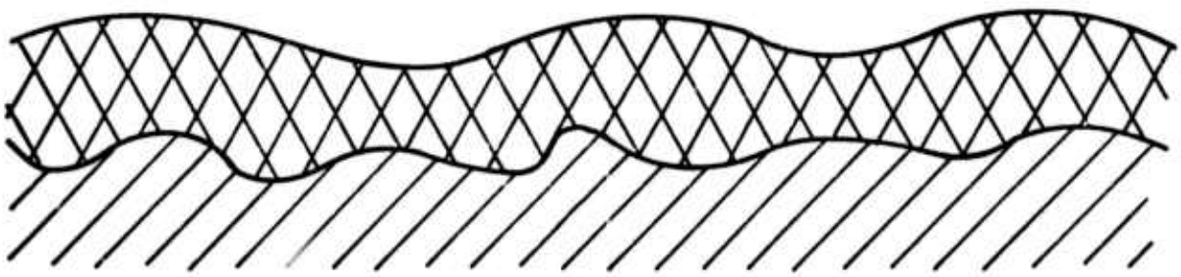
$$\zeta_1 = + \zeta_2$$

(b)



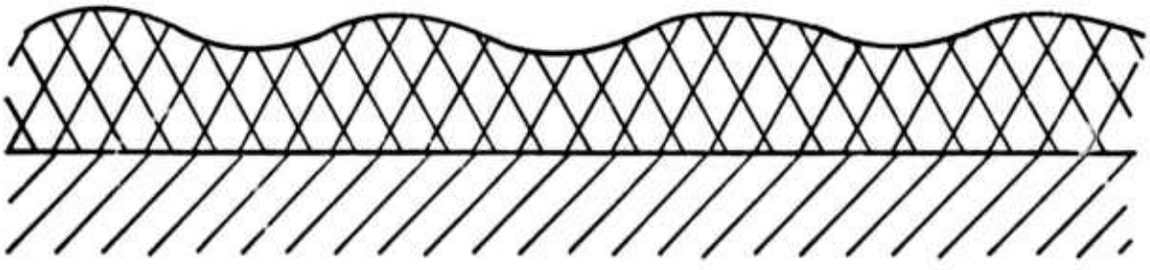
$$\zeta_1 = - \zeta_2$$

(c)

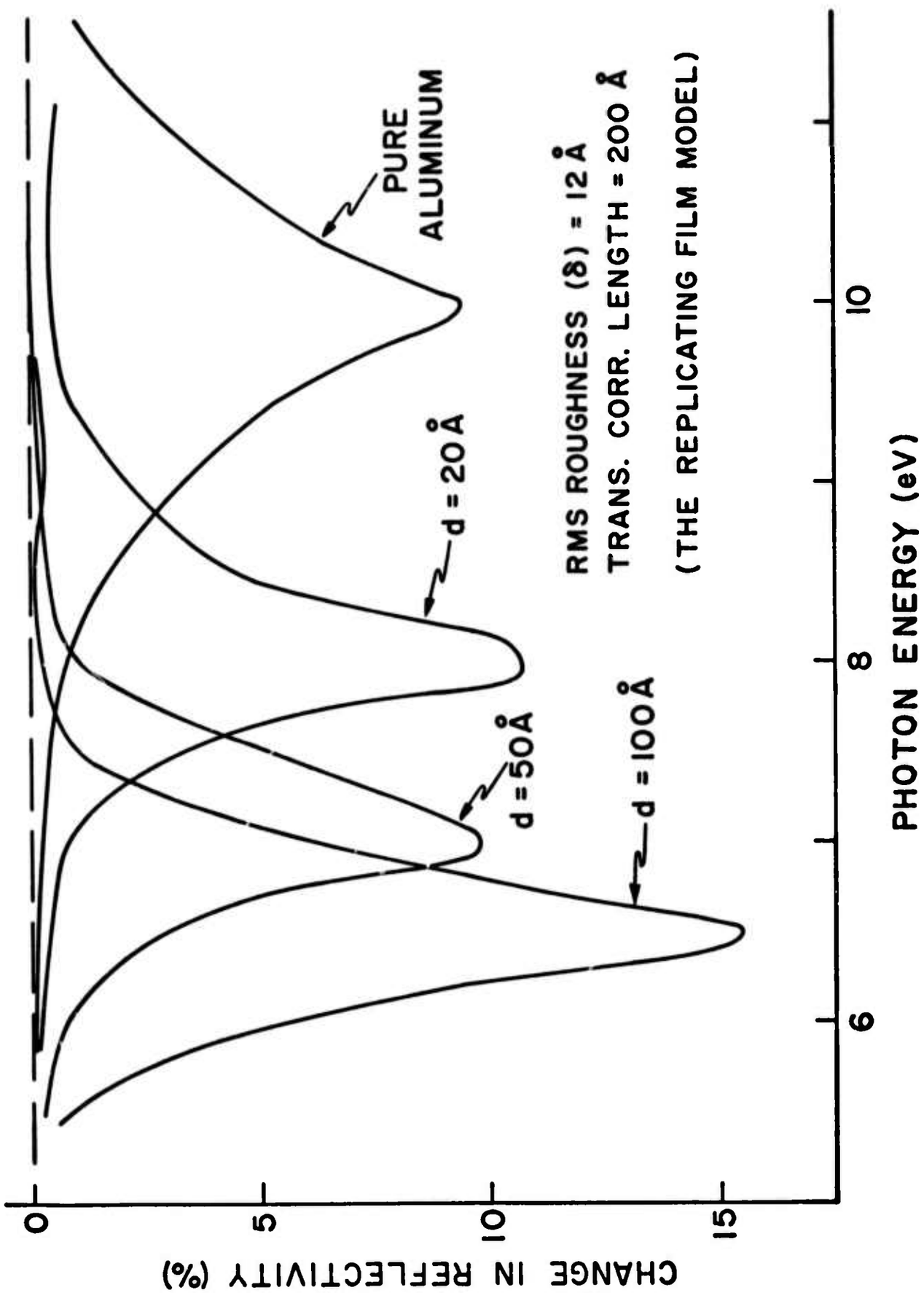


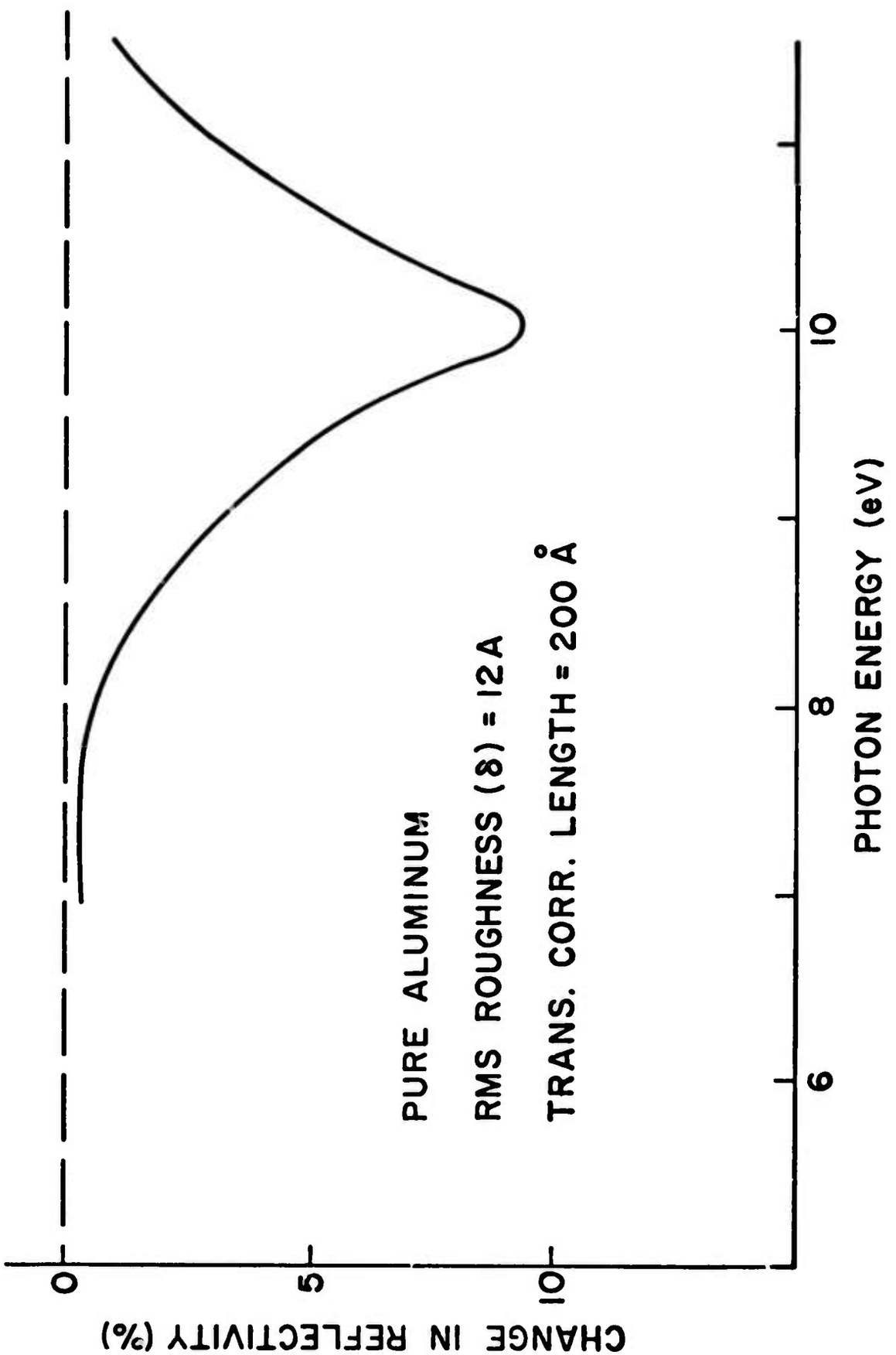
$$\langle \zeta_1^2 \rangle = \langle \zeta_2^2 \rangle \; ; \; \langle \zeta_1 \; \zeta_2 \rangle = 0$$

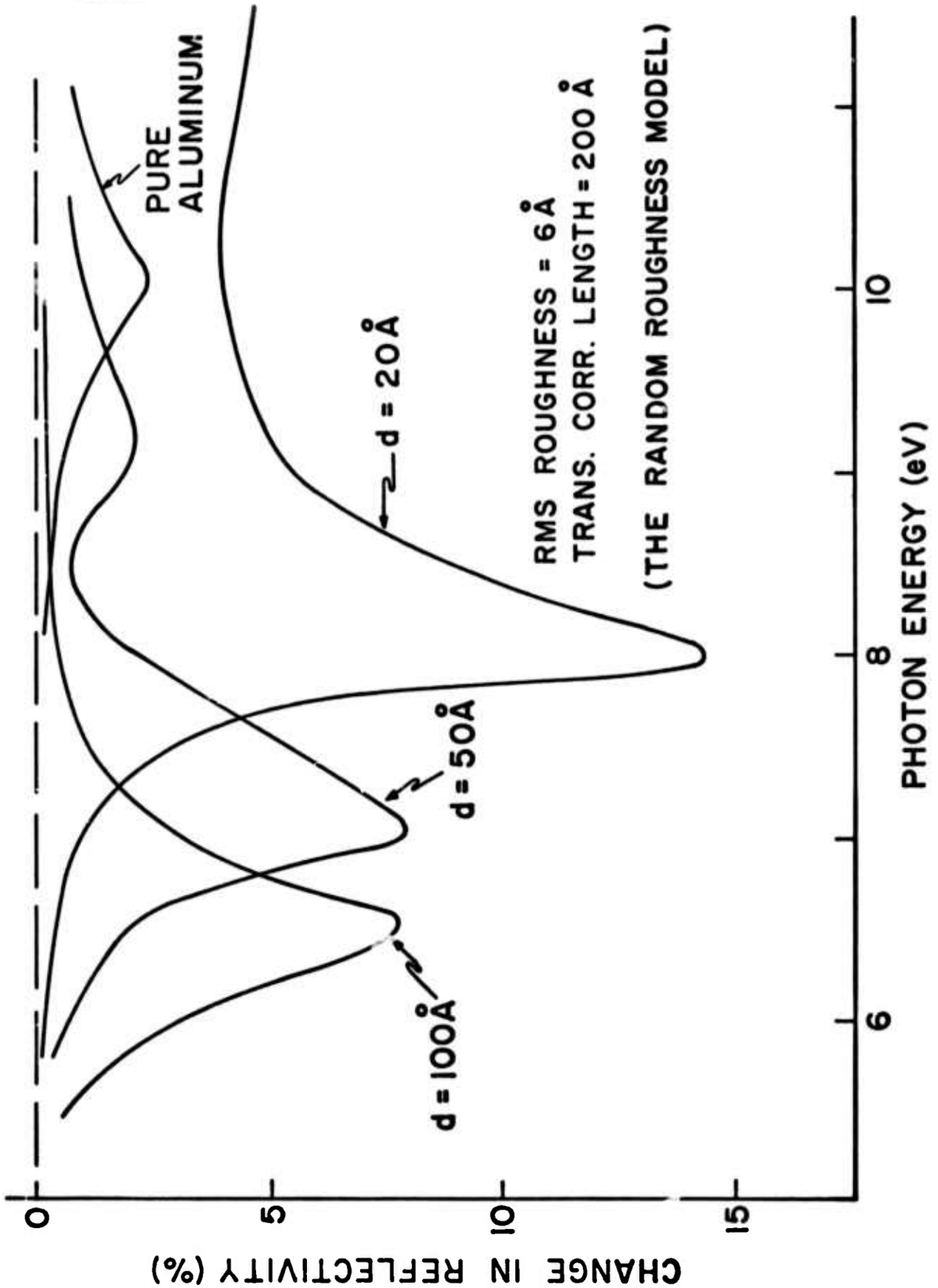
(d)

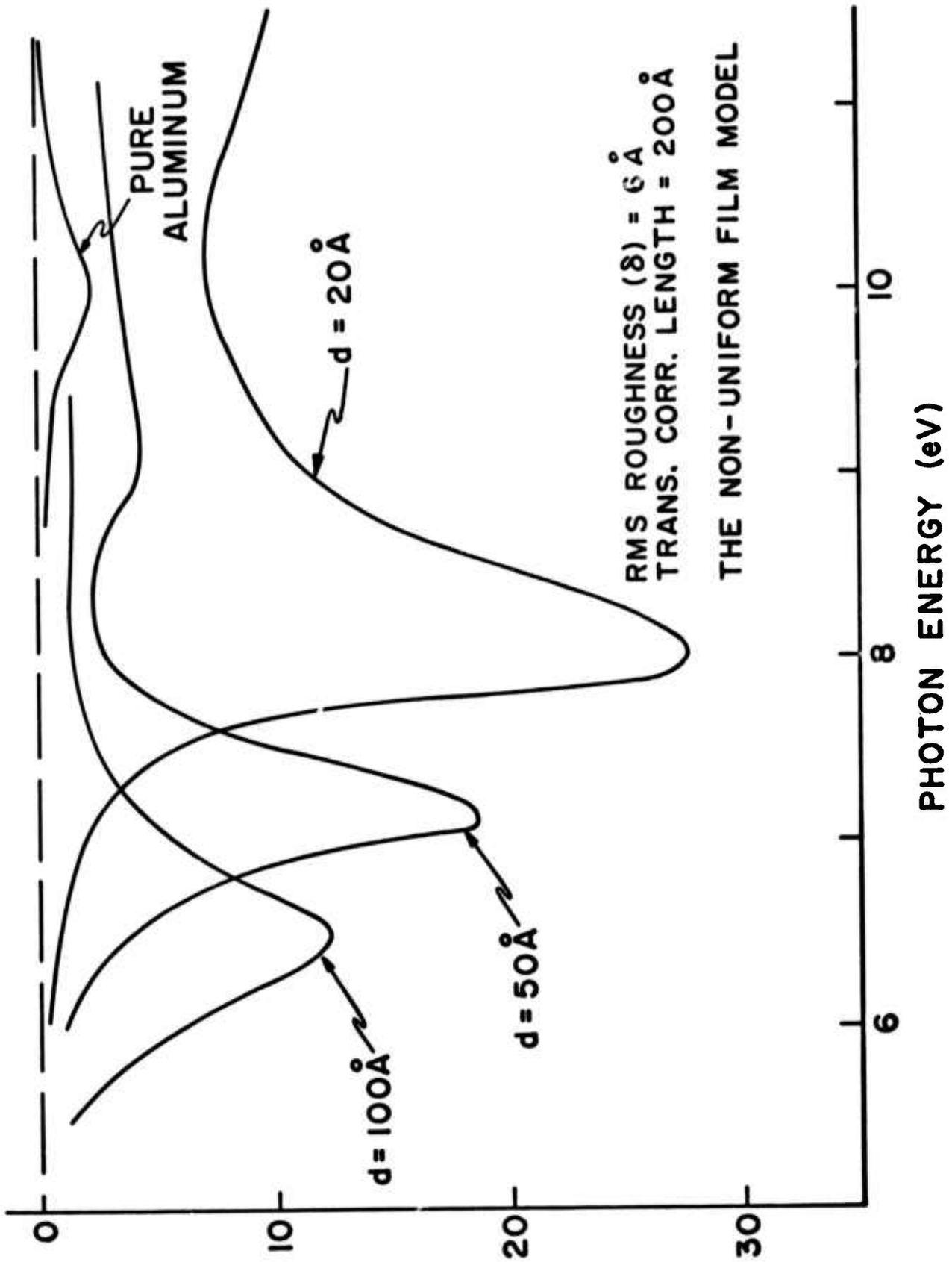


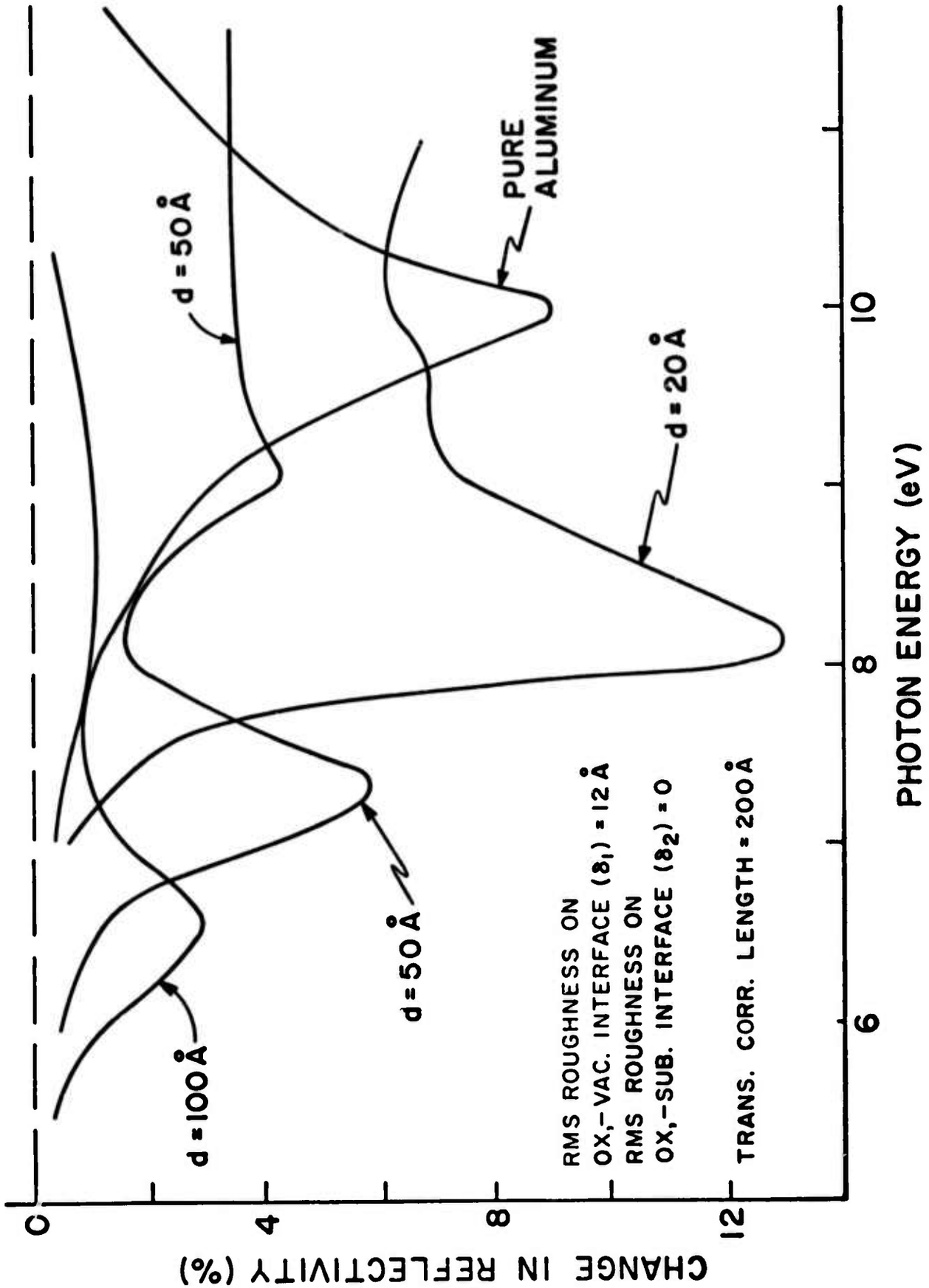
$$\zeta_1 = 0 \, , \qquad \zeta_2 \equiv 0$$

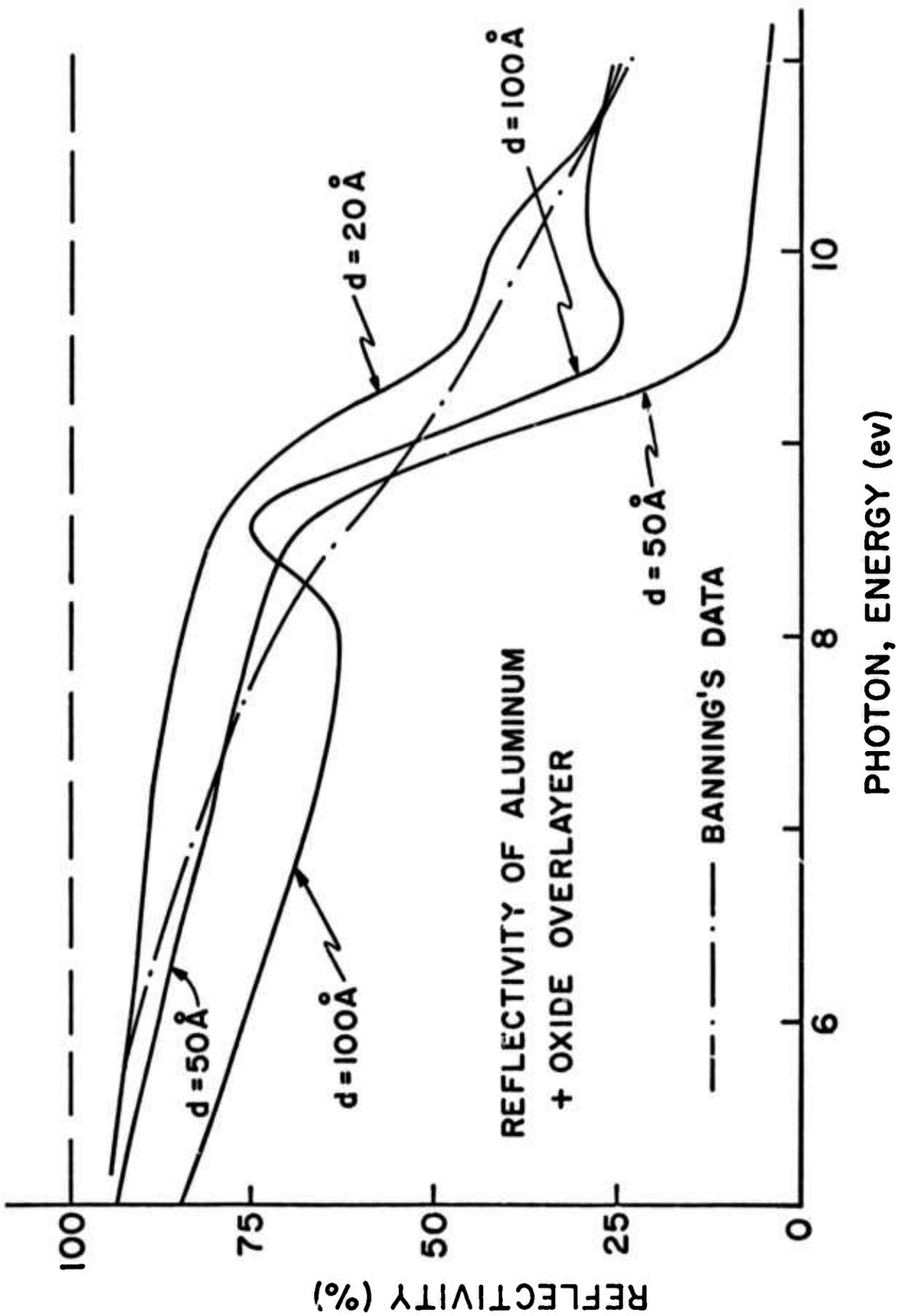

CHANGE IN REFLECTIVITY (%)
0
5
10
15
6
8
10
PHOTON ENERGY (eV)
PURE ALUMINUM
d = 20 Å
d = 50 Å
d = 100 Å
RMS ROUGHNESS (δ) = 12 Å
TRANS. CORR. LENGTH = 200 Å
(THE REPLICATING FILM MODEL)


CHANGE IN REFLECTIVITY (%)
0
5
10
6
8
10
PHOTON ENERGY (eV)
PURE ALUMINUM
RMS ROUGHNESS (δ) = 12A
TRANS. CORR. LENGTH = 200 Å


PURE ALUMINUM
d = 20Å
d = 50Å
d = 100Å
RMS ROUGHNESS = 6Å
TRANS. CORR. LENGTH = 200 Å
(THE RANDOM ROUGHNESS MODEL)
CHANGE IN REFLECTIVITY (%)
0
5
10
15
6
8
10
PHOTON ENERGY (eV)


CHANGE IN REFLECTIVITY (%)
0
10
20
30
6
8
10
PHOTON ENERGY (eV)
THE NON-UNIFORM FILM MODEL
RMS ROUGHNESS (δ) = 6 Å
TRANS. CORR. LENGTH = 200 Å
PURE ALUMINUM
d = 20 Å
d = 50 Å
d = 100 Å


CHANGE IN REFLECTIVITY (%)
0
4
8
12
6
8
10
PHOTON ENERGY (eV)
PURE ALUMINUM
d = 50 Å
d = 20 Å
d = 50 Å
d = 100 Å
RMS ROUGHNESS ON OX.-VAC. INTERFACE (σ1) = 12 Å
RMS ROUGHNESS ON OX.-SUB. INTERFACE (σ2) = 0
TRANS. CORR. LENGTH = 200 Å


REFLECTIVITY (%)
100
75
50
25
0
6
8
10
PHOTON, ENERGY (ev)
d = 50Å
d = 100Å
d = 20Å
d = 100Å
d = 50Å
REFLECTIVITY OF ALUMINUM
+ OXIDE OVERLAYER
BANNING'S DATA

F. THEORY OF LASER HEATING OF SOLIDS: I. METALS*

M. Sparks
Xonics, Incorporated, Van Nuys, California 91406

Calculations of the transient and steady-state temperature rise T of laser-irradiated metals indicate that the intensities I_f that cause failure of laser-irradiated metals have drastically different magnitudes and dependences on material parameters such as thermal conductivity K and heat capacity C, on the sample thickness ℓ and lateral dimension L, and on the operating conditions such as the laser-beam diameter D and pulse duration t. The behavior depends on the relative magnitudes of D, ℓ, L, and the thermal diffusion distance $d = (tK/C)^{1/2}$. The results suggest that the highest of the recently measured copper damage thresholds of 125 to 750 J/cm^2 for 0.6 μsec pulses at 10.6 μm are likely to be at or at least quite near the intrinsic li . set by the simple process of melting that results from the intrinsic absorption. The theoretical intensity at which the cavity mirrors of recently developed xenon uv lasers fail is in good agreement with the experimental values. The theoretical value of T for metals irradiated for 20 seconds with 10.6 μm radiation is two orders of magnitude too small to explain recent experimental results. It is suggested that the discrepancy is related to plasma ignition at the sample surface. The steady-state value of T for metals cooled with a surface-heat-transfer coefficient h is not reduced substantially by increasing the cooling efficiency past a certain point ($h > h_\ell \equiv K/\ell$). For $t < \tau$, where the characteristic time τ depends on both ℓ and h, cooling the metal is not effective in preventing the temperature rise.

Sec. F

I. INTRODUCTION

The general problems of heating of materials, including transparent window materials, reflecting materials for mirrors, and absorbing target materials, have become increasingly more important in the last four years as interest continues to increase in high-power laser systems. In addition to understanding the results of a given experiment, it is important to understand the effect of changing material parameters in order to predict and interpret the difference in performance of different materials. It is also important to understand the effects of changing the experimental conditions such as beam diameter, intensity, pulse duration, etc., especially in view of the fact that material tests are often performed under different conditions from those of normal operation. In particular, there is increasing interest in small-scale tests to represent large-scale operating conditions. Small-diameter beams are used to obtain high intensities with low powers.

Figures of merit[2] for transparent window materials have been developed for "pulsed" operation, that is pulses with duration of a second or two,[3] and for continuous operation, and there have been numerous studies of damage by single short pulses (typically ~10 nsec or 1 μsec). The relation between these various conditions has not been considered previously. A previous study[4] of surface cooling and boundary-layer heating emphasized the spatially averaged temperature, which is the important temperature in optical distortion[2] in transparent material having thickness ℓ much less than the laser beam diameter D.

The results of previous studies[5,6] of laser ignition of plasmas (laser supported detonation waves and laser supported combustion waves) at solid surfaces, which appear to have been overlooked in recent studies of laser damage to materials, are

shown to be important in obscuring the usual heating results in some cases, but negligible in others. Also, the results of the present study of heating are applicable to the laser-ignition studies.

The nature of the heating of a material by a laser beam is different, depending on the type of material, the operating conditions, and the mode of operation. Important operating conditions include the relative magnitudes of the laser-pulse duration and several characteristic times related to the sample dimensions, laser-beam diameter, and the amount of cooling. Modes of operation include single pulse, finite number of repeated pulses, infinitely repeated pulses, pulsed, and continuously operated systems. Furthermore, the relative importance of various failure mechanisms, such as material melting or fracture, avalanche breakdown,[7] or thermally induced optical distortion,[2] is different for different operating conditions. For example, sodium chloride is considered by many investigators to be a good candidate for windows for use in single-nanosecond-pulse high-power 10.6 μm systems; whereas it is a poor candidate for pulsed or continuous high-power 10.6 μm systems. This is because the high-intensity, single-pulse damage threshold is high, while the heating problem is severe as a result of the high intrinsic multiphonon absorption.

In the present series of two papers, an overview of laser-heating problems of current interest in high-power laser research is obtained by presenting several new results and summarizing several previously known results.[4,8,9] Previous results of two unpublished Parke Mathematical Laboratories reports[10] are included since the results are not readily available. Some of the present results were presented at the 1974 Advanced Research Projects Agency Materials Research Council Study on High Power Laser Mirror Heating, La Jolla, California.

The temperature rise ΔT as a function of time and the fluence F_f at which failure occurs are calculated for various experimental conditions of current interest. The results are sufficiently simple to show the effects of changing material parameters and experimental conditions, but are sufficiently accurate for estimating temperature rises and failure fluences. The relation between the steady-state results with cooling and the single-pulse and repeated-pulse results are considered.

An exact analysis of material damage by laser heating is quite complicated. The following simplifying assumptions render the problems tractable and afford reasonable estimates of the values of fluence or intensity at which failure occurs. Radiation and convection are neglected, which is a good approximation in most cases of current interest. For long pulses, high temperatures, and no intentional cooling, these cooling mechanisms could be important, of course. The temperature dependence of the thermal conductivity K and heat capacity C are neglected. For room-temperature ambient conditions the errors are of the order of 10 percent. For operation at low temperature, using constant average values of K and C could lead to greater errors. The absorptance A for metals and the absorption coefficient β for transparent materials vary more with temperature than do K and C in general. Nevertheless, constant average values of A and β are used. Further approximations are discussed as they arise.

The criterion for failure of a material varies with the application and is typically complicated. Here it is assumed that the failure occurs at a specified value ΔT_f of the temperature rise ΔT. Typical values of ΔT_f are the melting temperature ΔT_m (measured with respect to the ambient temperature), the fracture temperature, or the value of ΔT at which thermally induced optical distortion[2] becomes intolerable.

For optical components, the melting of the first small mass usually constitutes failure. Then the heats of fusion and vaporization and the heat required to raise the liquid phase to the boiling point are not important. In other applications, a large mass of the sample may be melted or vaporized. Melting is considered briefly in Sec. IV.

In this first paper, on laser heating of metals, fracture is not considered since metals are known to fail by melting. However, Wang and coworkers[8] have pointed out that fatigue may be important in repeated pulses. Thermally induced stresses have been considered elsewhere.[8,4] In a metal, the radiation is absorbed in a thin layer at the surface. The thickness of the layer, that is the skin depth, is typically a few hundred angstroms or less at optical frequencies. It will be shown that this is sufficiently thin to be completely negligible, and that the resulting heating can be considered as a surface source. There are other important systems in which energy is absorbed in a thin surface layer. Examples include absorption by a thin antireflection or protective coating on an optical element, such as a window or a lens, absorption in the thin metallic partial reflector in vuv lasers, absorption by a material having a very large optical absorption coefficient β, such as infrared absorption by optical elements that are transparent to visible, but highly absorbing at infrared wavelengths, absorption by painted surfaces, and absorption by dirty, contaminated, or damaged surfaces. Thus, the nonzero thickness of the absorbing layer will be included for generality.

The most important operating case is that of a large-diameter D beam ($D >> \ell$, where ℓ is the sample thickness). Furthermore, the thermal diffusion distance d is often much shorter than D or ℓ; then the results can be obtained from a simple linear-heat-flow model. In tests, smaller diameters are often used to obtain high

intensities with relatively low-power lasers. Then $D \ll \ell$ may be satisfied. The large-beam case $D \gg \ell$ is considered first. An exact infinite-series solution to the heat-flow equation (with no melting) is obtained for the problem of a slab of thickness ℓ originally at a constant temperature with heat produced at a constant rate for $t > 0$ in an absorbing layer of thickness $\delta (0 < z < \delta)$ with no heat flow at the surfaces at $z = 0$ and $z = \ell$. The approximation of constant heat production in the surface layer $(0 < z < \delta)$ is quite accurate in the case of an antireflection or protective coating of thickness δ with $\beta\delta \ll 1$, where β is the absorption coefficient. In other cases, such as the absorption by a homogeneous optical material with $\beta\ell \gg 1$, replacing the exponential heating by the step-function heating will cause some errors in the temperature, especially at small times. When greater accuracy is required for the exponential-heating case, the solution at the end of Appendix B can be used.

The central results for this case of $D \gg \ell$ for the premelting temperature distribution are illustrated in Figs. F1-3 of Sec. III, which show the temperature at the heated surface as a function of time and the temperature as a function of z for several values of time. The exact infinite-series solution for T (with no melting) is derived in Appendix A. In Sec. III, limiting cases of the exact results are discussed, and the results are explained by simple intuitive arguments. In Sec. IV, a simple treatment of melting is given. The small-beam case of $D \ll \ell$ is solved for the first time and the result explained physically in Sec. V. Steady-state results are derived in Sec. VI, and additional applications are considered in Sec. VII. The summary and discussion of Sec. VIII contains a tabulation of key-equation results. In Appendix B, an alternate approximation solution for T (with no melting), which includes the case of an exponential heat source $S \sim \exp(-z/\ell)$, is given.

Important results are denoted by underscored equation numbers.

II. CHARACTERISTIC TIMES, DISTANCES, AND HEAT-TRANSFER COEFFICIENT

In heat-flow and surface heat-transfer problems, three time constants

$$\tau_\ell = C\ell^2/K, \qquad \tau_s = CK/h^2, \qquad \tau_c = C\ell/h = (\tau_\ell \tau_s)^{1/2}, \qquad (2.1)$$

and the related distances

$$d = (Kt/C)^{1/2}, \qquad L_s = (K\tau_s/C)^{1/2} = K/h, \qquad L_c = (K\tau_c/C)^{1/2}, \qquad (2.2)$$

and the characteristic surface-heat-transfer coefficient

$$h_\ell = K/\ell \qquad (2.3)$$

are important in developing intuition. The first characteristic time τ_ℓ is the time required for heat to diffuse across the thickness ℓ of the disk, roughly speaking. The second time constant τ_s determines the approach of the surface temperature of a semi-infinite medium to its steady-state value. The third τ_c is the time constant for the exponential decay of the temperature of a thermally thin slab ($\ell \ll L_s$, as defined below). Factors of order unity ($4/\pi$ or $1/4$ for example) can appear with the time constants in (2.1) in specific problems. Solving the first equation in (2.1) for ℓ gives $\ell^2 = K\tau_\ell/C$ -- thus the definition of L_s and L_c in (2.2).

It will be shown below that the thermal behavior of the disk will be quite different for the two limiting cases of a <u>thermally thin</u> disk ($\ell \ll L_s$, or $\tau_\ell \ll \tau_s$) and a <u>thermally thick</u> disk ($\ell \gg L_s$, or $\tau_s \ll \tau_\ell$). Good thermal conductors (large K) with poor coolants (small h) tend to be thermally thin, and vice versa. Thermally thick also implies <u>well cooled</u> since $h \gg h_\ell$, where h is the surface heat-transfer

coefficient, is satisfied when $\ell >> L_s$ is satisfied. In fact, it is easy to show from the definitions (2.1)-(2.3) that

$$\ell / L_s = (\tau_\ell / \tau_s)^{1/2} = \tau_\ell / \tau_c = h / h_\ell = h\ell / K \quad . \qquad \underline{(2.4)}$$

Finally, for very short times, the time constant $\tau_\delta \equiv C\delta^2 / K$, where δ is the skin depth, is important.

III. LARGE BEAM, $D >> \ell$

First consider the temperature in an uncooled sheet of metal of thickness ℓ heated by a laser beam of constant intensity I over the face of the metal. This is a good approximation to the common operating case of $D >> \ell$. The heat leaving through the surfaces of the metal typically is negligible. Since the heat flow Q is equal to $-K\nabla T$, where ∇T is the gradient of the temperature, the boundary condition at the surfaces is $\hat{z} \cdot \nabla T = 0$, where $\hat{z}$ is the normal to the surface. Since there is no heat flow through the edges of the sample, the problem is one dimensional and the temperature distribution is the same as that of an infinite film of thickness ℓ with

$$dT/dz = 0 \qquad \text{at } z = 0, \ell \ .$$

The initial temperature distribution is assumed to be constant, and the value of the constant is taken as zero for convenience. Thus, T will always be measured with respect to the initial temperature, and ΔT can now be written as T.

The radiation is absorbed in a thin region, called the absorbing layer, at the irradiated surface of the metal at $z = 0$. For 10.6 μm radiation and values of the electrical conductivity appropriate to good metallic conductors, the skin depth typically is of the order of 10^{-2} μm. Since this depth is so small, the shape of the variation of the intensity with depth is unimportant except for very short times, as mentioned above and shown below. Thus, the source S of heat (W/cm^3) in the heat-flow equation will be taken as

$$S = (IA/\delta)\,\theta(\delta - z)\,\theta(t) \ , \tag{3.1}$$

where θ is the unit step function, $\delta \cong 10^{-2}\,\mu m$, I is the incident intensity, and A is the absorptance of the surface. It is assumed that K, A, and the heat capacity per unit volume C are constants. For large temperature changes this assumption is not well satisfied in general. Sufficient accuracy for many applications can be obtained by using average values of K, A, and C; otherwise numerical methods are required in general. The results of this model apply to cases other than that of infrared heating of metals, as discussed in Sec. I.

The differential equation and boundary conditions and initial conditions appropriate to this problem are

$$-K\,\partial^2 T/\partial z^2 + C\,\partial T/\partial t = S \ , \tag{3.2a}$$

$$T(z,0) = 0 \ , \qquad \partial T/\partial z = 0 \ , \quad \text{at } z = 0, \ell \text{ for all } t \ , \tag{3.2b}$$

and T and dT/dz are continuous at $z = \delta$; S is defined in (3.1). An exact solution for arbitrary δ is obtained in Appendix A. The results are illustrated in Figs. F1-3 for the case of $\delta \ll \ell$. Fig. F1 is a plot of the surface temperature $T(0,t)$ as a function of time from (A12). Fig. F2 is a plot of the spatial dependence of T near the interface of the absorbing layer and the bulk of the material from (A13), and Fig. F3 is a plot of T from (B2) adapted from Fig. 15, Chapter 3 of Carslaw and Jaeger,[9] showing the temperature distribution across the sample at several times.

These central results can be understood as follows: The thermal time constant

$$\tau_\delta \equiv C\delta^2/K \tag{3.3}$$

is the time constant for heat to diffuse a distance δ, roughly speaking. For $t \ll \tau_\delta$,

the heat generated in the absorbing layer $0 < z < \delta$ does not have time to diffuse out of the layer. Thus the temperature rise is the same as that of a uniformly heated slab of thickness δ with $dT/dz = 0$ at $z = 0$ and δ. From (3.2) with $-K\partial^2 T/\partial z^2 = 0$ and $S = IA/\delta$ in the absorption layer, the solution is

$$T \cong \frac{IA}{C\delta}\, t\theta(\ell - z)\,\theta(t), \qquad \text{for } t \ll \tau_\delta\ , \tag{3.4}$$

in agreement with the limit (A4b) of (A2). This result (3.4) also can be obtained from the definition of the heat capacity

$$CV = \Delta E/T\ , \tag{3.5}$$

where $\Delta E/tV$ is the energy added to volume V per unit time

$$\Delta E/tV = IA/\delta\ , \tag{3.6}$$

and T is measured with respect to the initial value. Combining these results gives $T = IAt/C\delta$ in the absorption layer, in agreement with (3.4). The effect of the heat diffusion at the edge $z = \delta$ of the absorbing layer is shown in Fig. F2.

The time constant $\tau_\ell = C\ell^2/K$ from (2.1) is the time constant for heat to diffuse across the thickness ℓ of the slab, roughly speaking. For $\tau_\delta \ll t \ll \tau_\ell$, the heat has diffused out of the absorbing layer, but has not yet reached the edge of the slab at $z = \ell$. The effective volume V_{eff} which is being heated is determined roughly by the relation $V_{eff} = \mathcal{A}d$, where $\mathcal{A}$ is the surface area and d is the distance the heat is diffused. With $d = (Kt/C)^{1/2}$ from (2.2) and with $V = \mathcal{A}d$, (3.5) and (3.6) give

$$T \cong (IA/C\delta)(\tau_\delta t)^{1/2} \quad , \tag{3.7}$$

which has the correct functional form, $T \sim t^{1/2}$, and agrees in magnitude within a factor of 1.1 with the limit (A4c) of the surface temperature.

Finally, for $t >> \tau_\ell$, the heat has diffused throughout the sample. As a first rough approximation, the temperarature is the same as if the sample had absorbed the radiation uniformly throughout its volume, i.e., it is the same as if $S = IA/\ell$ for $0 < z < \ell$. Then (3.5) or (3.2) with $-K\partial^2 T/\partial z^2 = 0$ gives

$$T \cong IA\, t\theta(t)/C\ell \quad . \tag{3.8}$$

This result neglects the temperature variation across the sample.

In order to obtain a better approximation, try a solution of the form

$$T = \alpha t + f(x) \tag{3.9}$$

since it is expected that the shape $f(x)$ reaches a temperature-independent form for $t >> \tau_{2\ell}$. Substituting (3.9) into (3.2) and setting

$$-K\partial T/\partial x = 0 \qquad \text{at } z = \ell$$

$$= IA \qquad z = 0$$

for $\delta << \ell$ (so that the heat capacity of the region $0 < z < \delta$ is negligible and the heat flow at $x = \delta$ is equal to the absorbed incident intensity IA) gives

$$T = IA/C\ell\, t + IA(z - \ell)^2/2K\ell + G \quad .$$

The value of G is determined by requiring that the average energy

$$E = C \int_0^{\ell} dx\, T$$

be equal to the energy IAt added to the sample (per unit area). This gives

$$T \cong IA/C\ell t + (IA/2K\ell) \left[(z - \ell)^2 - \tfrac{1}{3} \ell^2 \right] \qquad (3.10)$$

for $t >> \tau_\ell$ and $\delta << \ell$. This result is obtained by another method in Appendix B.

The temperature difference $T_{0\ell} \equiv T(0) - T(\ell)$ across the sample (for $t >> \tau_{2\ell}$) is, from (3.10)

$$T_{0\ell} = IA\ell/2K \ , \qquad (3.11)$$

and the temperature at $z = 0$ for $t >> \tau_\ell$ is, from (3.10)

$$T(0,t) \cong IAt/C\ell + IA\ell/3K = (IA/C\ell)(t + \tfrac{1}{3}\tau_\ell) \ . \qquad (3.12)$$

The value of $T(0,t)$ lies above the linear term $IAt/C\ell$ by the constant amount $IA\ell/3K$. By energy conservation, the average of $T(z,t)$ over all z must be $IAt/C\ell$ according to (3.8). The constant term $IA\ell/3K$ arises because T is parabolic in z according to (3.10); thus $T(0,t)$ is larger than the average of $T(z,t)$ over z.

The result (3.12) joins smoothly onto the result (A4c) at $t \cong \tau_\ell/3$ as seen in Fig. F1. At $t = \tau_\ell/3$, (3.12) gives $T(0,t) = 2A\tau_\ell/3C\ell$ and (A4c) gives $T(0,t) = 2IA\tau_\ell/(3\pi)^{1/2}C\ell$. Since these two values differ by only 2%, the relation

$$T(0,t) \cong (2IA/C\ell)(\tau_\ell\, t/\pi)^{1/2} \qquad \text{for } \tau_\delta << t \leq \tau_\ell/3$$

$$\cong (IA/C\ell)\,(t + \tfrac{1}{3}\,\tau_\ell) \qquad \text{for } t \geq \tau_\ell/3 \qquad \underline{(3.13)}$$

is valid for $\delta << \ell$ and $t >> \tau_\delta$ with $\sim 2\%$ accuracy at the crossover at $t \cong \tau_\delta/3$ and with increasing accuracy away from the crossover.

The large-slope, linear region $t \lesssim \tau_\delta$ in the inset in Fig. F1 corresponds to the curves in Fig. F2, for which $T \cong$ constant for $z < \delta$ and $T \cong 0$ for $z > \delta$. The $T \sim t^{1/2}$ region ($\tau_\delta << t \leq \tau_\ell$) in Fig. F1 corresponds to the curves in Fig. F3 with $T \cong 0$ at $z = \ell(t/\tau_\ell \lesssim 0.1)$, and the linear region $t > \tau_\ell/4$ in Fig. F1 corresponds to the curves in Fig. F3 in which the parabolic spatial temperature distribution moves up linearly in time ($t/\tau_\ell \gtrsim 0.2$).

IV. MELTING

The problem of melting will be considered only very briefly. After the surface temperature reaches the melting temperature T_m, the calculation of the subsequent temperature distribution becomes difficult, as discussed in Carslaw and Jaeger,[9] Chapter 11. In general, the liquid-solid interface travels across the sample, and the temperature of the liquid rises above the melting temperature T_m. For the case of "total ablation," in which the liquid is removed instantly after melting, the "burn-through" time required to melt the total mass of the slab can be computed quite simply from energy considerations.[11] The energy required to raise the temperature of the volume $\mathcal{A}\ell$, where $\mathcal{A}$ is the surface area, from the initial temperature T_0 to the melting temperature T_m is $C\mathcal{A}\ell(T_m - T_0)$, and the energy required to melt this volume of material after the melting temperature is reached is $H\mathcal{A}\ell$, where H is the heat of fusion. The energy added is $IA\,\mathcal{A}t_m$ where t_m is the time required to raise the temperature from T_0 to T_m and to melt the sample. Equating these energies gives

$$t_m = (\ell/IA)\left[H + C(T_m - T_0)\right] . \qquad (4.1)$$

This result has been obtained previously[12] by a more complicated method. For cases in which the material is not removed as soon as it melts, the value of t_m is larger than the value given by (4.1) since additional energy must be supplied to raise the temperature of the liquid.

The more general total-ablation case in which the absorption coefficient has the value A_s before the surface melts and A_ℓ afterwards is easily solved. Using the same energy balance method gives

$$IA_s t_m + IA_\ell t_H = \ell [H + C(T_m - T_0)] \quad , \tag{4.2}$$

where t_m is the time required to bring the surface to the melting temperature, and $t_H + t_m$ is the total time required to melt the sample. The value of t_m is determined by setting $z = 0$ and $t = t_m$ in (A8) and solving for t_m, and the value of t_H is then given by (4.2). Limiting values of t_m will be given in (4.6).

The relative sizes of the two terms on the right-hand side of (4.1) or (4.2) can be illustrated by defining a temperature T_H by the relation

$$H \equiv C T_H \quad . \tag{4.3}$$

For aluminum, $C = 2.4\ J/cm^3\,K$ and $H = 10^3\ J/cm^3$; thus $T_H \cong 400\,K$. The same amount of energy is required to raise the temperature of aluminum 400 K (assuming that C is independent of temperature) as is required to melt it after the melting temperature is reached.

The time required to melt a thin antireflection or protective coating also is of interest. For a thin coating, the time $H\delta/IA$ required to melt the coating after T_m is reached is negligible. For example, for NaCl with $H = 10^3\ J/cm^3$, $AI = 10^3\ W/cm^2$, and $\delta = 1\,\mu m$, $t_m = 10^{-4}$ sec. Thus the time required to melt the coating is just t_m, which is easily obtained as discussed previously.

Explicit expressions for t_m can be obtained in three limiting cases. The values of the temperature $T_{\ell o}$ and time $t_{\ell o}$ at which the low-temperature linear region of $T(0,t)$ as a function of t crosses the $t^{1/2}$ curve in the inset of Fig. F1 are easily found by equating (A4b) to (A4c). This gives

$$T_{\ell o} = 4IA\delta/\pi K \quad , \qquad t_{\ell o} = 4\tau_\delta/\pi \quad . \tag{4.4}$$

In (3.13), the $t^{1/2}$ curve was joined onto the high temperature linear curve (3.12) at $t = t_{hi} \equiv \tau_\ell / 3$. The corresponding value T_{hi} of T is

$$T_{hi} = 2IA\tau_\ell / 3C\ell \quad , \qquad t_{hi} = \tau_\ell / 3 \quad . \tag{4.5}$$

For T_m (measured with respect to the initial temperature of the slab) much less than $T_{\ell o}$, T is given by (A4b); for $T_{\ell o} << T_m < T_{hi}$, T is given by (A4c); and for $T > T_{hi}$, T is given by (3.12). Setting $T = T_m$ and solving for t_m in the three cases gives

$$t_m \cong CT_m / IA \quad , \qquad \text{for } T_m << T_{\ell o} \tag{4.6a}$$

$$\cong \pi CKT_m^2 / 4A^2 I^2 \quad , \qquad \text{for } T_{\ell o} << T_m < T_{hi} \tag{4.6b}$$

$$\cong C\ell T_m / IA - \tau_\ell / 3 \quad , \qquad \text{for } T_m > T_{hi} \quad . \tag{4.6c}$$

As an example, for aluminum 1 cm thick with $I = 10^3$ W/cm^2, $A = 0.2$, $C = 2.4$ J/cm^3 K, $K = 2.4$ W/cm C, and $\tau_\ell = 0.29$ sec, the value of T_{hi} is 65 K. For an initial temperature of 0 K, $T_m \cong 500$ K; thus $T_m > T_{hi}$, and t_m is given by (4.6c), which gives

$$t_m = 6 \text{ sec} .$$

The intensity required to melt the surface before a substantial amount of heat has diffused out of the absorbing layer is determined by the inequality in (4.6a), which can be written as

$$IA >> \pi K T_m / 4\delta \quad . \tag{4.7}$$

For metals the required intensity is great because δ is small. For example, for aluminum with T_m = 500 C, A = 0.2, and $\delta = 10^{-2}\,\mu m$, (4.7) gives $I >> 5 \times 10^9$ W/cm^2. Even for the case of an antireflection coating with $\delta = 2\,\mu m$, A = 0.5, and the same values of K and T_m as those of aluminum, I still must be large ($I >> 10^7$ W/cm^2) in order to burn off the coating without heating a substantial part of the bulk of the window.

V. SMALL BEAM, $D \ll \ell$

In testing materials, small beam diameters are often used in order to obtain higher intensities. Then $D \ll \ell$ may be satisfied. For times sufficiently short that the thermal diffusion is short with respect to D, the value of T at the heated surface is given by the linear-diffusion result

$$T = 2IA(t/\pi CK)^{1/2} \qquad (5.1)$$

from (8.1b).

For $D < d < \ell$, the heat diffuses hemispherically. Since the heat has not reached the back surface, the temperature distribution is the same as in a semi-infinite medium. By symmetry, this is the same as that for an infinite medium. Thus, the infinite-medium temperature distribution for a point δ-function source[9]

$$T(r,t) = (Q/4\pi Kr)\ \mathrm{erfc}\ (Cr^2/4Kt)^{1/2} \qquad (5.2)$$

can be used. Here Q is the energy per unit time added and erfc is the complementary error function. This is of course just the Green's function, and the maximum temperature (at $\underset{\sim}{r} = 0$) is therefore

$$T = \int dQ(4\pi K\rho)^{-1}\ \mathrm{erfc}\ (C\rho^2/4Kt) \quad . \qquad (5.3)$$

With $dQ = IA\,2\pi\rho\,d\rho$, for a Gaussian beam $I = I_0 \exp(-2\rho/D_e)^2$, where D_e is the 1/e diameter, (5.3) gives

$$T = (I_0 A D_e/4K) \int_0^\infty du\, e^{-u^2}\ \mathrm{erfc}\ (CD_e^2 u^2/16Kt) \quad .$$

For $t >> CD_e^2/16K$, erfc is approximately equal to unity everywhere except in a very small circle near the origin. Thus, using $\int_0^\infty du \exp(-u^2) = \sqrt{\pi}/2$ gives

$$T = \sqrt{\pi}\, I_0 A D_e / 8K \qquad (5.4)$$

for this case of $D << d \lesssim 2\ell$.

A physical explanation of this result is not difficult to construct. Simply evaluating the point source result (5.2) at the edge of the beam (at $r = \frac{1}{2}D_e$) and using $Q = \int d\rho\, 2\pi\rho A I_0 \exp(-2\rho/D_e)^2 = \frac{1}{2}\pi A I_0 D_e^2$ gives $T = I_0 A D_e / 4K$, in agreement with (5.4) to within a factor of $2/\sqrt{\pi} - 1.1$.

Next consider the case of $d >> \ell$. The method of images[13] gives a good approximation. With images (5.2) at $\pm 2\ell$, $\pm 4\ell$, $\pm 6\ell$, $\cdots$, the differential equation and boundary and initial conditions are satisfied. Thus, with $Q = \frac{1}{2}\pi A I_0 D_e^2$ and with D_e/ℓ neglected in the arguments of the images, for $d >> D$

$$T = (\sqrt{\pi} I_0 A D_e / 8K) + \sum_{n=1}^{\infty} \frac{I_0 A D_e^2}{8nK\ell} \operatorname{erfc}\left(\frac{C\ell^2 n^2}{Kt}\right) . \qquad (5.5)$$

Since $\operatorname{erfc}(u) \cong 1$ for $u >> 1$, while $\operatorname{erfc}(u) << 1$ for $u << 1$, (5.4) can be approximated by

$$T = (\sqrt{\pi} I_0 A D_e / 8K)\left[1 + (D_e/\sqrt{\pi}\ell)\left(1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \tfrac{1}{m}\right)\right] ,$$

where $m = (Kt/C\ell^2)^{1/2} = (t/\tau_\ell)^{1/2}$ is the number of pairs of images contributing. [See the argument of the complementary error functions in the sum in (5.5).] Approximating the sum by $\ell n\, m + 0.6$ gives, for this case of $D << \ell << d$

$$T = (\sqrt{\pi} I_0 A D_e / 8K)\left\{1 + (D_e/\sqrt{\pi}\ell)\left[\ell n(Kt/C\ell^2) + 0.6\right]\right\} . \qquad (5.6)$$

The term containing the factor $D_e/\sqrt{\pi}\ell$, which is the term from the images, often is smaller than the first term, which is the direct source term. For example, for $D_e = 0.084$ cm, $\ell = 0.6$ cm, $K = 3.9$ W/cm K, $t = 20$ sec, and $C = 3.4$ J/cm^3 K as in the experiments of Saito, Charlton, and Loomis[14] on copper,

$$(D_e/\sqrt{\pi}\ell)\,[\,\ell n(Kt/C\ell^2) + 0.6\,] = 0.38 \ .$$

This small size of the image term indicates that the resulting temperature is not much greater than it would have been in a semi-infinite medium. Most of the temperature rise occurs in time $t = \tau_{D/2}$. Furthermore, experiments with time varied to determine the damage thresholds should be avoided in this case of $\frac{1}{2}D \ll \ell$ and $t \gg \tau_\ell$ since t appears only in the small term, and is in the argument of a logarithm. Changing the beam diameter D_e or the intensity I_0 would be effective since D_e and I_0 appear linearly in the large term.

The result (5.6) has the following physical interpretation. At time $t = \tau_\ell$, the heat has diffused out in a hemisphere of radius ℓ which touches the back face of the sample. At that time, $T = \sqrt{\pi} I_0 A D_e/8K$, according to (5.4). The subsequent diffusion is radially outward, as a cylinder of expanding radius. The additional temperature rise can be approximated by using the result[9] for the case of heat added at a uniform rate q per unit length inside of a cylinder of radius ρ_0:

$$T(\rho) = -(qC/4\pi K)\,Ei(-C\rho^2/4Kt) \ , \qquad (5.7)$$

where Ei is the exponential integral[15] and the radius ρ satisfies $\rho \geq \rho_0$. With $q = I_0 A(\pi D_e^2/2)/\ell$ and $\rho = \ell$ in the limit $t \gg C\rho^2/4K$ (then $Ei(-u) \cong \gamma + \ell n\, u$ for small u, where $\gamma = 0.5772$), (5.7) gives

$$T = (I_0 A D_e^2/8K\ell)\,[\,\ell n(Kt/C\ell^2) + 0.8\,] \ . \qquad (5.8)$$

In (5.8) we used $\ell n\, 4 - \gamma = 0.8$. This result (5.8) is the same as the image term in (5.6) except for the negligible difference between the terms 0.6 and 0.8. The simple model works quite well.

These results (5.1), (5.4), and (5.6) are schematically illustrated in Fig. F4a. For short-pulse operation, it is useful to consider the fluence F at which the failure temperature T_f is reached. Solving each equation (5.1), (5.4), and (5.6) for F and sketching the results gives the curve in Fig. F4b.

VI. COOLED SAMPLES, STEADY-STATE SOLUTIONS

Next consider the effect of cooling the back side of a laser-irradiated slab. For times shorter than the thermal diffusion time $\tau_{2\ell} = C4\ell^2/K$, the cooling has negligible effect because the heat has not diffused to the cooled back surface (and reacted back to the heated front surface). For example, for copper with K = 3.9 W/cm K and C = 3.4 J/cm^3 K, the cooling has little effect for $t < \tau_{2\ell}$, with $\tau_{2\ell}$ = 0.9 sec for ℓ = 0.5 cm or $\tau_{2\ell} = 9 \times 10^{-3}$ sec for ℓ = 1 mm. For nanosecond or microsecond pulses, cooling has little effect for samples of practical thickness.

For short times $t < \tau_{2\ell}$ the temperature is governed by the transient solutions of Secs. III and V, while for long times T approaches the steady-state value. For the large-diameter case ($D >> \ell$), the steady-state value T_{ss} of T is easy to determine. It is assumed that at t = 0, the sample and the coolant are at $T = T_c$, and that the coolant temperature remains at T_c. In the steady state, the heat flow Q is continuous. At the front surface the heat flow is

$$Q \equiv Q_0 = IA .$$

Across the sample,

$$Q = T_{0\ell} K/\ell ,$$

where $T_{0\ell} = T(0) - T(\ell)$, and at the back surface

$$Q = T_{\ell c} h ,$$

where $T_{\ell c} = T(\ell) - T_c$, with T_c the coolant temperature. Equating these values of Q and solving for $T_{ss} \equiv T_{0c} = T_{0\ell} + T_{\ell c}$ gives, for this case of $D >> \ell$,

$$T_{ss} = (IA/h)(1 + h/h_\ell) \quad , \qquad (6.1)$$

and $T_{0\ell} = IA\ell/K$ and $T_{\ell c} = IA/h$, where $h_\ell \equiv K/\ell$. The transient solution from Sec. III and the approach to the steady-state value in (6.1) are shown schematically in the solid curve in Fig. F5a for the thermally thin case ($h \ll h_\ell$). Notice that $T_{ss} = T_{transient}$ at $t = \tau_c$, which is not surprising since τ_c is the appropriate time constant for the thermally thin case. The results for the corresponding thermally thick case of $h > h_\ell$ are sketched in Fig. F5b.

An important feature of the result (6.1) is that the temperature T_{ss} is not decreased substantially by increasing the cooling efficiency (increasing h) once h is greater than h_ℓ. For $h \gg h_\ell$, (6.1) gives $T_{ss} = IA/h_\ell$, which is independent of h. This is because the temperature drop T_{0c} is then controlled by the drop $T_{0\ell}$ across the slab, the drop $T_{\ell c}$ at the interface being negligible.

For the small-diameter case of $D \ll \ell$ with cooling, the analysis is somewhat more complicated. Since there are no experimental results for this case, the following approximate analysis should serve as a guideline to the effects of changing material and experimental parameters and serve as a first approximation to the value of the temperature rise.

The heat diffusing from the heated spot of diameter $\pi D^2/4$ leaves the back surface and enters the coolant in an area at the back surface considerably greater than $\pi D^2/4$. As an approximation for the case of good cooling ($h \gg h_\ell$), it is assumed that the area $\pi(2\ell)^2/4$ is at temperature T_ℓ and that heat is transferred to the coolant at the rate $hT_{\ell c}\pi\ell^2$, where $T_{\ell c} = T_\ell - T_c$. Equating this to steady-state rate of heating $IA\pi D^2/4$ by the beam gives

$$T_{\ell c} \cong IAD^2/4\ell^2 h \quad . \qquad (6.2)$$

With the inequality $h >> h_\ell$, (6.2) gives $T_{\ell c} << IAD^2/4\ell K$. With $T = T_{0\ell} + T_{\ell c}$, where $T_{0\ell}$ is given by (5.4), and with $D/\ell << 1$, this gives

$$T \cong \sqrt{\pi} I_0 A D_e / 8K \qquad (6.3)$$

for this case of $h >> h_\ell$.

For the case of poor cooling ($h << h_\ell$), the temperature distribution approximately attains the cylindrically diffusing value (5.6) before the heating "takes effect." Thus, setting $t = \tau_c = \ell C/h$ in (5.6) gives

$$T \cong (\sqrt{\pi} I_0 A D_e / 8K) \left\{ 1 + (D_e/\sqrt{\pi}\ell \left[\ell n (h_\ell / h) + 0.6 \right] \right\} \qquad (6.4)$$

for this case of $h << h_\ell$. This result can be verified by showing that the integrated heat flow $IA\pi D^2/4$ at $z = 0$ is equal to that

$$h \int_0^\infty d\rho\, 2\pi\rho T = -h \int_0^\infty d\rho\, 2\pi\rho \left(\frac{IAD^2}{16\ell K} \right) Ei\left(-\frac{\rho^2 C}{4K\tau_c} \right)$$

at $z = \ell$. The integral is evaluated by changing the variable to $u \equiv \rho^2 C/4K\tau_c$ and using $-\int_0^\infty du\, Ei(-u) = 1$.

VII. ADDITIONAL APPLICATIONS

In addition to the examples in the previous sections, there are several other interesting applications. Consider the Hughes Research Laboratories infrared-mirror damage test.[8] Copper and molybdenum mirrors were irradiated with 10.6 μm radiation having 0.6 μsec nominal pulse duration. In applying the theoretical results to these experiments it is important to include the temperature dependence of the absorptance and consider the 0.2 μsec pulse peak in the nominal 0.6 μsec pulse.

From (2.2),

$$d = (Kt/C)^{1/2} = 8.3\,\mu m$$

for copper with $K = 3.9\,W/cm\,K$ and $C = 3.4\,J/cm^3\,K$. This thermal diffusion distance is much smaller than the beam diameter D and the sample thickness ℓ. Thus, the top equation in (3.13) is valid. Solving for the fluence $F = It$ gives

$$F = T(\pi CKt)^{1/2}/2A \quad . \qquad (7.1)$$

The temperature dependence of C and K is not very important ($\sim 10\%$ variation), but the pulse shape and the temperature dependence of A are important. The absorptance, which is proportional to the electrical resistivity, is sketched roughly in Fig. F6 for pure and impure copper. The value of A at the melting temperature $T_m = 1083$ C is approximately six times greater than the room temperature value (for pure copper). From the typical curves shown in Fig. F6 it is seen that the average value of A on the interval 300 K to T_m is not drastically different for pure and impure samples. A reasonable average value is one half the value at T_m or $A = 3A_{pure}$, where A_{pure} is the room temperature value of A for pure copper.

Next consider the pulse shape in these Hughes experiments, as sketched in Fig. F7. As a first approximation,[8] this pulse shape can be replaced by a pulse of height I_{mx} and width $0.6\,\mu$sec, as shown in the dashed line. A slight improvement is obtained by approximating the pulse by two pulses, the first of height I_{mx} and duration $t_{p1} = 0.2\,\mu$sec (or fluence $F_1 = \frac{1}{3}$ the total fluence F_{tot}) and the second of fluence $F_2 = 2F_1$ and duration $0.2\,\mu$sec. According to (7.1), $T \sim F/t_p^{1/2}$. Thus the first pulse gives a greater temperature rise, and the second pulse is neglected in the present rough estimate. Putting these results into (7.1) gives

$$F_{tot} \cong T(\pi C K t_{p1})^{1/2}/2A_{pure} \quad . \tag{7.2}$$

With $T = 1080 - 20 = 1060$ K, $C = 3.4$ J/cm^{3}K, $K = 3.9$ W/cm K, $t_{p1} = 0.2\,\mu$sec, and $A_{pure} = 4 \times 10^{-3}$, (7.2) gives

$$F_{tot} \simeq 400 \text{ J/cm}^2 \quad . \tag{7.3}$$

The experimental value of the intrinsic threshold is not known, but the theoretical estimate of 400 J/cm^2 appears to be within a factor of two of the experimental value. The highest fluence measured for polished mirrors was 125 J/cm^2, but a chemically etched surface, which was not a good reflector in the visible, withstood 750 J/cm^2. In any event, the theoretical result suggests that the intrinsic limit could well be the simple process of melting the mirror material by the usual absorption processes, in which case the currently measured values are near the intrinsic limit. In passing it is mentioned that the value of d in (7.1) corresponding to the shorted pulse time of $0.2\,\mu$sec is $5\,\mu$m.

Next consider the failure of the cavity mirrors in the first xenon lasers.[16] The following simple calculation shows that this failure is simply the melting of the aluminum substrate as a result of the large absorptance in the vacuum ultraviolet. The

mirrors were made by depositing approximately 100 A of aluminum onto magnesium fluoride substrates. The laser intensity is absorbed in this thin aluminum film, but the intensity required to melt the film is controlled by the diffusion of heat from the film onto the magnesium fluoride substrate. The volume of material heated, most of which is magnesium fluoride, is equal to the product of the surface area illuminated times the effective thickness d, where d is the distance into the magnesium fluoride that the heat diffuses. The value of d for a pulse duration of 10 nsec is, from (2.2),

$$d = (Kt/C)^{1/2} = 0.2\,\mu m \quad ,$$

where $K \cong 0.2\,W/cm\,K$ and $C = 3.14\,J/cm^2\,K$ for MgF_2. Thus d is much less than both the beam diameter $D \cong 1\,cm$ and sample (MgF_2) thickness $\ell \cong \frac{1}{2}\,cm$, and the linear heat-flow result (5.1) applies. As discussed in Sec. I, for high-power optical components, a reasonable criterion for failure is the melting of a negligibly thin layer at the surface. Thus, with $T = T_m = 640\,K$ (above room temperature ambient) and $A = 0.2$, (5.1) gives

$$I_f = (T_m/2A)(\pi CK/t)^{1/2} = 20\ MW/cm^2 \quad ,$$

or a fluence of $0.2\ J/cm^2$.

The value of $A = 0.2$ is the value quoted by the manufacturer (Acton) of the mirrors. Since the source of this absorptance at 7.2 eV is unknown and well above the intrinsic value[17] of 0.08, the temperature dependence of A is unknown. Thus the published room-temperature value of 0.2 was used. The agreement of the theoretical result of $I_f = 20\ MW/cm^2$ with the experimentally estimated value of $40\ MW/cm^2$ is within the accuracies of the experimental value and the theoretical value (with the large uncertainty in the value of A).

Finally, consider the experiments of Saito, Charlton, and Loomis[14] in which $D_e = 0.084$ cm, $\ell = 0.6$ cm, $t = 20$ sec, and $I_0 = 10^5$ W/cm^2 for damage to copper. Since $D << \ell$ and $d = (Kt/C)^{1/2} = 4.8\,\text{cm} >> \ell$, T is given by (5.6). With $K = 3.9$ W/cm K, $C = 3.4$ J/cm^3 K, and $A = 0.01$ for copper, (5.6) gives

$$T = 6.6\ \text{K}\ ,$$

which is a factor of 160 too low to melt copper (melting temperature of $T_m = 1080 - 20 \cong 1060$ K above ambient). The source of this great discrepancy between the experimental and theoretical results is not known. The low experimental threshold may be related to surface-ignited plasmas,[5,6] possibly the initial stages of this phenomena. The power of approximately 1 kW in these experiments is at the measured threshold of 1.4 - 2 kW for laser ignition of plasmas at solid surfaces for $D < 1$ mm.

The fact that dust on the windows in the experiments was observed to reduce the damage threshold to values below 100 W/cm^2, compared with the above experimental value of 10^5 W/cm^2 and the theoretical value of 10^7 W/cm^2 for laser breakdown in dirty air, suggests that absorption associated with surface contamination may be involved. The characteristic dust-damage patterns observed at the 100 W/cm^2 damage level were not observed in the 10^5 W/cm^2 damage patterns. This lack of the dust-burn pattern is consistent with a plasma phenomenon.

Before considering the plasma ignition further, it should be mentioned that it is unlikely that hot spots in the beam are responsible for the low threshold. For example, from (5.6) with the image term neglected, if the intensity were 100 times greater than I_0 in an area 100 times smaller than $\frac{1}{2}\pi D_e^2$, then T would be a factor of $100/\sqrt{100} = 10$ times greater, or $T = 66$ K. The fact that molybdenum did

not damage, whereas copper did, now has little significance until source of the discrepancy between experiment and theory is resolved.

The present theoretical results have important consequences in the problem of explaining the laser ignition of plasmas at solid surfaces. First, ordinary heating of the solid by the laser beam is not sufficient to ignite the plasma, since the temperature of 6.6 K obtained above is orders of magnitude too small for ignition. On the other hand, a thermally isolated imperfection on the surface of the solid could cause ignition. The temperature rise T of such an imperfection is given by $CV = \Delta E / T$, from (3.5), with $\Delta E = I \sigma_{rel} \pi a^2 t$, where a is the radius of the imperfection (assumed to be spherical for simplicity) whose absorption cross section is σ_{rel} times the geometrical cross section πa^2. Combining these two results gives

$$T = 3 \sigma_{rel} I t / 4 C a \quad . \qquad (7.4)$$

For $I = 10^5 \, W/cm^2$, $t = 20$ sec, $C \cong 3 \, J/cm\,K$, $a = 10 \, \mu m$, and $\sigma_{rel} \cong 1$, (7.4) gives $T = 10^9$ K. This extremely great temperature rise suggests that thermally isolated or partially thermally isolated surface imperfections should be quite effective in igniting plasmas. Dr. Craig Walters has suggested and studied the mechanism of plasma ignition by thermally isolated regions of aluminum on rolled aluminum samples by shorter pulses (40 nsec rise time) of $\sim 10^8 \, W/cm^2$ intensity, and it appears that this is the controlling mechanism in his case. It should be mentioned that a region of the sample that is thermally isolated for nanosecond pulses may not be thermally isolated for microsecond pulses or for essentially cw operation since there may be thermal-diffusion channels whose characteristic times are greater than nanoseconds but less than microseconds.

Finally for picosecond pulses, say 10^{-11} sec, with $K/C \cong 1$, (2.2) gives $d \cong 3 \times 10^{-6}$ cm. This small thermal diffusion distance indicates that for picosecond pulses great temperature rises do not require thermally isolated regions since the heated volume already is small.

VIII. SUMMARY AND DISCUSSION

The results derived for the temperature are summarized for the convenience of the reader:

For $\ell \ll D$:

$$T = IAt/C\delta\,, \qquad \text{for } t \ll \tau_\delta \qquad (8.1a)$$

$$= 2IA(t/\pi CK)^{1/2}\,, \qquad \text{for } \tau_\delta \ll t \leq \tfrac{1}{3} t_\ell \qquad (8.1b)$$

$$= (IA/C\ell)(t + \tfrac{1}{3}\tau_\ell)\,, \qquad \text{for } \tfrac{1}{3} t_\ell \leq t \qquad (8.1c)$$

$$= IA(1/h + 1/h_\ell)\,, \qquad \text{for } t \to \infty, \text{ cooled}. \qquad (8.1d)$$

$$T(z) = (IA/C\ell)\left\{ t + (C/2K)\left[(z-\ell)^2 - \tfrac{1}{3}\ell^2 \right] \right\}\,, \quad \text{for } \tau_\ell \ll t \qquad (8.2)$$

$$T_{0\ell} \equiv T(0,t) - T(\ell,t) = IA\ell/2K\,, \qquad \text{for } \tau_\ell \ll t\,. \qquad (8.3)$$

For $D \ll \ell$:

$$T = 2IA(t/\pi CK)^{1/2}\,, \qquad \text{for } t \ll \tau_D \qquad (8.4a)$$

$$T = \sqrt{\pi} I_0 A D_e/8K\,, \qquad \text{for } \tau_D \ll t \ll \tau_\ell \qquad (8.4b)$$

$$T = \frac{\sqrt{\pi} I_0 A D_e}{8K}\left[1 + \frac{D}{\sqrt{\pi}\ell}\left(\ell n \frac{Kt}{C\ell^2} + 0.6 \right) \right]\,, \quad \text{for } \tau_\ell \ll t \qquad (8.4c)$$

$$T = \sqrt{\pi} I_0 A D_e/8K\,, \qquad \text{for } t \to \infty, \text{ well cooled} \qquad (8.4d)$$

$$T = \frac{\sqrt{\pi} I_0 A D_e}{8K}\left\{ 1 + \frac{D_e}{\sqrt{\pi}\ell}\left[\ell n\, h_\ell/h + 0.6 \right] \right\} \quad \text{for } t \to \infty, \text{ poorly cooled} \qquad (8.4e)$$

In interpreting results, the following time constant, distances, and heat transfer coefficients are useful:

$$\tau_\ell = C\ell^2/K \ , \qquad \tau_s = CK/h^2 \ , \qquad \tau_c = C\ell/h = (\tau_\ell \tau_s)^{1/2} \ ,$$

$$d = (Kt/C)^{1/2} \ , \qquad L_s = (K\tau_s/C)^{1/2} \ , \qquad L_c = (K\tau_c/C)^{1/2} \ ,$$

$$h_\ell = K/\ell \ .$$

These temperature-rise results include both the transient and steady-state temperatures. The approach to the steady state when the laser is tuned on follows the transient result for short times and approaches the steady-state limit for long times, of course, and the time at which the transient curve crosses the constant steady-state value is a measure of the system time constant τ_{ss}. For $t \ll \tau_{ss}$, the cooling is not effective in keeping the temperature low. The value of τ_{ss} cannot be smaller than $\sim \tau_{2\ell}$ since the heat has not then diffused to the cooled surface and back to the surface at which the temperature T is measured, roughly speaking. (This result is obvious also from an examination of image solutions).

Another important and interesting cooling effect is that increasing the cooling efficiency above the critical value $h_\ell = K/\ell$ does not result in a farther decrease in T. For $h > h_\ell$, the temperature is controlled by the temperature drop $T_{0\ell}$ across the sample, the drop $T_{\ell c}$ at the coolant interface being negligible.

Damage thresholds can be obtained from these results (8.1) to (8.4) by solving the appropriate equations for the intensity I or fluence F. The numerical results and the dependence on material parameters such as K and C are different, depending on the relative size of the pulse duration t and the characteristic times τ_δ, τ_ℓ, and τ_D and on the size of the laser-beam diameter D with respect to the sample thickness ℓ.

Application of the results to several recent experiments has interesting results: The mirror-damage thresholds measured[8] with 10.6 μm pulses of duration 0.6 μm on copper and molybdenum with $D >> \ell$ are near or possibly at the intrinsic limits set by heating. Thus, nonlinear effects need not be invoked in order to explain the results, as has been thought previously.

The experimental damage intensities[14] for copper by twenty-second pulses of 10.6 μm radiation with $D << \ell$ are a factor of 160 lower than the theoretical value. Several possible explanations of this large discrepancy were examined. All were ruled out by simple arguments, with the possible exception of a phenomenon related to plasma ignition. The evidence for such a phenomenon is inconclusive at best.

The laser ignition of plasmas at solid surfaces by cw radiation or long pulses is not understood. The present theoretical results suggest that the temperature rise caused by ordinary heating of the solid by the laser beam is orders of magnitude too small to ignite the plasma, but that thermally isolated surface imperfections could cause ignition.

In another application, the model explains the failure of the aluminum-coated magnesium fluoride mirrors in the recently developed vuv xenon lasers operating at 7.2 eV. In a repeated-pulse system having many pulses per second, the heating effects are essentially the same as those of a cw system having the same average intensity unless the damage occurs in the first pulse. In this latter case, the single-pulse results apply, of course. It should be kept in mind that the increase in intensity resulting from reducing the pulse repetition rate and maintaining the average power constant can result in failure from nonlinear or other explicit high-power effects.

IX. ACKNOWLEDGMENTS

Dr. C. Walters and Dr. D. Smith very kindly discussed their results on laser ignition of plasmas. Appreciation is expressed to Dr. P. Hoff for several discussions of the Livermore ultraviolet-laser program.

APPENDIX A. DERIVATION OF EXACT SOLUTION FOR $D \ll \ell$

First consider the problem of a semi-infinite medium heated by a source that is turned on at time $t = 0$ and has a constant value IA/δ for $0 < z < \delta$ and vanishes for $x > \delta$. The surface of the medium is at $z = 0$. The initial temperature is zero everywhere, and there is no heat flow at $z = 0$. The differential equation is given in (3.2a) and (3.1), and the boundary and initial conditions are

$$T(z,0) = 0 \quad , \qquad \partial T/\partial z = 0 \qquad \text{at } z = 0 \text{ for all } t \tag{A1a}$$

$$\lim_{z \to \infty} T = 0 \quad , \tag{A1b}$$

where T and dT/dz are continuous at $z = \delta$. Taking the Laplace transform of (3.2a), solving the simple differential equation in z, and inverting the Laplace transform gives[18]

$$T_\infty = T_< \equiv \frac{S_0 t}{C} \left| 1 - 2\,i^2\,\mathrm{erfc}\left(\frac{\tau_\delta - z}{4t}\right)^{1/2} - 2\,i^2\,\mathrm{erfc}\left(\frac{\tau_\delta + z}{4t}\right)^{1/2} \right| , \tag{A2a}$$

$$\text{for } 0 < z < \delta$$

$$T_\infty = T_> \equiv \frac{2S_0 t}{C} \left| i^2\,\mathrm{erfc}\left(\frac{\tau_z - \delta}{4t}\right)^{1/2} - i^2\,\mathrm{erfc}\left(\frac{\tau_z + \delta}{4t}\right)^{1/2} \right| \tag{A2b}$$

$$\text{for } z > \delta \quad ,$$

where $i^2\mathrm{erfc}$ is the second integral of the complementary error function, the subscript ∞ denotes that the solution (A2) is for a semi-infinite medium, and

$$\tau_z^{1/2} \equiv (C/K)^{1/2}\, z \quad . \tag{A3}$$

The maximum value of T, at z = 0, is

$$T_\infty(0,t) = \frac{IAt}{C\delta}\left[1 - 4i^2\mathrm{erfc}\,(\tau_\delta/4t)^{1/2}\right] \qquad \text{(A4a)}$$

$$\cong IAt/C\delta\ , \qquad \text{for } t \ll \tau_\delta \qquad \text{(A4b)}$$

$$\cong (2IA/C\delta)(t\tau_\delta/\pi)^{1/2} = 2IA(t/\pi CK)^{1/2}, \quad \text{for } t \gg \tau_\delta\ . \qquad \text{(A4c)}$$

This is the solution for a semi-infinite medium. For a medium of finite thickness ℓ with

$$\partial T/\partial z = 0\ , \qquad \text{at } z = 0, \ell \qquad \text{(A5)}$$

$$T(z,0) = 0\ , \qquad \text{(A6)}$$

the solution can be obtained by the method of images as follows: The infinite-medium solution $T_>(z)$ in (A2b) still satisfies the differential equation (A1) when displaced by an arbitrary distance d and/or reflected in the z = 0 plane; that is, $T_>(z - d)$ and $T_>(-z)$ are solutions to (3.2a). The function $T_\infty(z) + T_>(-z + 2\ell)$ satisfies the differential equation and the boundary condition (A5) at $z = \ell$. But it does not satisfy the boundary condition (A5) at z = 0 because of the "backward-propagating" image term $T_>(-z + 2\ell)$, which corresponds to the solution of the problem of a semi-infinite medium extending from $z = 2\ell$ to $z = -\infty$ with $S = IA/\delta$ for $2\ell - \delta < z < 2\ell$ and S = 0 otherwise, $\partial T/\partial z = 0$ at $z = 2\ell$, T(z,0) = 0, and $T \to 0$ as $z \to -\infty$. Adding a "forward-propagating" image at $z = -2\ell$, $T_>(z + 2\ell)$, gives

$$T_\infty(z) + T_>(-z + 2\ell) + T_>(z + 2\ell) \qquad \text{(A7)}$$

which satisfies the boundary condition (A5) at $z = 0$, but not at $z = \ell$. Continuing this procedure generates the solution as an infinite sum of images $T = T_\infty + T_\Sigma$, which can be written as

$$\begin{aligned} T(z,t) &= T_<(z,t) + T_\Sigma(z,t) && \text{for } 0 < z < \delta \\ &= T_>(z,t) + T_\Sigma(z,t) && \text{for } z > \delta \ , \end{aligned} \tag{A8}$$

where

$$T_\Sigma = (2\Lambda t/C\delta) \sum_{p=0}^{1} \sum_{n=1}^{\infty} \left\{ i^2 \operatorname{erfc} \Lambda \left[(-1)^p z + 2n\ell - \delta \right] - i^2 \operatorname{erfc} \Lambda \left[(-1)^p z + 2n\ell + \delta \right] \right\} , \tag{A9}$$

with $\Lambda \equiv (C/4Kt)^{1/2}$. Equations (A7) and (A9) are related by the expression

$$T_>(\pm z + 2n\ell) = (2S_0 t/C) \left[i^2 \operatorname{erfc} \Lambda (\pm z + 2n\ell - \delta) - i^2 \operatorname{erfc} \Lambda (\pm z + 2n\ell + \delta) \right] . \tag{A10}$$

The behavior of T as a function of z and t is easily obtained from the image interpretation. For $t \ll \tau_\ell$ the image terms are small; thus

$$T \cong T_\infty \qquad \text{for } t \ll \tau_\ell \ .$$

For $t \ll \tau_{2\ell}$, T is approximately equal to the sum of the T_∞ term and the first image:

$$T = T_\infty + T_>(-z + 2\ell) \ , \qquad \text{for } t \ll \tau_{2\ell} \ .$$

Similarly, for $t \ll \tau_{m\ell}$, only m images are required. Mathematically, the convergence of (A9) is very good because

$$i^n \operatorname{erfc} \xi \approx \frac{1}{\xi^{n+1}} e^{-\xi^2}, \qquad \text{for } \xi \gg 1 . \tag{A11}$$

The temperature $T(0,t)$ at $z = 0$ is obtained by setting $z = 0$ in (A8):

$$T(0,t) = \frac{IAt}{C\delta} \left\{ 1 - 4i^2 \operatorname{erfc} \frac{\delta}{\ell} + 4 \sum_{n=1}^{\infty} \left[i^2 \operatorname{erfc} \frac{(2n\ell - \delta)}{\ell} - i^2 \operatorname{erfc} \frac{(2n\ell + \delta)}{\ell} \right] \right\} . \tag{A12}$$

In the limits $t \lesssim \tau_\delta / 8$ and $\delta \ll \ell$, all terms in (A8) are negligible except the following:

$$\tilde{T} = \tilde{t} \left\{ 1 - 2i^2 \operatorname{erfc} \left[\tilde{t}^{-1/2} (1 - \tilde{z}) \right] \right\} \qquad \text{for } 0 < \tilde{z} < 1$$

$$= \tilde{t}\, 2i^2 \operatorname{erfc} \left[\tilde{t}^{-1/2} (\tilde{z} - 1) \right] \qquad \text{for } \tilde{z} > 1 , \tag{A13}$$

where $\tilde{T} \equiv C\ell / IA$, $\tilde{t} \equiv t / \tau_\delta$, and $\tilde{z} \equiv z / \ell$. This result (A13) is plotted in Fig. F2.

APPENDIX B. USEFUL APPROXIMATE SOLUTION FOR $\delta \ll \ell \ll D$

An accurate approximate solution for the case of $\delta \ll \ell$ can be obtained as follows: For $t \lesssim \tau_\ell / 2\pi^2$, the back surface at $z = \ell$ has no effect on the temperature distribution since the heat has not had time to diffuse to this surface. Thus, the solution T_∞ in (A2) for a semi-infinite film is quite accurate for $t \lesssim \tau_\ell / 2\pi^2$ in the slab of thickness ℓ.

For times $t \gg \tau_\delta$, the heat has diffused well out of the thin absorbing layer. Also, for $t \gg \tau_\delta$, the temperature drop across the layer is negligible, as easily seen from (A2a), which gives

$$\frac{T_\infty(0,t) - T_\infty(\delta,t)}{T_\infty(0,t)} = \frac{\frac{1}{2} + 2\,i^2\,\mathrm{erfc}\,2\eta_\delta - 4\,i^2\,\mathrm{erfc}\,\eta_\delta}{1 - 4\,i^2\,\mathrm{erfc}\,\eta_\delta}$$

$$\cong \frac{1}{2}\,, \qquad \text{for } t \ll \tau_\delta$$

$$\cong \theta(\eta_\delta) \ll 1\,, \qquad \text{for } t \gg \tau_\delta\,, \tag{B1}$$

where $\eta_\delta = (\tau_\delta / 4t)^{1/2}$. Thus the source S can be replaced by a constant heat flux 1A into the slab at the surface at $z = 0$. Clearly, if $\delta \to 0$ and $S \to \infty$ in such a way that $S\delta$ remains equal to a constant $(S\delta)_0$, then $S = (S\delta)_0\,\delta(x)$, where $\delta(x)$ is the Dirac delta function. From the theory of Green's function solutions to boundary value problems, it is well known that a delta-function source at a boundary is equivalent to an inhomogeneous boundary condition -- in the present case, a constant heat flow at the surface. Thus, the solutions

$$T_{\delta\to 0} = 2IA\left(\frac{t}{KC}\right)^{1/2}\sum_{n=0}^{\infty}\left\{ i\,\mathrm{erfc}\left[\left(\frac{\tau_\ell}{4t}\right)^{1/2}(2n+\frac{z}{\ell})\right] + i\,\mathrm{erfc}\left[\left(\frac{\tau_\ell}{4t}\right)^{1/2}(2nt2-\frac{z}{\ell})\right]\right\}$$

$$= \frac{IAt}{C\ell} + \frac{IA\ell}{K}\left\{\frac{3\zeta^2-\ell^2}{6\ell^2} - \frac{2}{\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\,e^{-t/\tau_n}\cos\frac{n\pi\zeta}{\ell}\right\} \qquad (B2)$$

where $\zeta \equiv \ell - z$, $\tau_n \equiv C\ell^2/n^2\pi^2K$, and $IA = (S\delta)_0$, to the heat-flow equation (3.2a) with the boundary and initial condition

$$-K\frac{\partial T}{\partial z} = IA \quad , \qquad \text{at } z = 0 \quad ,$$

$$\frac{\partial T}{\partial z} = 0 \quad , \qquad \text{at } z = \ell \quad , \qquad (B3)$$

$$T(z,0) = 0 \quad ,$$

are accurate approximations to the solution of (3.1) and (3.2) for

$$t >> \tau_\delta \qquad \delta << \ell \quad . \qquad (B4)$$

It is easy to show directly that (A9) reduces to (B2) when the inequalities (B4) are satisfied.

Summarizing, for $\delta << \ell$, T can be approximated by

$$T \simeq T_\infty \quad , \qquad \text{for } t < \tau_\ell/2\pi^2$$

$$\simeq T_{\delta\to 0} \quad , \qquad \text{for } t >> \tau_\delta/4 \quad , \qquad (B5)$$

where T_∞ and $T_{\delta\to 0}$ are defined in (A2) and (B2), respectively. In the region $\tau_\delta << t < \tau_\ell/2\pi^2$, both (A2) and (B2) are accurate approximations.

The result (3.10) derived in Sec. III can be obtained from (B2). For $t >> \tau_\ell$, all the terms in the sum on n in (B2) are negligible; thus, with $\zeta = z - \ell$, (B2) reduces to (3.10).

This method of approximate solution affords the advantage that it is easily extended to sources S other than the spatial step function in (3.1). For example

$$S = S_0 \exp(-\alpha z) \tag{B5}$$

is more realistic than the step function for metals and for homogeneous materials with $\beta\ell >> 1$. In a semi-infinite medium ($z > 0$) originally at $T = 0$ with no heat flow at $z = 0$ and with the heat source (B5),[9]

$$T_\infty = \frac{2S_0}{\alpha K}(\kappa t)^{1/2}\, i\,\mathrm{erfc}\,\frac{z}{2(\kappa t)^{1/2}} - \frac{S_0}{\alpha^2 K}\, e^{-\alpha z}$$

$$+ \frac{S_0}{2\alpha^2 K}\, e^{\alpha^2 \kappa t - \alpha z}\,\mathrm{erfc}\left|\alpha(\kappa t)^{1/2} - \frac{z}{2(\kappa t)^{1/2}}\right|$$

$$+ \frac{S_0}{2\alpha^2 K}\, e^{\alpha^2 \kappa t + \alpha z}\,\mathrm{erfc}\left|\alpha(\kappa t)^{1/2} + \frac{z}{2(\kappa t)^{1/2}}\right| . \tag{B6}$$

where $\kappa = K/C$. For a slab of finite thickness $\ell >> \delta \equiv 1/\alpha$, an accurate approximation to T is given by (B6) for $t < \tau_\ell / 2\pi^2$ and by (B2) for $t >> \tau_\delta$. An image solution could be generated from (B6) just as the image solution was generated from (A2) in Appendix A.

REFERENCES

*This section will be submitted for publication in the Journal of Applied Physics.

1. M. Sparks and H. C. Chow, J. Appl. Phys. 45, 1510 (1974); M. Sparks and H. C. Chow, Third Conference on High Power Infrared Laser Window Materials, C. A. Pitha, H. Posen, and A. Armington, editors, AFCRL-TR-74-0085, Vol. III, 1083 (Air Force Cambridge Research Laboratories, Bedford, Mass., 1974); M. Sparks and M. Cottis, J. Appl. Phys. 44, 787 (1973).
2. M. Sparks, J. Appl. Phys. 42, 5029 (1971).
3. Henceforth, pulsed will mean this one-second type of operation, while single pulse will mean a single short pulse, with typical duration of nanoseconds to microseconds.
4. M. Sparks, J. Appl. Phys. 44, 4137 (1973).
5. D. C. Smith and M. C. Fowler, Appl. Phys. Lett. 22, 500 (1973).
6. C. T. Walters and R. H. Barnes, Battelle Columbus Laboratories Semiannual Report, Contract Number DAAH01-73-C-0776, June (1974).
7. L. V. Keldysh, Zh. Eksp. Teor. Fiz. 47, 1945 (1964) [Sov. Phys. - JETP 20, 1307 (1965)]; G. M. Zverev, T. N. Mikhailova, V. A. Pashkov, and N. M. Solov'eva, Zh. Eksp. Teor. Fiz. 53, 1849 (1967) [Sov. Phys. - JETP 26, 1053 (1968)]; D. W. Fradin, E. Yablonovitch, and M. Bass, Appl. Opt. (to be published).
8. V. Wang, A. I. Braunstein, M. Braunstein, and J. Y. Wada, Laser Induced Damage in Optical Materials, NBS Special Publication 372, 183 (U.S. Department of Commerce, 1972).
9. H. S. Carslaw and J. C. Jaeger, Conduction of Heat in Solids, 2nd Edition (Oxford-Clarendon, 1959).

10. M. Sparks, Parke Mathematical Laboratories, Inc. Reports No. TM-1 and No. TM-2, 1971 (unpublished).

11. R. Hindley (private communication) and others have realized that simple energy-balance considerations can be used to obtain the burn-through time.

12. J. E. Rogerson and G. A. Chayt, J. Appl. Phys. 42, 2711 (1971).

13. The problem solved by images in Appendix A illustrates the method.

14. T. T. Saito, G. B. Charlton, and J. S. Loomis, Laser Induced Damage in Optical Materials, NBS Special Publication, 1974 (to be published)

15. Handbook of Mathematical Functions, edited by M. Abramovitz and I. A. Stegun, National Bureau of Standards, Applied Mathematics Series No. 55 (U.S. GPO, Washington, D.C., 1964).

16. N. Basov, V. Denilychev, and Y. Popov, Sov. J. Quantum Electron. 1, 18 (1971); P. W. Hoff, J. C. Swingle, and C. K. Rhodes, Appl. Phys. Lett. 23, 245 (1973); J. B. Gerardo and A. W. Johnson, J. Appl. Phys. 44, 4120 (1973); W. M. Hughes, J. Shannon, A. Kolb, E. Ault, and M. Bhaumik, Appl. Phys. Lett. 23, 385 (1973).

17. H. C. Chow and M. Sparks, J. Appl. Phys. (to be published); M. Sparks and C. J. Duthler, Xonics, Incorporated Third Technical Report, Contract No. DAHC15-73-C-0127, June (1974).

18. This result (A2) is quoted in Reference 9.

Figure Captions

Fig. F1. Time dependence of the temperature $T(0,t)$ at the illuminated surface (hottest plane). The inset shows the small-time behavior of $T(0,t)$.

Fig. F2. Spatial dependence of the normalized temperature $KT/I\Lambda\delta$ in and near the absorbing layer showing the near-step-function profile for $t/\tau_\delta \ll 1$; the broadening of the region of rapid temperature change as t/τ_δ increases; and the near-constant temperature in the absorbing layer ($0 < z < 1$) where $z \equiv z/\delta$, for $t \gg \tau_\ell$.

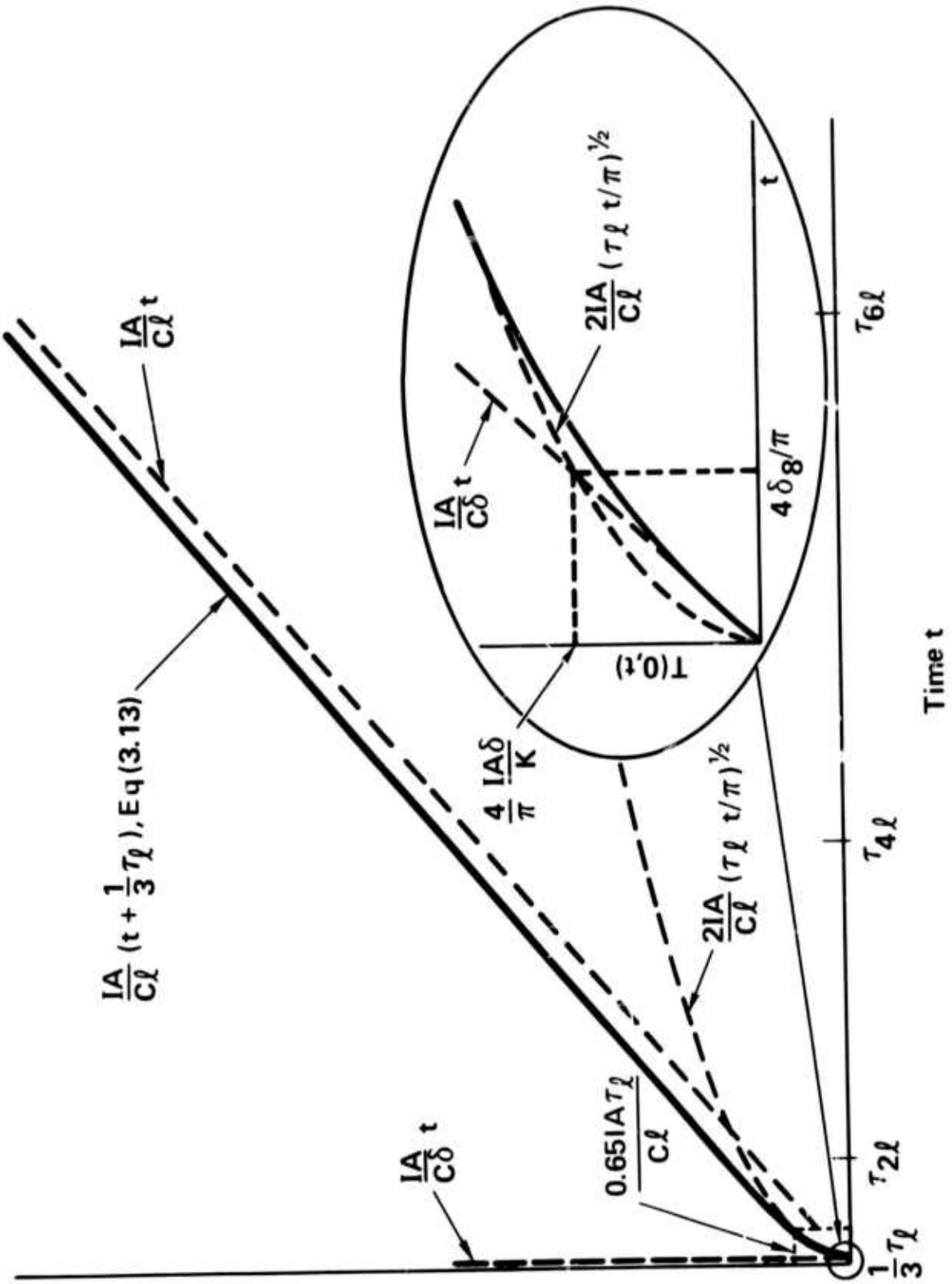
Fig. F3. Spatial dependence of the temperature for the case of $\delta \ll \ell$ and $t \gg \frac{1}{4}\tau_\delta$ showing: the isolation of the temperature change near the illuminated surface ($z = 0$) for $t \ll \frac{1}{4}\tau_\ell$ (the curve marked $t/\tau_\ell = 0.01$, for example); and the approach to a parabolic distribution as t/τ_ℓ increases. For $t/\tau_\ell = 0.30$ (top curve) the distribution is very nearly parabolic.

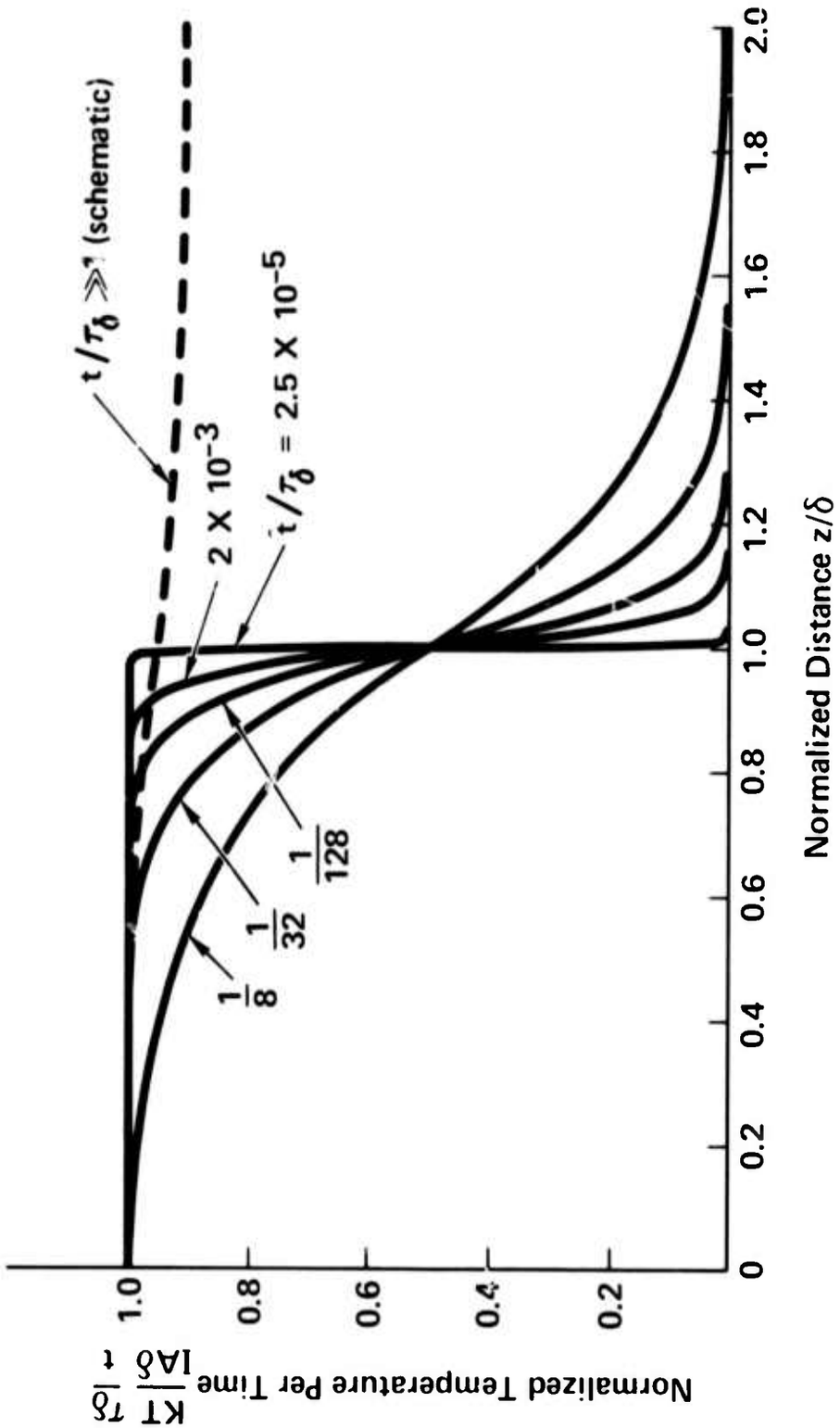
Fig. F4. (a) Temperature rise for the case of I fixed and t varied. (b) Fluence for the case of T fixed and t varied. In both cases, $D \ll \ell$.

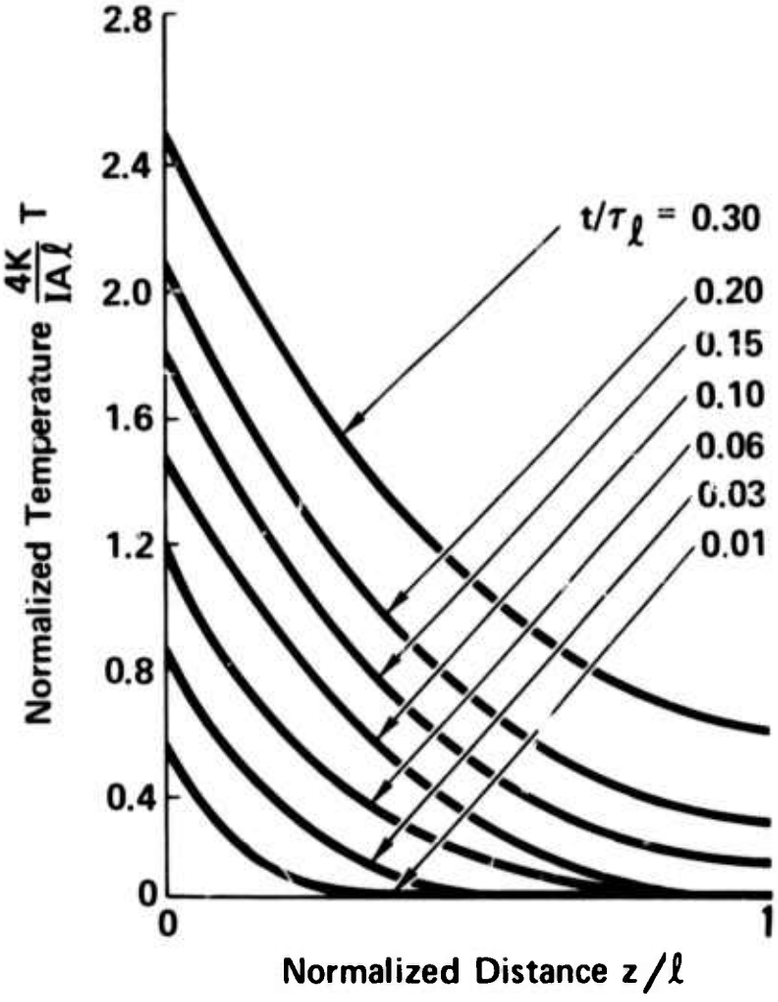
Fig. F5. (a) Temperature rise for the case of I fixed and t varied. (b) Fluence for the case of T fixed and t varied. In both cases, $\ell \ll D$ and $h \gg h_\ell$.

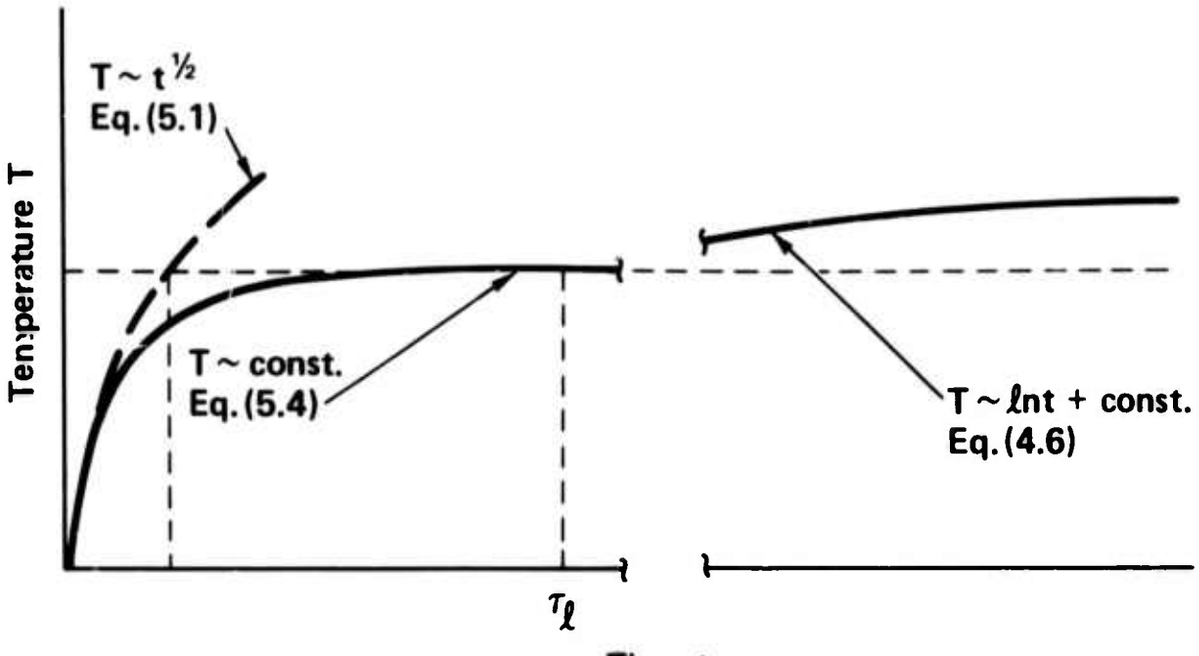
Fig. F6. Rough sketch of the absorptance A of copper ($A \sim \rho$).

Fig. F7. Sketch of pulse shape in the Hughes experiments (Ref. 8).

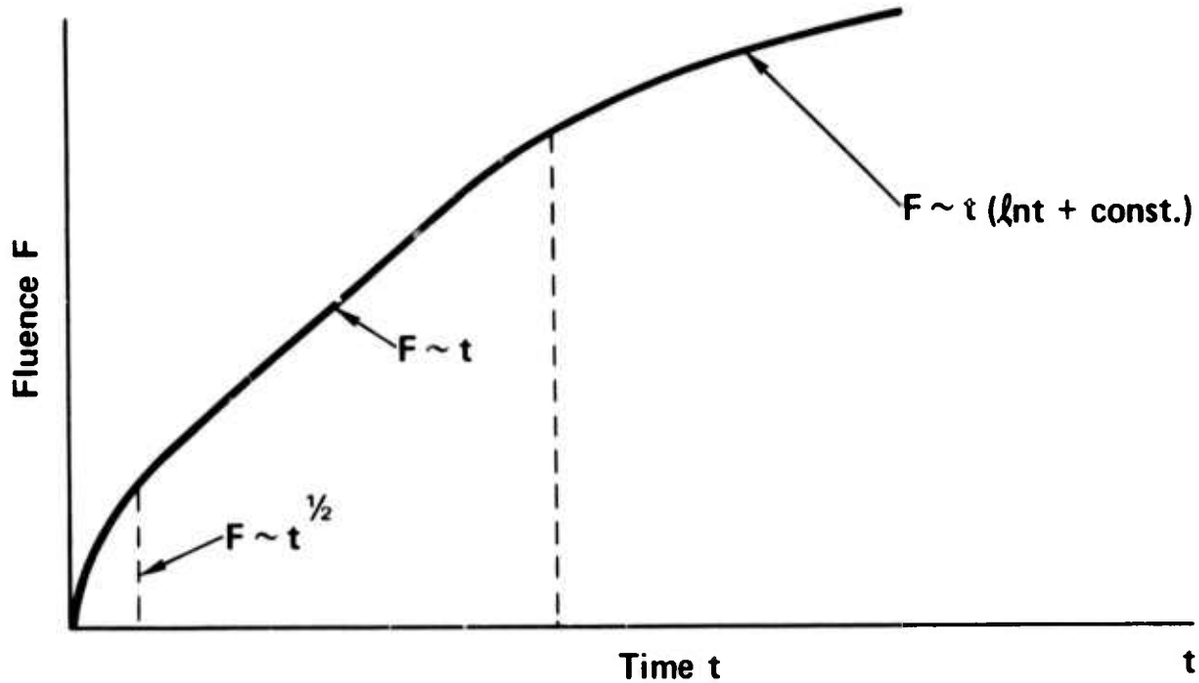

Temperature T(0,t) of illuminated surface
Time t
$\frac{IA}{C\delta}t$
$\frac{0.651IA\tau_\ell}{C\ell}$
$\frac{1}{3}\tau_\ell$
$\tau_{2\ell}$
$\tau_{4\ell}$
$\tau_{6\ell}$
$\frac{IA}{C\ell}(t+\frac{1}{3}\tau_\ell)$, Eq(3.13)
$\frac{IA}{C\ell}t$
$\frac{4}{\pi}\frac{IA\delta}{K}$
$\frac{2IA}{C\ell}(\tau_\ell\, t/\pi)^{1/2}$
$4\delta_8/\pi$
T(0,t)
t


Normalized Temperature Per Time $\frac{KT}{IA\delta}\frac{\tau_\delta}{t}$
1.0
0.8
0.6
0.4
0.2
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
Normalized Distance z/δ
$t/\tau_\delta \gg 1$ (schematic)
2×10^{-3}
$t/\tau_\delta = 2.5 \times 10^{-5}$
$\frac{1}{8}$
$\frac{1}{32}$
$\frac{1}{128}$

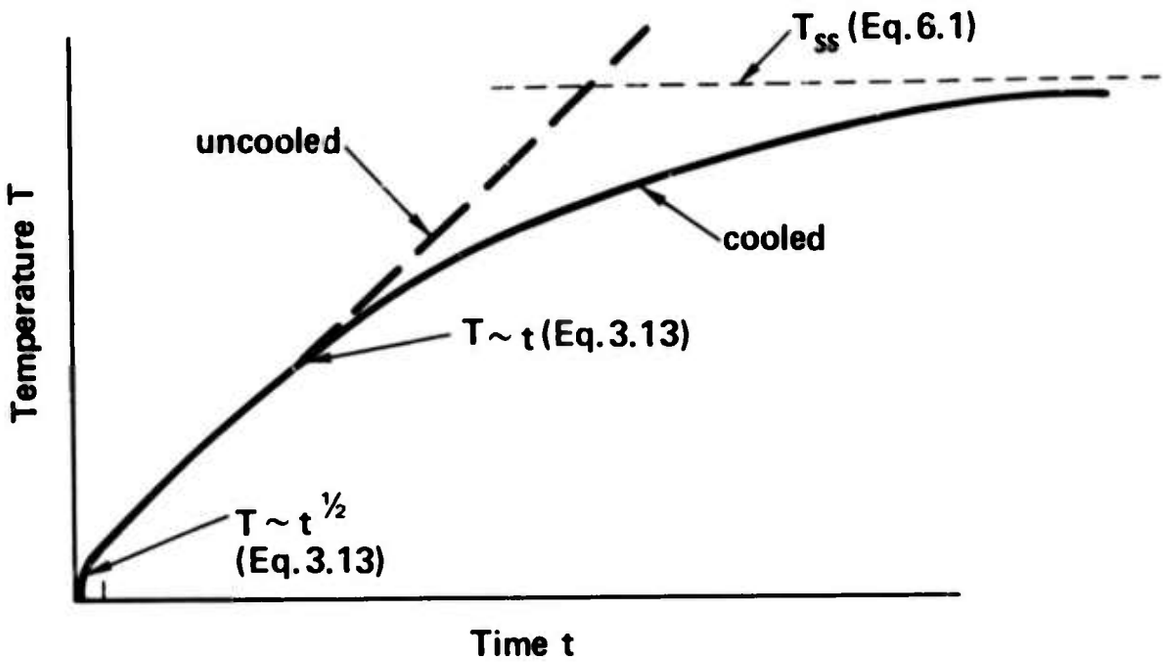

2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
0
1
Normalized Temperature $\frac{4K}{IA\ell}$ T
Normalized Distance z/ℓ
t/τ_ℓ = 0.30
0.20
0.15
0.10
0.06
0.03
0.01


T ~ t^½
Eq. (5.1)
Temperature T
T ~ const.
Eq. (5.4)
T ~ ℓnt + const.
Eq. (4.6)
τℓ
Time t

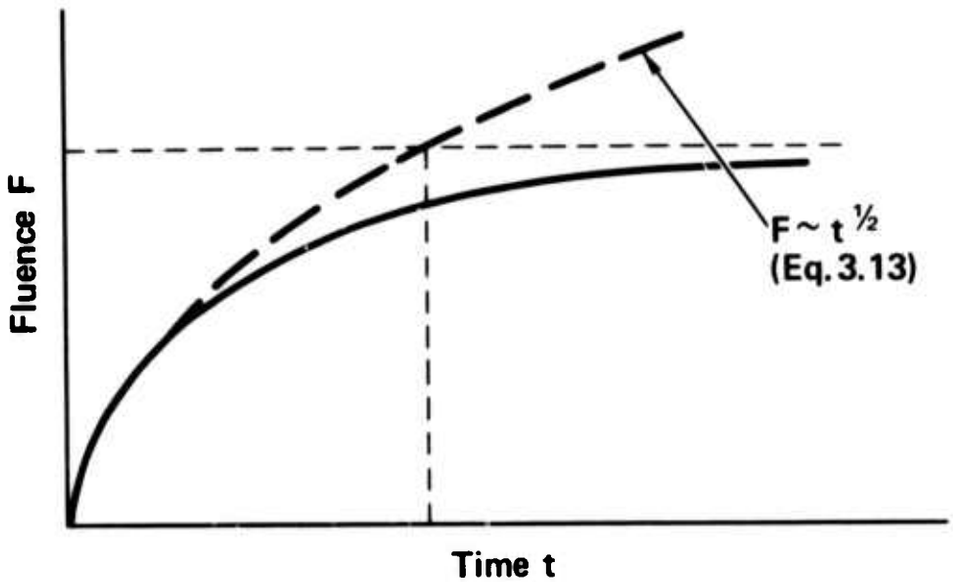
(a)


Fluence F
F ~ t (ℓnt + const.)
F ~ t
F ~ t^½
Time t
t

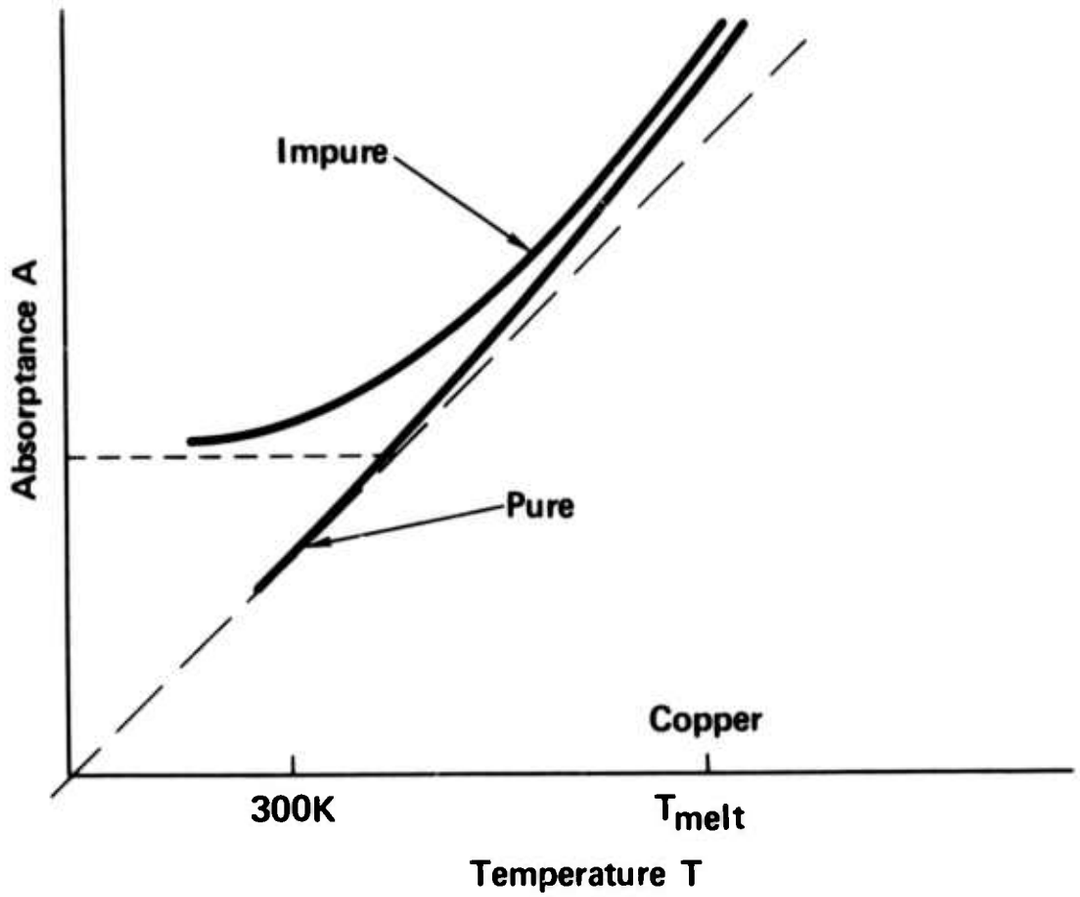
(b)

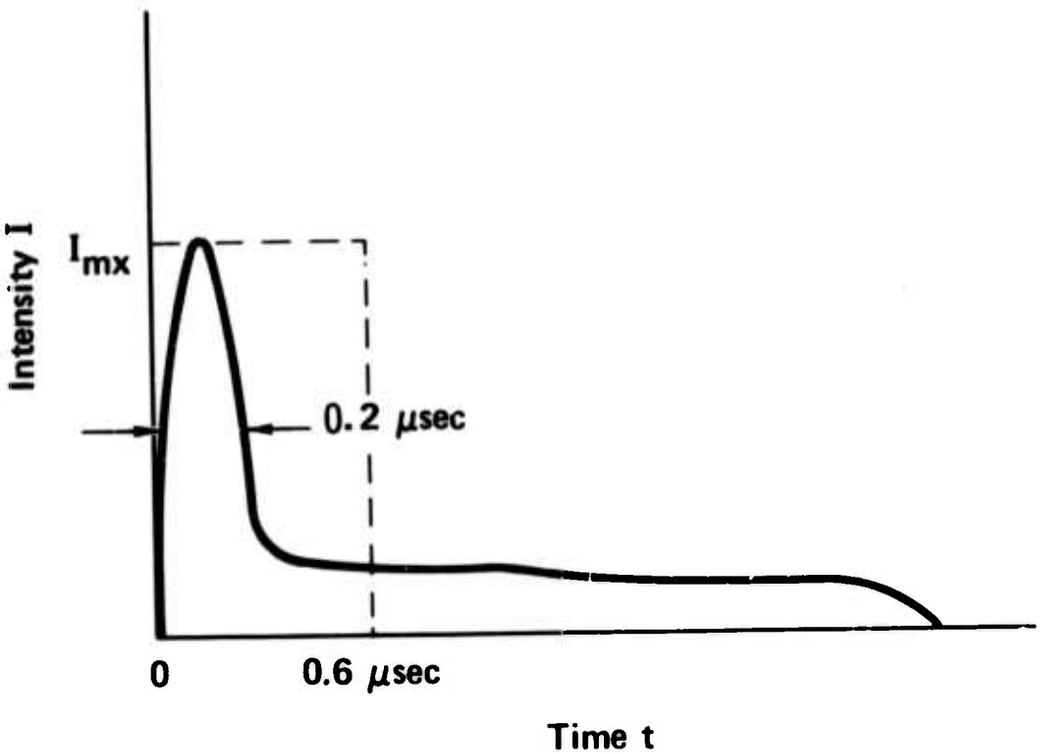

T_{ss} (Eq. 6.1)
uncooled
cooled
T ~ t (Eq. 3.13)
T ~ t ½
(Eq. 3.13)
Temperature T
Time t

(a)


F ~ t ½
(Eq. 3.13)
Fluence F
Time t

(b)


Impure
Pure
Absorptance A
Copper
300K
T_{melt}
Temperature T


Intensity I
I_{mx}
0.2 μsec
0
0.6 μsec
Time t

G. CURRENT STATUS OF HIGH-INTENSITY VACUUM ULTRAVIOLET MATERIALS*

M. Sparks and C. J. Duthler
Xonics, Incorporated, Van Nuys, California 91406

The recent successful operation of vuv lasers has attracted wide attention to high-intensity vuv materials since the first lasers were materials limited. In contrast to the IR case, the reflector problem in the vuv is more severe than the window problem, the number of candidate materials is much smaller, and optical-tolerance requirements are more severe. With extreme care in sample preparation the absorptance of aluminum, the only suitable metallic reflector, approaches 8%, which we show to be the intrinsic value. The absorption coefficients of transparent materials are currently extrinsic, with typical values of 0.1-1 cm^{-1}.

Calculations of the thresholds for damage in transparent materials indicate that with a single 10 nsec duration xenon-laser pulse, the failure intensity I_f is: ~20 MW/cm^2 for melting of a 100 A-thick film of aluminum on magnesium fluoride; ~70 MW/cm^2 for optical distortion by two-photon heating in windows and total-internal-reflection devices; ~200 MW/cm^2 for transparent-material surface damage; ~1 GW/cm^2 for thermal fracture by two-photon heating; ~1 GW/cm^2 for fracture from enhanced stimulated Raman scattering in Raman active materials; ~500 MW/cm^2 for optical distortion by one-photon heating with absorption coefficient $\beta = 0.1\ cm^{-1}$; and ~100 GW/cm^2 for thermal fracture by one-photon heating with $\beta = 0.1\ cm^{-1}$. The failure intensities for thermal fracture and optical distortion are

in general orders of magnitude lower in cw and repeated-pulse operation than those given above for a single 10 nsec pulse. Multilayer dielectric reflectors have greater theoretical values of I_f, but technical problems also must be considered.

A first estimate of two-photon absorption in LiF gives $\beta = 1\ cm^{-1}$ as the theoretical value of the absorption coefficient for $I = 2\ GW/cm^2$ at the photon energy $\hbar\omega = 7.2\ eV$ of the xenon laser. With this value of β, the optical distortion caused by two-photon heating of transparent optical components has the lowest failure threshold of all processes considered.

The recent successful operation of vacuum ultraviolet lasers has attracted wide attention to the problem of high-intensity materials. The first xenon lasers in fact were limited by the materials problem of the burning off of the aluminum films used as cavity mirrors. It is quite likely that we are now at the beginning of a new era of materials development in the vacuum ultraviolet.

The current status of the ARPA-sponsored vacuum ultraviolet materials program will be sketched. It is emphasized that the program is in the early stage of development. To date we have tried to anticipate the major materials problems that will arise as higher power becomes available in the vacuum ultraviolet, make order-of-magnitude estimates of the important processes, then study the individual problems in more detail. In this way an overview of the field was obtained, and the various parts of the program were kept in perspective. Hopefully this has avoided making analyses that later turn out to have little impact. For example, at the beginning of the program exhaustive studies of a number of metallic reflectors could have been made. It now appears that this would have been of little use, as discussed below.

Since this is a conference proceeding on infrared materials, comparisons of the ultraviolet materials problems to the infrared materials problems will be made whenever possible. The problems of obtaining high-power vacuum ultraviolet (vuv) materials are more difficult and quite different from those in the infrared.

In Fig. G1 the absorption coefficient is shown schematically as a function of frequency. To the left there is the familiar infrared peak in the absorption coefficient β at the Reststrahl frequency, followed by a nearly exponential decrease in β as the frequency increases. These features are understood in terms of multiphonon absorption. As the intrinsic multiphonon infrared absorption decreases with increasing frequency to a sufficiently small value, it is overshadowed by the extrinsic absorption.

At frequencies generally much higher than the fundamental Reststrahl frequency in the infrared, there is another strong absorption edge due to the electronic absorption which limits ultraviolet transmission. In the vuv, the emphasis will be on 7.2 eV (compare 0.1 eV for 10.6 μm) since this is the frequency of the xenon laser. Below the electronic absorption peak, as the intrinsic absorption coefficient reaches low values, the absorption is dominated by extrinsic absorption as in the IR case. Currently the extrinsic absorption is greater in the vuv (typically 0.1 to 1 cm^{-1}) than in the IR (typically 10^{-4} cm^{-1}).

In the infrared region it is possible to handle greater powers with metallic reflectors than with transparent materials because the absorptance of metals in the infrared is extremely low and because it is much easier to extract the absorbed heat from metals than from transparent materials. Thus in the infrared plane reflectors are made from metals and focusing of beams has been accomplished by curved metallic reflectors rather than by lenses. By contrast, in the

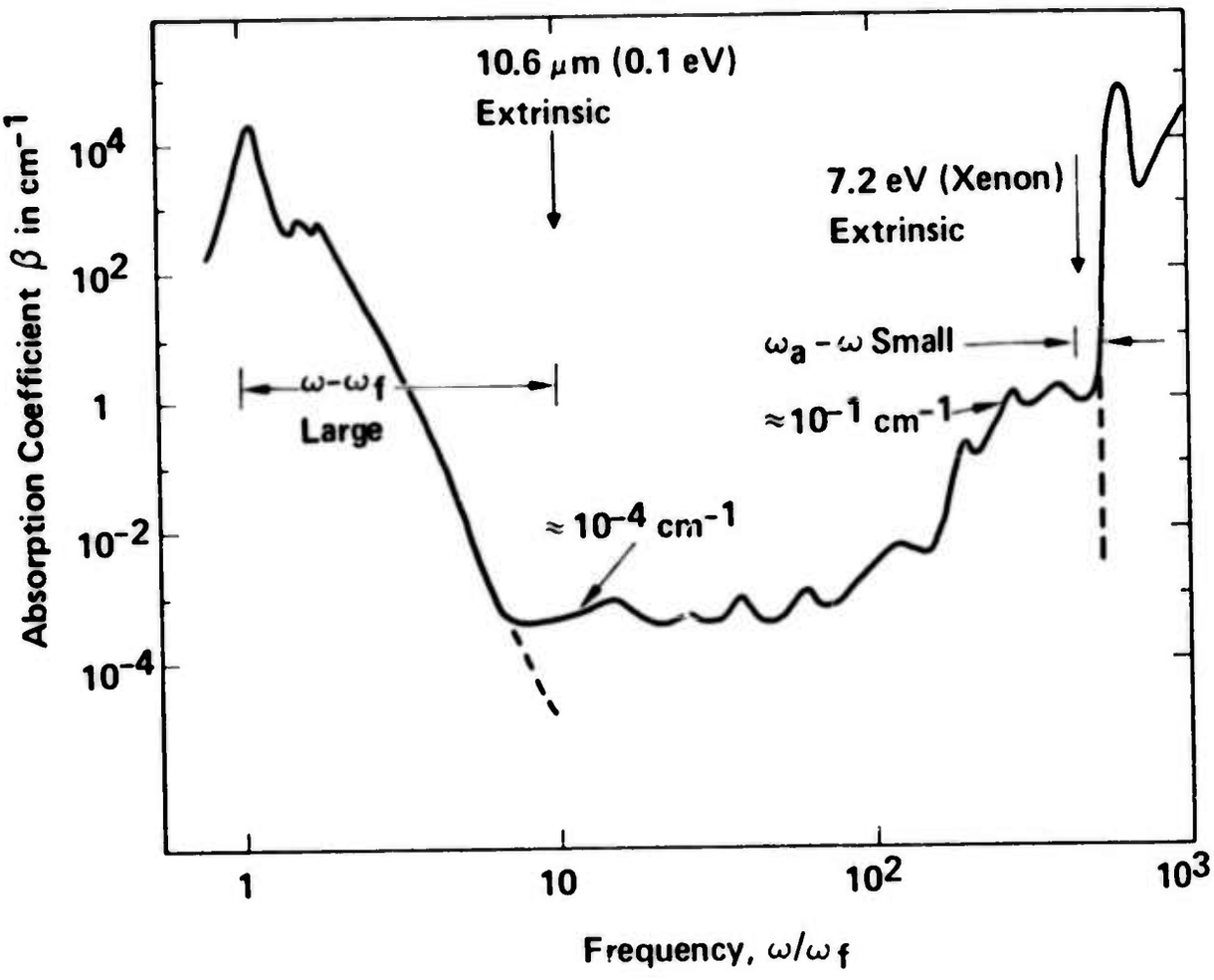


Fig. G1. Schematic optical absorption in KC1 from the infrared to the ultraviolet.

vacuum ultraviolet transparent materials can handle higher intensities than can metallic reflectors. Thus it is anticipated that lenses will be more important as focusing elements than curved metallic reflectors and dielectric-stack and total-internal-reflection devices will probably be more important than metallic reflectors for plane surface reflectors.

The reason for this difference is that the absorptivity of metals is generally high in the vacuum ultraviolet. Aluminum is the best vuv metallic reflector, and it is unlikely that a more suitable metal will be found since the plasma frequency must be high (three electrons for aluminum) and the interband transitions must be weak (nearly free-electron behavior for aluminum).

With extreme care in sample preparation, the absorptance of aluminum approaches 8 percent. This has only been possible in laboratory environments where thin films can be deposited rapidly at extremely high vacuums on supersmooth substrates and then overcoated by a protective layer such as MgF_2 or LiF. The best commercially available reflectors by contrast have an absorptance of about 20 percent and unless considerable care is exercised in preparing the samples the absorptance can be 40 percent or greater.

In Fig. G2 the reflectance of aluminum as a function of frequency is shown. In the infrared region the reflectance approaches 100 percent for the scale of the figure, while the reflectance at the xenon laser frequency of 7.2 eV approaches approximately 8 percent. As the frequency passes through the plasma frequency, the reflectance drops to a value near zero, of course. This top (8 percent) curve is for the case of extreme care in sample preparation.

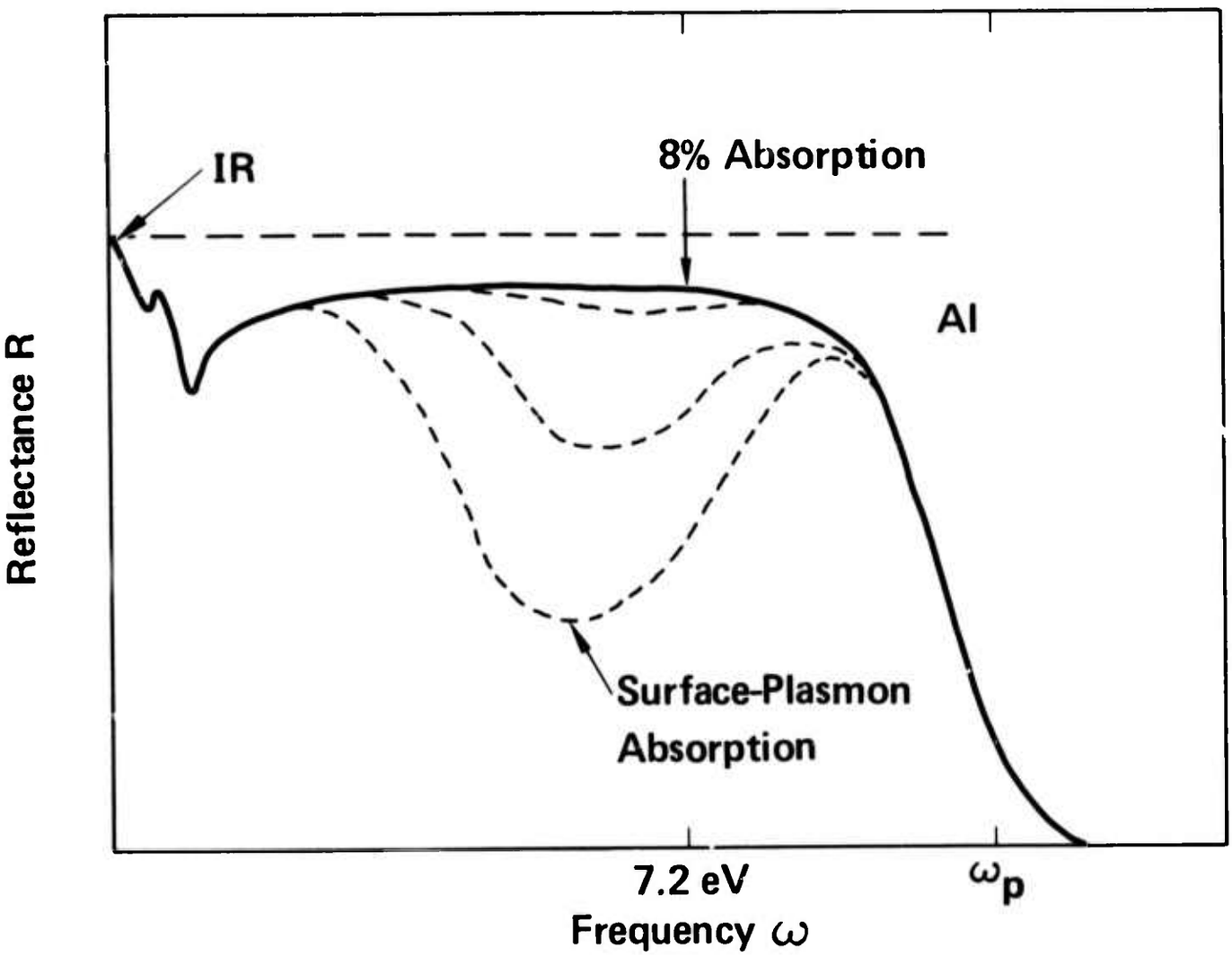


Fig. G2. Reflectance of aluminum showing the interband and surface plasmon absorption bands.

The effect of surface roughness on the reflectance is shown by the series of lower-reflectance curves. The greater the surface roughness, the greater the absorption as shown. For high-intensity reflectors, the limiting value of 8 percent absorptance is much greater than the infrared value of a small fraction of a percent absorptance. M. Sparks and H. C. Chow[1] extended the previous calculations of Ashcroft and Sturm[2] made for frequencies below 4 eV to the higher frequency range of current interest and found that the absorptance of 8 percent is intrinsic. Since aluminum is the best reflector in the vacuum ultraviolet and since it is extremely unlikely that other better metallic reflectors will be found, it is important that reflectors other than metallic reflectors be developed.

Before going on to the transparent materials which would be used in these dielectric reflectors as well as in windows and other optical components, consider the failure of the mirrors in the first xenon lasers. A simple calculation shows that this failure is simply the melting of the aluminum substrate as a result of the large absorptance. The mirrors were made by depositing approximately 100 A of aluminum onto magnesium fluoride substrates. The laser intensity is absorbed in this thin aluminum film, but the intensity required to melt the film is controlled by the diffusion of heat from the film onto the magnesium fluoride substrate. The volume of material heated, most of which is magnesium fluoride, is equal to the product of the surface area illuminated times an effective thickness d, where d is the distance into the magnesium fluoride that the heat diffuses. For a given pulse length t_p, the standard heat diffusion result is $t_p = 4Cd^2/\pi K$, where C is the heat capacity per unit volume and K is the thermal conductivity. From this result and the heat capacity equation $CV\Delta T = IA$, where ΔT is the temperature rise resulting from the pulse of intensity I on the metal having absorptivity A, solving for I gives

$$I_f = \frac{\Delta T_f}{2A} \left(\frac{\pi C K}{t_p} \right)^{1/2} \qquad (1)$$

where the subscripts f denote failure intensity and failure temperature rise.

For a pulse duration of 10 nsec and the value of C and K for MgF_2, Eq. (1) gives a failure intensity of approximately 20 MW/cm^2, which is equivalent to 0.2 J/cm^2 for this pulse duration. Both the failure intensity and the failure fluence change as the pulse length changes, since the diffusion distance d depends on the pulse duration.

Next consider transparent materials. As in the case of reflectors, the problem of transparent materials in the ultraviolet is much more difficult than in the infrared. Materials must have large bandgaps since the absorption edge must be above the operating frequency which is 7.2 electronvolts for the xenon laser. There are only a few materials with sufficiently large bandgaps, and most of these are unsatisfactory for use as practical optical components. See Table G1. Furthermore, the optical tolerances are much more severe in the vacuum ultraviolet, and the paucity of materials makes it difficult to obtain satisfactory dielectric reflectors and antireflection and protection coatings. There is no experimental information at present on the high-intensity failure of transparent materials since high-intensity sources have not been available for a sufficiently long time for measurements to have been made. Thus one of the most important parts of the program has been to make preliminary calculations of the values of intensity at which materials are expected to fail by various mechanisms.

A tentative result of these calculations is that the two-photon absorption is the most important failure mechanism for single short-pulse systems (pulse length 10 nsec). Energy bands of a transparent solid are shown schematically in Fig. G3.

Table GI. Electronic bandgaps of candidate materials for high-intensity vuv optical components. The buttons designate the more practical materials.

• LiF	~13	• SrF_2	>9	NaBr	7.7
• MgF_2	~11	• BaF_2	>9	• SiO_2	~7.7
KF	10.9	• NaCl	8.6	• MgO	7.3
NaF	>10.5	• KCl	8.5	KI	>6.2
RbF	10.4	LiBr	~8.5	LiI	>5.9
CsF	10	• Al_2O_3	8.3	NaI	>5.8
LiCl	~10	RbCl	8.2	Diamond	5.33
CaF_2	~10	KBr	7.8		

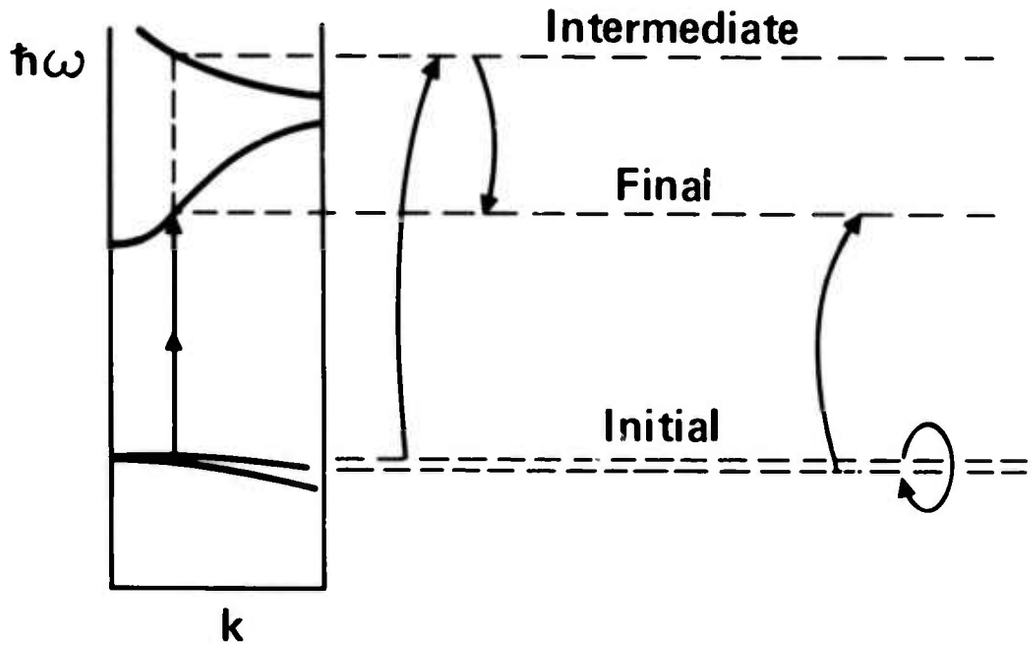


Fig. G3. Schematic illustration of two-photon absorption.

The bandgap must be greater than the photon energy; otherwise the one-photon absorption would be extremely large. Several materials have sufficiently large bandgaps to avoid one-photon absorption at the xenon-laser frequency, as shown in Table G I. However, the bandgap of all known materials is less than twice the xenon-laser frequency so that two-photon absorption is always possible.

Two-photon absorption is illustrated schematically by the double arrow in Fig. G3. The absorption process is a second-order process in the sense of perturbation theory. The first photon causes a virtual transition to an intermediate state as shown by the first curved arrow, and the second photon causes a virtual transition from this intermediate state to the final state as shown by the second curved arrow. Calculation of the absorption coefficient for this process is rather complicated. Since it became apparent at an early stage in the program that a precise calculation would be time consuming, a rough estimate of the absorption coefficient was first obtained by making some simple refinements to a previous zeroth-order theory of Braunstein.[3] It was found that the absorption coefficient was proportional to the intensity and had a value of 1 cm^{-1} for an intensity of 2 GW/cm^2. Thermal distortion from two-photon absorption has the lowest value of the failure intensity (70 MW/cm^2) of any of the processes considered. Thermally induced fracture occurs at $I_f \simeq 1\,GW/cm^2$. In view of these facts and the consideration that the accuracy of this estimate is less than an order of magnitude, refinement of two-photon absorption theory is one of the main goals of the program.

Next consider the values of the failure intensities for other mechanisms in transparent materials. The surface-damage mechanism, which occurs for example by the absorption of a macroscopic inclusion, can vary drastically from

sample to sample depending on the type of inclusion or the type of damage mechanism, just as in the cases of visible and IR radiation. A typical value for alkali halides is 200 MW/cm^2, or just 2 J/cm^2.

Damage can also result from Raman scattering since phonons are generated, which means that heat is generated in the crystal. If there were only the usual stimulated Raman scattering, the intensity threshold for damage by this heat generation would be quite large. However, it has recently been shown[4,5] that the Raman process has a parametric instability, which is analogous to instabilities that are well known in the cases of ferromagnetic resonance, higher-order absorption processes such as parallel pumping in ferromagnetism, in plasma physics, in atmospheric propagation, and in a number of other fields. The effect of this instability is that as the intensity reaches a threshold value which has a typical value of 1 GW/cm^2, the incident laser beam is converted to Stokes radiation in a very short distance near the surface of the crystal. Since one phonon is created for every Stokes photon created, the heat generated in this narrow layer near the surface is extremely large, and surface damage is expected to occur.

The final mechanism is that of the simple one-photon absorption by impurities. For an impurity absorption coefficient of 0.1 cm^{-1}, the failure intensity for thermal distortion is approximately 500 MW/cm^2. The intensity for fracture is $\sim$100 GW/cm^2. This value of the absorption coefficient of 0.1 cm^{-1} is at the lower range of 0.1 to 1 cm^{-1} for a numbe. of materials for which the measurements have been reported. Reducing the impurity absorption to lower levels will not result in further improvement in single-pulse operation since the absorption is dominated by the two-photon process at these intensities. The values of absorption coefficients for the various processes are shown in the bar graph representation in Fig. G4.

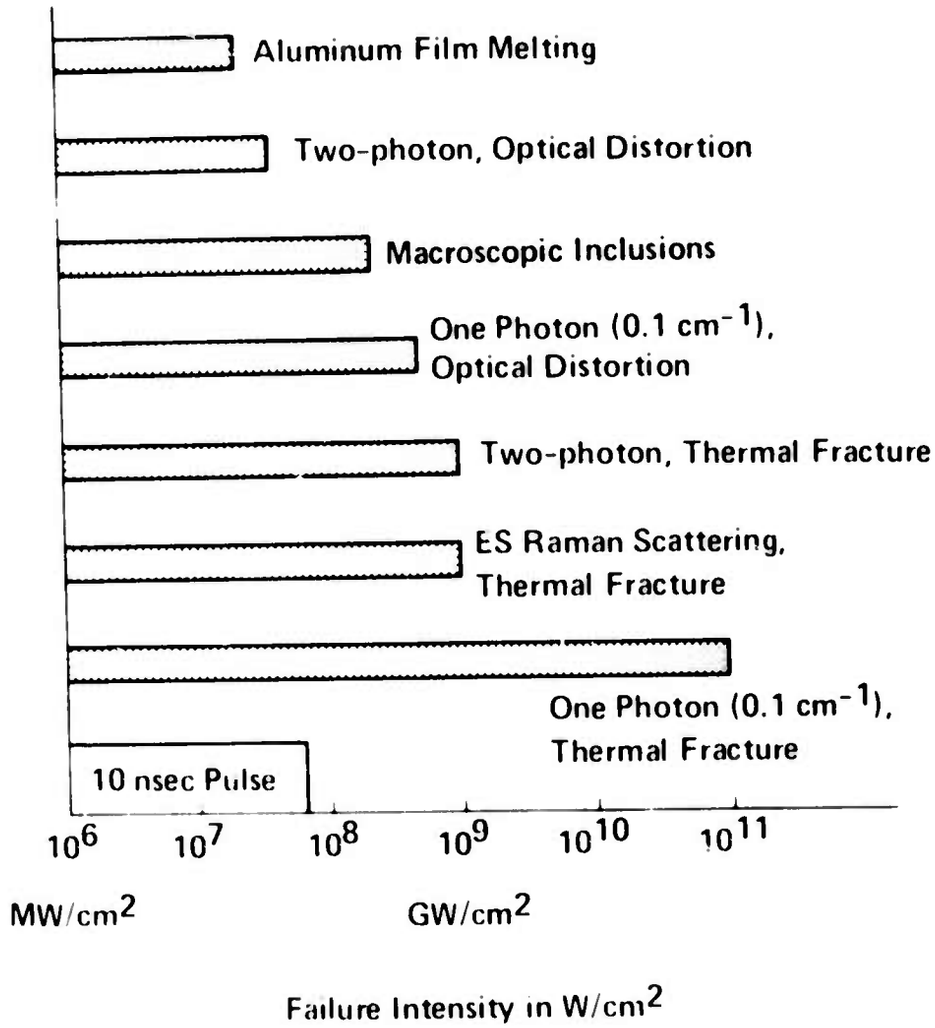


Fig. G4. Summary of values of failure intensities.

REFERENCES

*This paper was presented at the ARPA Fourth Laser Window Conference, Tucson, Arizona, 18-20 November 1974.

1. M. Sparks and H. C. Chow, Phys. Rev. B 10, 1699 (1974).

2. N. W. Ashcroft and K. Sturm, Phys. Rev. B 3, 1898 (1971).

3. R. Braunstein, Phys. Rev. 125, 475 (1962); R. Braunstein and N. Ockman, Phys. Rev. 134, A499 (1964).

4. M. Sparks, Phys. Rev. Lett. 32, 450 (1974); M. Sparks, Phys. Rev., in press.

H. IMPURITY ABSORPTION IN HALIDE WINDOW MATERIALS*

C. J. Duthler
Xonics, Incorporated, Van Nuys, California 91406

Molecular-ion impurities that substitute for the halide ion in alkali-halide and alkaline-earth-halide crystals constitute an important class of impurities limiting infrared transmission in laser-window materials. A literature survey indicates that concentrations of less than 0.1 ppm of NO_2^-, HCO_3^-, SO_4^{2-}, and CrO_4^{2-} will yield an absorption coefficient $\beta > 10^{-4}\,cm^{-1}$ at 10.6 μm in KCl or KBr. The hydroxyl ion (OH^-) will severely limit transmission near 3 μm.

I. INTRODUCTION

Among the most promising materials for use with high-power 10.6 μm wavelength CO_2 lasers are the alkali halides, in particular KCl and KBr. The mechanism for intrinsic infrared absorption in these materials is multiphonon absorption with the fundamental Reststrahl frequency ν_ℓ of these crystals being six to eight times smaller than the CO_2 laser frequency of 943 cm^{-1}. It has been established experimentally[1] and theoretically[2] that for $\nu > \nu_\ell$, the intrinsic multiphonon absorption coefficient β_{int} decreases nearly exponentially with increasing frequency. Extrapolation of the measured low-frequency intrinsic absorption coefficients yields $\beta_{int} = 8 \times 10^{-5}\,cm^{-1}$ for KCl and $\beta_{int} = 5 \times 10^{-7}\,cm^{-1}$ for KBr at 943 cm^{-1}. However, absorption in even the best currently available materials is impurity dominated with $\beta > 10^{-4}\,cm^{-1}$ at the CO_2 laser frequency.

The present paper provides evidence that an important class of infrared-absorbing impurities is polyatomic molecular anions that substitute for the halide ions in the crystal lattices. Several ionic impurities that absorb strongly at 943 cm^{-1} are identified. The influence of absorbing inclusions on the bulk absorption coefficient and the failure of materials containing inclusions have been considered in previous publications.[3,4]

II. IMPURITY SPECTRA

As an example, the case of NO_2^- substituting for the Cl^- ion in KCl is drawn to scale in Fig. H1.[5] This ion fits loosely into the lattice and suffers little distortion. Consequently the internal vibrations of NO_2^- and other substitutional anions are only slightly affected by the presence of the host crystal. Frequency shifts of the molecular modes and splittings of degenerate modes are small. Infrared inactive modes may be weakly absorbing in the crystal.

Because of the small interaction with the host crystal, spectroscopists have used the technique of alkali-halide matrix isolation for a number of years to concentrate and isolate ions for convenient study using conventional absorption spectroscopy. A literature survey has been made of those ions that have been studied in KCl and KBr crystals. Those that most severely limit 10.6 μm transmission are presented below.

To estimate the absorption at the laser frequency of $\nu = 943\ cm^{-1}$ due to an impurity mode at ν_0, a Lorentzian line shape

$$\beta(\nu) = (\text{const}) \frac{\Delta\nu}{(\nu - \nu_0)^2 + (\Delta\nu/2)^2} \tag{1}$$

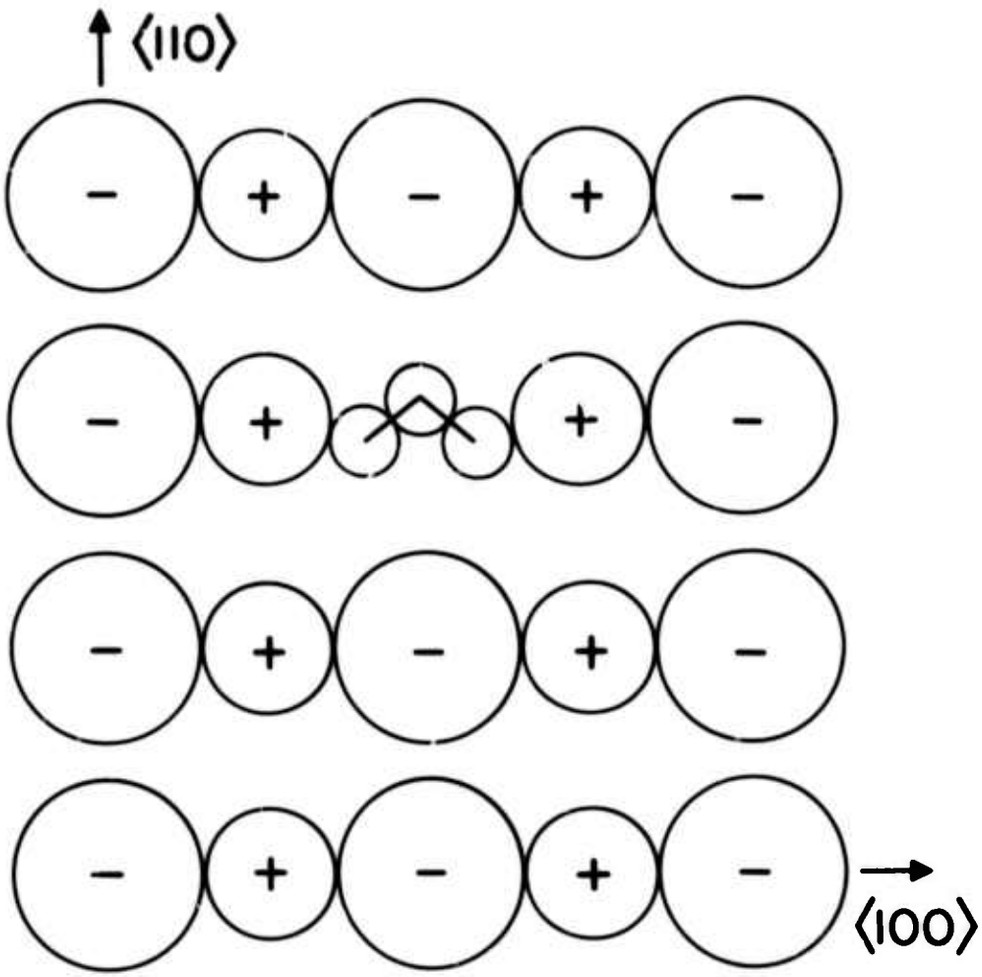


Fig. III. Orientation of NO_2^- impurity in KCl at room temperature.

is assumed. In Equation (1), $\Delta\nu$ is the full width at half maximum, which is assumed to be frequency independent. When available, measured line strengths and line widths were used for extrapolation to 943 cm^{-1}. Typically for a crystal with 0.1% impurity concentration, the absorption coefficient $\beta(\nu_0)$ at the center of an allowed transition ranges from 10^2 to 10^3 cm^{-1}, with $\Delta\nu$ ranging from 4 to 50 cm^{-1} depending on the particular molecular species and host.

Impurities that substitute for lattice ions can be grouped into a few general types. First, atomic impurities can either yield an infrared active local mode of the lattice, or can activate normally infrared inactive phonons. However, because the frequencies of these modes are generally small compared to 943 cm^{-1}, little absorption occurs at the CO_2 laser frequency. The U center at 500 cm^{-1} has the highest frequency of this group and we estimate that more than 10 ppm H^- is needed to produce $\beta(\nu = 943\ cm^{-1}) = 10^{-4}\ cm^{-1}$.[6]

Diatomic ionic impurities that have been studied in KCl and KBr crystals, on the other hand, have absorption frequencies that are too large to result in significant absorption at 943 cm^{-1} in spite of their large room-temperature line widths. The hydroxyl ion (OH^-) with $\nu = 3640\ cm^{-1}$, is a member of this group. There has been concern regarding the hydroxyl ion because of the difficulty in removing this ion from alkali-halide crystals.[7] However, using $\Delta\nu = 50\ cm^{-1}$ in Eq. (1), we estimate that greater than 100 ppm of OH^- at isolated lattice sites is needed to yield $\beta(943\ cm^{-1}) = 10^{-4}\ cm^{-1}$.[8] At high impurity concentrations, OH^- may combine with other impurities to absorb more strongly at 943 cm^{-1}.

At chemical laser wavelengths near 3 μm diatomic ionic impurities will severely limit laser transmission. Besides OH^- with $\lambda = 2.7\ \mu m$, SH^- and CN^-

are observed to have strong absorption lines at 3.9 μm and 4.8 μm, respectively, in both KCl and KBr.[9] Since the absorption wavelengths are a property of the impurity ion rather than the host material, the above diatomic ions, when present in alkaline-earth halides, should absorb at nearly the same wavelengths.

Polyatomic impurities have several infrared active modes, some of which may occur near 943 cm^{-1} and produce considerable CO_2 laser absorption. This group can be further subdivided into small and large ions. Small polyatomic ions such as NO_2^- fit loosely into the lattice and have large, temperature dependent line widths due to rotational and translational degrees of freedom. Larger ions such as HCO_3^- fit tightly into the lattice and have several narrow, temperature insensitive lines.

In addition to the above groups, polyvalent ions such as SO_4^{2-} can substitute for the singly charged halide ion. In order to preserve the charge neutrality of the crystal, these ions must be accompanied by a compensating vacancy or by a divalent cation impurity such as Mg^{2+}. The infrared spectrum and 943 cm^{-1} absorption is dependent on the particular compensating species.

Our literature survey indicates that 0.03 ppm of either NO_2^-, HCO_3^-, SO_4^{2-}, or CrO_4^{2-} will result in $\beta(\nu = 943\ cm^{-1}) = 10^{-4}\ cm^{-1}$. Calculated spectra using an assumed Lorentzian line shape along with experimental frequencies and line widths are shown in Figure H2(A-C) for NO_2^-, HCO_3^-, and SO_4^{2-}.[10]

In each of the cases shown in these figures, the laser frequency, 943 cm^{-1}, falls at the edge of a small number of peaks representing the infrared active internal modes of the impurity. It would seem from looking at these figures that it would be easy to measure the spectrum of a crystal containing an unknown impurity and hence determine the impurity. However, there are two difficulties.

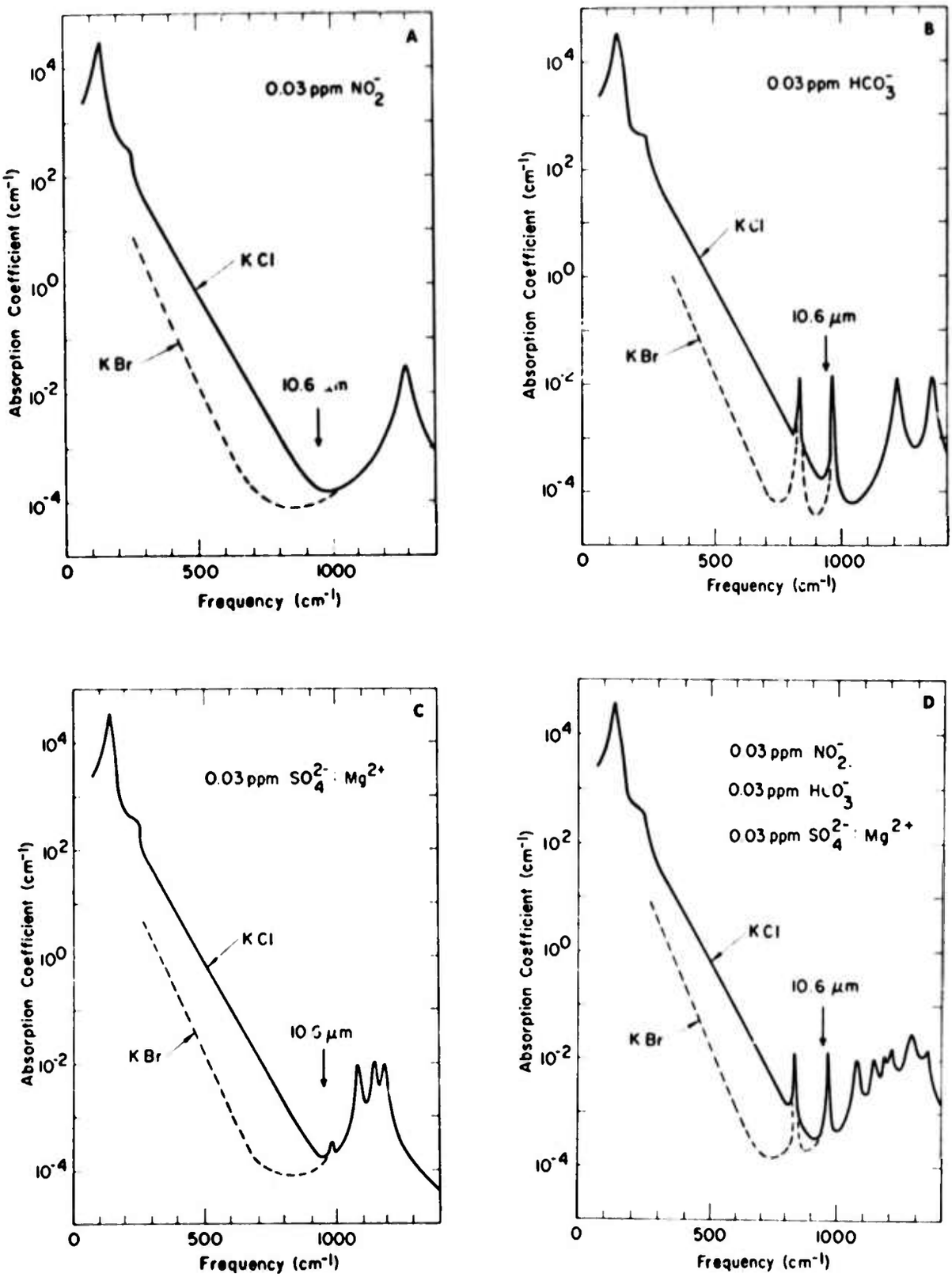


Fig. H2. Calculated absorption coefficients of KCl and KBr crystals containing various impurities.

First, the absorption coefficient at the peaks is only 10^{-2} cm^{-1} so that it would be difficult to make high resolution measurements using conventional spectroscopy. Second, and more important, several impurities with overlapping absorption lines will be present in a real crystal.

In Fig. H2-D we represent a real crystal by adding the absorption coefficient in the previous three figures. The individual lines near 1200 cm^{-1} have merged to form a relatively featureless, broad peak. The only distinctive lines remaining are the two HCO_3^- lines at 840 cm^{-1} and 971 cm^{-1}. If we next decrease the amount of HCO_3^- and add different compensating divalent cations such as Pb^{2+} along with a small amount of CrO_4^{2-}, even the calculated spectrum would appear completely featureless. This overlapping of absorption lines from several polyatomic ions (all of which contain oxygen) could account for the broad unidentified absorption peak near 9.5 μm that has been observed experimentally.[11]

The above examples serve to illustrate the importance of molecular-ion impurities on laser transmission in ionic solids. Besides the impurities given above, it has recently been reported that ClO_3^- will severely limit 10.6 μm transmission through alkali halides.[12] CO_2^- and CO_3^{2-} will absorb near 6 μm but should have little effect on 10.6 μm transmission.[13] A more complete survey of the ions that have been studied in KCl and KBr has been published elsewhere.[14]

REFERENCES

*This paper was presented at the ARPA Fourth Laser Window Conference, Tucson, Arizon., 18-20 November 1974.

1. T. F. Deutsch, J. Phys. Chem. Solids 34, 2091 (1973).

2. For theories of intrinsic multiphonon absorption see M. Sparks and L. J. Sham, Phys. Rev. B 8, 3037 (1973) and references therein.

3. M. Sparks and C. J. Duthler, J. Appl. Phys. 44, 3038 (1973).

4. C. J. Duthler, Appl. Phys. Lett. 24, 5 (1974).

5. A. R. Evans and D. B. Fitchen, Phys. Rev. B 2, 1074 (1970).

6. A. A. Maradudin in Solid State Physics, edited by F. Seitz and D. Turnbull (Academic, New York, 1966), Vol. 18, p. 274 and Vol. 19, p. 1.

7. R. C. Pastor and M. Braunstein, Hughes Research Laboratories Technical Report No. AFWL-TR-72-152, Vol. II (1973).

8. B. Wedding and M. V. Klein, Phys. Rev. 177, 1274 (1969); M. V. Klein, B. Wedding, and M. A. Levine, Phys. Rev. 180, 902 (1969); D. F. Smith, J. Overend, J. C. Decius, and D. J. Gordon, J. Chem. Phys. 58, 1636 (1972).

9. Data for CN^- from W. D. Seward and V. Narayanamurti, Phys. Rev. 148, 463 (1963), and G. R. Field and W. F. Sherman, J. Chem. Phys. 47, 2378 (1967). Data for SH^- from C. K. Chi and E. R. Nixon, J. Phys. Chem. Solids 33, 2101 (1972).

10. The data used for NO_2^- in KCl are obtained from V. Narayanamurti, W. D. Seward, and R. O. Pohl, Phys. Rev. 148, 481 (1966); V. P. Dem'yanenko and Yu. P. Tsyashchenko, Sov. Phys.-Solid State 11, 3043 (1970); R. Kato and J. Rolfe, J. Chem. Phys. 47, 1901 (1967); R. Bonn, R. Metselaar, and J. van der Elsken, J. Chem. Phys. 46, 1988 (1967); A. R. Evans and D. B. Fitchen, Phys. Rev. B 2, 1074 (1970). The HCO_3^- data from D. L. Burnitt, K. O. Hartman, and I. C. Hisatsune, J. Chem. Phys. 42, 3553 (1965). The SO_4^{2-} data from J. C. Decius, E. H. Coker, and G. L. Brenna, Spectrochim. Acta 19, 1281 (1963); E. H. Coker, J. C. Decius, and A. B. Scott, J. Chem. Phys. 35, 745 (1961); D. N. Mirlin and I. I. Reshina, Sov. Phys.-Solid State 10, 895 (1968).

11. M. Hass, J. W. Davisson, P. H. Klein, and L. L. Boyer, J. Appl. Phys. 45, 3959 (1974); T. F. Deutsch, Appl. Phys. Lett. 25, 109 (1974).

12. H. G. Lipson, J. J. Larkin, and B. Bendow, Third Conference on High Power Infrared Laser Window Materials, Air Force Cambridge Research Labs, Vol. I, p. 237 (1973).

13. H. W. Morgan and P. A. Staats, J. Appl. Phys. 33, 364 (1962); K. O. Hartman and I. C. Hisatsune, J. Chem. Phys. 44, 1913 (1966); V. P. Dem'yanenko, Yu. P. Tsyashchenko, and E. M. Verlan, Sov. Phys.-Solid State 13, 767 (1971).

14. C. J. Duthler, J. Appl. Phys. 45, 2668 (1974).

I. LIST OF RECENT PUBLICATIONS

New publications and those whose status has changed since the Third Technical Report of 30 June 1974 are included in the following list:

1. M. Sparks and H. C. Chow, "Parametric Instabilities of Phonons: Nonlinear Infrared Absorption," Phys. Rev. B 10, 1699 (1974).

2. D. L. Mills and A. A. Maradudin, "The Absorption Coefficient of Alkali Halides in the Multiphonon Regime: Effects of Nonlinear Dipole Moments," Phys. Rev. B 10, 1713 (1974).

3. M. Sparks, "Infrared Absorption by the Higher-Order-Dipole-Moment Mechanism," Phys. Rev. B 10, 2581 (1974).

4. M. Sparks and C. J. Duthler, "Current Status of High-Intensity Vacuum Ultraviolet Materials," ARPA 4th Laser Window Conference, Tucson, Arizona, 18-20 November, 1974.

5. C. J. Duthler, "Impurity Absorption in Halide Window Materials," ARPA 4th Laser Window Conference, Tucson, Arizona, 18-20 November 1974.

6. M. Sparks, "Stimulated Raman Scattering: Enhanced Stokes Gain and Effects of Anti-Stokes and Parametric Phonon Processes," Phys. Rev., in press.

7. A. Karo, M. Sparks, and L. J. Sham, "Infrared Multiphonon Absorption Calculations," in preparation.

8. A. A. Maradudin and D. L. Mills, "The Scattering and Absorption of Electromagnetic Radiation by a Semi-Infinite Crystal in the Presence of Surface Roughness," Phys. Rev., in press.

9. M. Sparks and C. J. Duthler, "Intensity Limits of High-Intensity Vacuum Ultraviolet Materials," to be published.

10. M. Sparks and J. H. Wilson, "Enhanced Stimulated Raman Scattering and General Three-Boson Parametric Instabilities," submitted to Phys. Rev.

11. H. C. Chow and M. Sparks, "Calculated Reflectance of Aluminum in the Vacuum Ultraviolet," J. Appl. Phys., in press.

12. M. Sparks, "Theory of Laser-Materials Damage by Enhanced Stimulated Raman Scattering," submitted to J. Appl. Phys.

13. M. Sparks, "Theory of Laser Heating of Solids: I. Metals," to be submitted to J. Appl. Phys.

14. D. L. Mills and A. A. Maradudin, "Surface Roughness and the Optical Properties of a Semi-Infinite Material: The Effect of a Dielectric Overlayer," submitted to Phys. Rev.