BELIC, MILIVOJ RADIVOJ

SOME ASPECTS OF THE NONLINEAR INTERACTION OF LIGHT AND MATTER

City University of New York PH.D. 1980

SOME ASPECTS OF THE NONLINEAR INTERACTION OF LIGHT AND MATTER

by

MILIVOJ R. BELIC

A dissertation submitted to the Graduate Faculty in Physics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1980

This manuscript has been read and accepted for the Graduate Faculty in Physics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

March 14. 1980
date

Prof. Joel I. Gersten
Chairman of Examining Committee

March 14. 1980
date

Prof. Frank Martino
Executive Officer

Prof. Robert R. Alfano

Prof. Melvin Lax

Prof. Joseph B. Krieger

Prof. Harvey Kaplan
Supervisory Committee

The City University of New York

ABSTRACT

SOME ASPECTS OF THE NONLINEAR INTERACTION OF LIGHT AND MATTER

by

MILIVOJ R. BELIC

Adviser: Professor Joel I. Gersten

The present thesis is concerned with various aspects of the interaction of light with matter. It is divided into four parts. In part I we study the competition between self-phase modulation (SPM) and stimulated Raman scattering (SRS). In part II we study a realistic model for the unstable loaded optical resonator in the paraxial wave approximation. Part III is concerned with the possibility of utilizing optical power to directly drive a refrigerator. Part IV is concerned with the nature of the energy band structure of solids in an external laser field.

ACKNOLEDGMENT

I wish to thank Professor Joel Gersten and Professor Melvin Lax for their precious help, generous support and endless patience which made completion of this work possible.

Contents :

0.1 Introduction	1
Chapter 1 : Combined SRS and SPM	3
1.1 Introduction	3
1.2 Theory	8
1.3 Results	17
Chapter 2 : Unstable optical resonators	27
2.1 Introduction	27
2.2 Theory	32
2.3 Resonator model	40
2.4 Modified predictor-corrector	45
2.5 Fast Fourier transform	52
2.6 Programming considerations and results	70
Chapter 3 : Solar heat pump	82
3.1 Introduction	82
3.2 Kinetic theory	86
3.3 Phenomenological equations	93
3.4 Analysis of the equations	102
3.5 Thermal currents	106
3.6 The device	112
Chapter 4 : Band spectrum in a laser field	127
4.1 Introduction	127
4.2 Modified Kronig-Penney model	129
4.3 General theory	137

List of Tables and Figures

Table 1.1	21
Figure 1.1	23
Figure 1.2	24
Figure 2.1	75
Figure 2.2	76
Figure 2.3	77
Figure 2.4	78
Figure 2.5	79
Figure 3.1	122
Figure 3.2	123
Figure 3.3	124
Figure 3.4	125
Figure 4.1	148
Figure 4.2	149
Figure 4.3	150

0.1 Introduction

The present thesis is concerned with various aspects of the interaction of light with matter. It is divided into four parts. Parts I and II are concerned primarily with the nonlinear propagation of electromagnetic waves through matter. Part III is concerned with thermodynamic questions pertaining to the propagation of light through matter. Part IV involves the question of how matter is altered when it is subjected to a strong electromagnetic perturbation.

In part I we study the competition between self-phase modulation and stimulated Raman scattering. Aside from the intrinsic interest in these mechanisms, per se, they are important because they limit the ability to design high power lasers. An understanding of the systematics of these effects is cruicial to the optimization of laser design.

In part II we study a realistic model for the unstable loaded optical resonator in the paraxial wave approximation. An analysis of this kind is important in the theoretical understanding and numerical handling of high power lasers.

Part III is concerned with the possibility of utilizing optical power to directly drive a refrigeratior. A solar powered semiconductor heat pump device is proposed, based on the combination of photovoltaic and Peltier effects.

Part IV is concerned with the nature of the energy band structure of solids in an external laser field. We investigate the mathematical nature of the spectrum, with the goal of obtaining realistic band structures to be used in calculations of optical and spectroscopic properties of solids.

CHAPTER 1. Combined SRS and SPM

1.1 Introduction

Two effects of interest in understanding the propagation of intense optical pulses through matter are stimulated Raman scattering and continuum self-phase modulation. Both may be categorized as a third order nonlinear optical phenomenon, and their theory is well established.[1-4]

In a light scattering process, a photon at ω_1 is absorbed and a photon at ω_2 is emitted while the material makes a transition from the initial state Ψ_i to the final state Ψ_f. The material excitation involved can be an entropy wave (Rayleigh scattering), a pressure wave (Brillouin scattering), a phonon, magnon, plasmon or electronic excitation (Raman scattering), a concentration variation (concentration scattering), etc. The transition probability[5] per unit time per unit volume from Ψ_i to Ψ_f can be obtained from a perturbation calculation and is proportional to $m_1 (m_2 + 1)$, where m_1 is the number of photons in the incoming radiation mode, and m_2 is the corresponding number in the outgoing (scattered)

mode. Spontaneous scattering corresponds to $m_2 = 0$ and stimulated to $m_2 \neq 0$. Here we shall limit our discussion to stimulated Raman scattering (SRS) with optical phonons. Since the Raman gain is usually very small, SRS can only be observed with a high-power laser beam. It has, however, been seen in some liquids at a much lower power level. In these cases SRS is actually initiated by self-focusing of the laser beam, and the apparently lower threshold corresponds to the self-focusing threshold[6,7]. It was combined nonlinear action like this one that prompted our attention. However, a general, unified theory of transient self-focusing and SRS turned out to be a too tough nut to crack. It requires a solution to the coupled system of nonlinear differential equations which include transverse derivatives as well as longitudinal derivatives. Therefore, we decided to tackle in this chapter only some aspects of the combined action - like the effect of self-phase modulation (SPM) on the SRS. Some preliminary analysis of this problem has been carried out by the Soviets[7], but within a different context and without too many details. A similar problem has been addressed by Carman et al.[2] in 1970. However, their procedure was numerical and simplified, in that the effects of laser-pump depletion and saturation of the material system were neglected.

Physically, self-focusing may occur when the refractive index n of the nonlinear medium increases with the beam

intensity. Consider a laser beam with Gaussian transverse profile propagating into the medium, as in Fig.1. The central part of the beam sees a larger refractive index than the edge and therefore propagates with a lower velocity. Consequently, as the beam traverses the medium, the original wave front gets more and more distorted - the central portion lags behind. Since the rays should propagate in a direction perpendicular to the wave front, the beam appears to focus itself. A beam with a finite cross section will of course also diffract. Self-focusing occurs only if the self-focusing action is stronger than the diffraction action. Otherwise the beam will still diffract , even though the diffraction would be weaker.

In order to account for the transient nature of the process, assume our beam to be a strong, short pulse (Fig. 2). Qualitatively, the leading edge of the pulse (section a in the figure) sees only a small induced change in the refraction index Δn and shows weak or no self-focusing, but it affects strongly the self-focusing dynamics of the lagging part. The next part (section b) sees somewhat larger Δn and accordingly diffracts not as strongly. The Δn induced by the front part is then large enough to cause the lagging part (sections c-f in tne figure) to self-focus. However, as the beam propagates on , it sees a gradually decreasing Δn because the front part has diffracted. As a result, self-focusing of the lagging part

tends to be more gradual. If no other strong nonlinear effect were to suddenly cut off the self-focusing action, diffraction from the focal region will also be gradual, leading to a long focus. Knowing how the various parts of the pulse propagate in the medium, we can find how the transverse profile of the pulse gets deformed during propagation. This is sketched in Fig. 2. where we assume there is a limiting diameter for the focus. It is seen that the pulse is quickly deformed into a horn shape. Obviously, the neck of the horn sweeping along the axis leads to an intense filament and the strong phase modulation acquired by the neck gives rise to large spectral broadening. Both filament and spectral broadening are characteristics of picosecond transient self-focusing[8].

As our strong optical pulse traverses the medium, energy is transfered from a primary mode, which is at the laser frequency, to other modes of the system. In the case of stimulated Raman scattering , the shift in frequency takes place in units of the vibrational frequencies of the medium. Thus, the primary mode becomes depleted and various Stokes modes are intensified. When spectral broadening occurs, it occurs in a quantized manner. In the case of self-phase modulation , however, the repopulation of the spectral intensity takes place in a more gradual manner. Due to the nonlinearity of the medium, the pulse heterodynes against itself and gradually increases its spectral width.

There is a continuum of frequencies produced in this process.

Since both the stimulated Raman scattering and self-phase modulation spread the spectral width of the pulse, one can ask the question if they occur independently of each other or if they compete and interfere. Clearly, if the SRS effect is effective in transfering energy to the Stokes branches very rapidly, SPM will not occur to any significant extent. On the other hand if SPM is the dominant effect, a Stokes shift of the primary mode must be replaced by a Stokes shift of a broadened primary mode, and the spectrum will lose its discrete character. Thus the two effects influence each other and should be studied simultaneously. Our goal here will be to calculate the extent of this interaction in various limiting cases.

In the next section we introduce a simplified model [9] which captures the essence of both processes. We will limit our attention to one dimensional propagation, even though the transverse character of the pulse is needed for a more complete theory (such as one which would include self-focusing). Likewise dispersion effects will be neglected, as these mainly affect the spatial evolution of the pulse. A number of other simplifying assumptions will be made to bring the equations to a tractable level.

CHAPTER 1. Combined SRS and SPM

1.2 Theory

The system of interest consists of an electromagnetic field interacting with the molecular vibration (phonon) field of a liquid. The appropriate Lagrangian density is

$$\mathcal{L} = \frac{\epsilon E^2 - B^2}{8\pi} + \frac{\mu}{2}\left[\dot{Q}^2 - \delta^2 Q^2\right] \qquad (1.2.1)$$

where E, B and Q are the electric, magnetic and phonon fields respectively. The phonon frequency is denoted by δ, its effective oscillator mass by μ, and its group velocity is taken to be zero. The dielectric constant is of the form

$$\epsilon = \epsilon_0 + \epsilon_2 E^2 + \gamma Q \qquad (1.2.2)$$

where ϵ_2 is the nonlinear coefficient and γ determines the coupling of the phonons to the electromagnetic field. In this analysis dispersion effects will be neglected, so ϵ_0 is simply a constant. Linear polarization will be assumed,

with the electromagnetic potential A and E along x-axis and B along y-axis, z-axis being the propagation direction. The vector potential is assumed to be the sum of a laser pulse at frequency ω_1

and a Stokes wave at frequency $\omega_2 = \omega_1 - \delta$:

$$A = a_1 e^{i(k_1 z - \omega_1 t)} + a_2 e^{i(k_2 z - \omega_2 t)} + c.c. \qquad (1.2.3)$$

We neglect the anti-Stokes and higher order Stokes waves. Furthermore, let the phonon field be described by

$$Q = q e^{i(k_0 z - \delta t)} + c.c. \qquad (1.2.4)$$

The wave vector of the phonon is such as to satisfy the phase matching $k_2 = k_1 - k_0$. The amplitudes a_1 , a_2 and q are assumed to be slowly varying functions of space and time and the variation of these amplitudes in the transverse directions are neglected.

Several approximations are made which reduce the problem's complexity. We insert Eqs. (3) and (4) into Eq. (1) and eliminate those terms which oscillate rapidly. We also discard those terms in the Lagrangian which are quadratic in spatial and/or temporal derivatives. Since $\epsilon_2 E^2$ and γQ are likely to be small, we shall eliminate terms involving these parameters when they multiply derivative terms. These approximations will be justified later. It

will finally be assumed that $\delta/\omega_1 \ll 1$, so[10] $\omega_1 \approx \omega_2 \equiv \omega$. The Lagrangian density then becomes:

$$\mathcal{L} = \frac{1}{8\pi}\left\{-\frac{4\epsilon_0\omega}{c^2}\,\mathrm{Im}\left[\dot{a}_1 a_1^* + \dot{a}_2 a_2^*\right] - 4k\,\mathrm{Im}\left[a_1' a_1^* + a_2' a_2^*\right]\right.$$

$$+\frac{4\gamma\omega^2}{c^2}\,\mathrm{Re}\left(Q a_2 a_1^*\right) - 2\mu\delta\,\mathrm{Im}\left(\dot{Q}Q^*\right) \qquad (1.2.5)$$

$$\left.+\frac{6\epsilon_2\omega^4}{c^4}\left[|a_1|^4 + |a_2|^4 + 4|a_1|^2|a_2|^2\right]\right\} \quad ,$$

where primes denote spatial derivatives and dots denote time derivatives. Use is made of the dispersion formula $k^2 = \frac{\epsilon_0\omega^2}{c^2}$. The Euler-Lagrange equations of motion yield

$$\dot{a}_1 + v a_1' = i\left[\frac{3\epsilon_2\omega^3}{\epsilon_0 c^2}\left(|a_1|^2 + 2|a_2|^2\right)a_1 + \frac{\gamma\omega}{2\epsilon_0} Q a_2\right] \quad , \qquad (1.2.6)$$

$$\dot{a}_2 + v a_2' = i\left[\frac{3\epsilon_2\omega^3}{\epsilon_0 c^2}\left(|a_2|^2 + 2|a_1|^2\right)a_2 + \frac{\gamma\omega}{2\epsilon_0} Q^* a_1\right] \quad , \qquad (1.2.7)$$

and

$$\dot{Q} = i\,\frac{\gamma\omega^2}{8\pi\mu c^2\delta}\,a_2^* a_1 \qquad (1.2.8a)$$

where $v = c/\sqrt{\epsilon_0}$. Let us introduce a phenomenological damping term into the last equation to account for the dephasing of the molecular vibrations due to collisions[11]. Thus Eq. (8a) becomes

$$\dot{Q} + \Gamma Q = i\,\frac{\gamma\omega^2}{8\pi\mu c^2\delta}\,a_2^* a_1 \quad . \qquad (1.2.8b)$$

Equations (6), (7) and (8b) are the fundamental equations of interest[12]. We notice three types of terms appearing on the right hand sides of Eqs. (6) and (7). The first term produces self-phase modulation. The second term we shall refer to as the cross-phase modulation term[13]. The third term is the stimulated Raman-parametric driving term. Thus, in solving Eqs. (6)-(8) we shall be studying the competition and interplay between these three nonlinear phenomena.

Let us introduce new dependent and independent variables and cast these equations into a more compact form, i.e.

$$\xi = \Gamma\left(t - \frac{z}{v}\right) \quad , \tag{1.2.9}$$

$$\eta = \Gamma \frac{z}{v} \quad , \tag{1.2.10}$$

$$\varphi = \frac{\gamma\omega}{2\Gamma\epsilon_0} q \quad , \tag{1.2.11}$$

$$\beta = \frac{3\epsilon_2\omega^3}{\Gamma\epsilon_0 c^2} \quad , \tag{1.2.12}$$

$$\delta = \frac{\gamma^2\omega^3}{16\pi\mu_0\epsilon_0 c^2\Gamma^2} \quad . \tag{1.2.13}$$

The equations (6)-(8) reduce to

$$\partial_\eta a_1 = i\left\{\beta\left[|a_1|^2+2|a_2|^2\right]a_1 + \varphi a_2\right\} , \tag{1.2.14}$$

$$\partial_\eta a_2 = i\left\{\beta\left[|a_2|^2+2|a_1|^2\right]a_2 + \varphi^* a_1\right\} , \tag{1.2.15}$$

$$\partial_\xi \varphi + \varphi = i\delta a_2^* a_1 \quad . \tag{1.2.16}$$

The boundary conditions at z=0 are: the input laser pulse a_1 is given by

$$a_1(\xi,\eta)\Big|_{\eta=0} = F(\xi) \quad , \tag{1.2.17}$$

and the Stokes wave a_2, is due to noise

$$a_2(\xi,\eta)\Big|_{\eta=0} = f(\xi) \quad . \tag{1.2.18}$$

In principle we can also specify a boundary condition[14] for φ, but this is unnecessary. The energy/area in the input pulse is given by

$$U = \frac{c}{4\pi}\int \overline{E^2}\,dt = \frac{\omega^2}{2\pi c}\int |F|^2 dt \tag{1.2.19}$$

For the special case of a Gaussian pulse

$$F(\xi) = F_0 \exp\left[-\left(\xi/\Gamma\tau\right)^2\right] \quad , \tag{1.2.20}$$

where τ is the pulse duration and F_0 is the amplitude of

the vector potential ($F_0 = cE_0/\omega$). Combining Eqs. (19) and (20) gives

$$F_0 = \left[\frac{cU\sqrt{8\pi}}{\omega^2 \varepsilon}\right]^{1/2} . \qquad (1.2.21)$$

A formal integration of Eq. (16) leads to

$$\varphi(\xi,\eta) = i\delta \int_{-\infty}^{\xi} d\xi'\, a_2^*(\xi',\eta)\, a_1(\xi',\eta)\, e^{\xi'-\xi} . \qquad (1.2.22)$$

Let us now assume that the time scale over which a_1 and a_2 vary is long compared with the phonon relaxation time. Due to the exponential in Eq. (22) we then get the main contribution to the integral from the vicinity $\xi' \approx \xi$ so

$$\varphi(\xi,\eta) \approx i\delta\, a_2^*(\xi,\eta)\, a_1(\xi,\eta) . \qquad (1.2.23)$$

Eqs. (14) and (15) become

$$\partial_\eta a_1 = i\beta\left[|a_1|^2 + 2|a_2|^2\right]a_1 - \delta\,|a_2|^2 a_1 \qquad (1.2.24)$$

and

$$\partial_\eta a_2 = i\beta\left[|a_2|^2 + 2|a_1|^2\right]a_2 + \delta\,|a_1|^2 a_2 . \qquad (1.2.25)$$

Let us express the field amplitudes as

$$a_1 = \sqrt{\rho_1}\, e^{i\phi_1} , \qquad (1.2.26)$$

$$a_2 = \sqrt{\varrho_2}\, e^{i\phi_2} \quad . \tag{1.2.27}$$

Then we find for the phases

$$\phi_1 = \int_0^\eta \beta\, d\eta' (\varrho_1 + 2\varrho_2) \quad , \tag{1.2.28}$$

and

$$\phi_2 = \int_0^\eta \beta\, d\eta' (\varrho_2 + 2\varrho_1) \quad . \tag{1.2.29}$$

The cross-phase modulation term broadens twice as fast as the self-phase modulation term. The magnitudes satisfy the Volterra-Lotka[15] equations:

$$\partial_\eta \varrho_1 = -2\delta \varrho_1 \varrho_2 \tag{1.2.30}$$

and

$$\partial_\eta \varrho_2 = 2\delta \varrho_1 \varrho_2 \tag{1.2.31}$$

which describes the predator-prey relatioship between the Stokes and incident pulses. These equations are solved by first noting that the sum of ϱ_1 and ϱ_2 is constant:

$$\varrho_1 + \varrho_2 = F^2(\xi) + f^2(\xi) \quad . \tag{1.2.32}$$

Then dividing Eq. (30) by ρ_1 and Eq. (31) by ρ_2 and subtracting gives, upon integration

$$\ln\left(\frac{\rho_2}{\rho_1}\right) = 2\delta\eta\left[F^2(\xi)+f^2(\xi)\right] + 2\ln\left(\frac{f}{F}\right) \qquad (1.2.33)$$

Hence

$$\rho_1 = \frac{F^2+f^2}{1+\left(\frac{f}{F}\right)^2 \exp\left[2\delta\eta(F^2+f^2)\right]} \quad , \qquad (1.2.34)$$

and

$$\rho_2 = \frac{F^2+f^2}{1+\left(\frac{F}{f}\right)^2 \exp\left[-2\delta\eta(F^2+f^2)\right]} \quad . \qquad (1.2.35)$$

Eqs. (26)-(29) and (34)-(35) provide an analytic solution to the problem. The intensity of the incident pulse decays while the intensity of the Stokes wave is parametrically amplified and ultimately saturates[16].

Let us restrict our attention to the case where $\rho_2 \ll \rho_1$. Then

$$\rho_1 \approx F^2(\xi) \quad , \qquad (1.2.36)$$

and

$$\rho_2 \approx f^2 \exp\left[2\delta\eta F^2(\xi)\right] \quad , \qquad (1.2.37)$$

so

$$a_1(\xi,\eta) = F(\xi)\, e^{i\beta\eta F^2(\xi)} \,, \tag{1.2.38}$$

$$a_2(\xi,\eta) = f(\xi)\, e^{i(2\beta - i\delta)\eta F^2(\xi)} \tag{1.2.39}$$

which we will use in the evaluation of the spectral characteristics of the process.

CHAPTER 1. Combined SRS and SPM

1.3 Results

In this section we present some theoretical predictions and results which are then applied to a specific example. The frequency spectrum of the electromagnetic field is proportional to

$$I(\omega) = \left| \int_{-\infty}^{\infty} d\xi \left[a_1 e^{i(\omega-\omega_1)\frac{\xi}{\Gamma}} + a_2 e^{i(\omega-\omega_1+\delta)\frac{\xi}{\Gamma}} \right] \right|^2 . \qquad (1.3.1)$$

The two terms will contribute independently, since F and f are incoherent. We shall focus our attention on the Raman term first. From Eqs. (2.40) and (2.20) one obtains

$$I_R = \left| \int_{-\infty}^{\infty} d\xi f(\xi) \exp i \left[\eta(2\beta - i\delta) F_o^2 e^{-2(\xi/\Gamma\tau)^2} + (\omega-\omega_1+\delta)\frac{\xi}{\Gamma} \right] \right|^2 . \qquad (1.3.2)$$

Let us neglect the time-dependence of f and expand the exponential in a power series:

$$I_R = \left| f\tau \sum_{n=0}^{\infty} \frac{[i\eta F_o^2(2\beta - i\delta)]^n}{n!} \sqrt{\frac{\pi}{2n}}\, e^{-\frac{(\omega-\omega_1+\delta)^2\tau^2}{8n}} \right|^2 \qquad (1.3.3)$$

For large η the important contributions to this sum stem from large values of n, so the sum may be replaced by an integral and Stirling's approximation may be employed. Thus

$$I_R = \left| \int_0^\infty \frac{f\mathcal{E}\Gamma}{2} \frac{dn}{n} \exp\left[\phi(n) - \frac{(\omega-\omega_1+\delta)^2\tau^2}{8n}\right]\right|^2 \qquad (1.3.4)$$

where

$$\phi(n) = n - n\ln n + n\ln\left[i\eta F_0^2(2\beta - i\delta)\right] . \qquad (1.3.5)$$

The point where ϕ is maximum is given by

$$n_0 = i\eta F_0^2(2\beta - i\delta) \qquad (1.3.6)$$

We first assume that the frequency shift satisfies the inequality

$$\left|\frac{(\omega-\omega_1+\delta)^2\tau^2}{8n_0}\right| \ll |n_0| \quad , \qquad (1.3.7)$$

so the second term in the exponential in Eq. (4) may be neglected. Then, from the method of steepest descent, we find

$$I_R = \frac{\pi}{2}\left(\frac{f\Gamma\mathcal{E}}{F_0}\right)^2 \frac{1}{\eta} \frac{e^{2\eta\delta F_0^2}}{\sqrt{4\beta^2+\delta^2}} \qquad (1.3.8)$$

For frequencies such that Eq. (7) is not satisfied, the

intensity is highly supressed due to the second exponential in Eq. (4). In this limit we are probing the far spectral wings of the pulse and the spectral intensity drops rapidly. Thus the Raman spectrum is predicted to be rather flat with its intensity given by Eq. (8) and a half width

$$\Delta\omega_R = \frac{\eta F_o^2}{\sigma}\sqrt{\beta(4\beta^2+\delta^2)} \quad . \tag{1.3.9}$$

For pulses of constant energy per unit area we see from Eqs. (2.21) and (9) that the Stokes width varies as the inverse square of the primary pulse duration. The broadening is due to a combination of the cross-phase modulation and the Raman terms.

Next let us examine the first term (laser) in Eq. (1). The analysis proceeds along much the same line, and we find

$$I_L(\omega) \approx \pi\Gamma^2\sigma^2/\beta\eta \tag{1.3.10}$$

for frequencies such that

$$\Delta\omega_L \equiv |\omega-\omega_1| \lesssim \frac{2\beta\eta F_o^2}{\sigma} \tag{1.3.11}$$

and I_L drops rapidly for frequencies greater than this value. This is the conventional self-phase modulation result.

In summary the theory predicts that the primary laser

pulse will undergo self-phase modulation . The Stokes pulse will be parametrically amplified from the noise. The Stokes pulse broadens due to a combination of cross-phase modulation and Raman parametric amplification. The spectral pulse widths increase as the pulses progress down the fluid and grow dramatically as the duration of the pulse is shortened. Theory predicts a ratio of the Stokes spectral width to the primary spectral width given by

$$\frac{\Delta\omega_R}{\Delta\omega_L} = \left\{ 8 \left[1 + \left(\delta/2\beta \right)^2 \right] \right\}^{1/2} \qquad (1.3.12)$$

Thus by the time self-phase modulation has broadened the primary pulse to the point where it overlaps the Stokes line, little trace of a distinct Raman Stokes line should be observed.

As a numerical example, let us compute $\Delta\omega_R$ and $\Delta\omega_L$ for two typical liquids, CS and ethanol, in the field of Neodynium glass laser. The parameters relevant to such a system are tabulated in Table 1, along with the computed values of $\Delta\omega_R$ and $\Delta\omega_L$. These values should be readily observable in an experiment.

Table I. Parameters for liquids in a typical laser field (estimated values)

Quantity	Value in c.g.s. units	
	CS_2	ethanol
ϵ_0	1.7	1.3
ϵ_2	10^{-11}	10^{-12}
ω	3.6×10^{15} (1.9×10^{4} cm^{-1})	3.6×10^{15} (1.9×10^{4} cm^{-1})
σ	1.2×10^{14} (656 cm^{-1})	5.5×10^{14} (2928 cm^{-1})
μ	1.3	.81
Γ	9.4×10^{10} (.5 cm^{-1})	3.4×10^{12} (18 cm^{-1})
γ	10^{6} (estimated)	10^{6} (estimated)
τ	6×10^{-12}	6×10^{-12}
F_0	.1	.1
υ	2.3×10^{10}	2.6×10^{10}
N	10	10
$\Delta\omega_R$	3.8×10^{15} (1.9×10^{4} cm^{-1})	4.4×10^{14} (2.2×10^{3} cm^{-1})
$\Delta\omega_L$	13.3×10^{14} (7.1×10^{3} cm^{-1})	1.5×10^{13} (810 cm^{-1})

Figure captions

Figure 1. Sketch showing the distortion of the wave front and self-focusing of a laser beam in a NL medium.

Figure 2. Sketch showing transient self-focusing of a laser pulse in a NL medium. Such a focusing is always accompanied with a broad phase modulation. Different parts of the pulse (a,b,c, etc.) focus and defocus along different ray paths. The pulse first gets deformed into a horn shape and then propagates on without much further change.

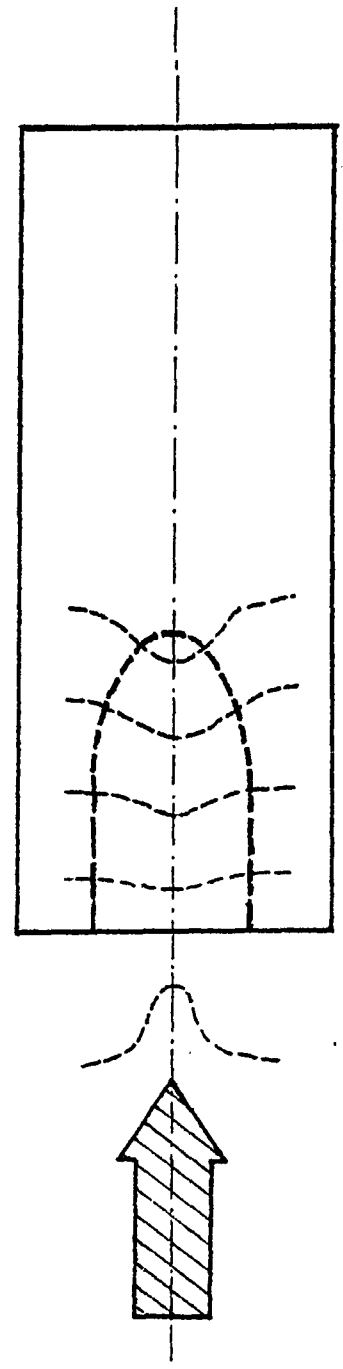


Figure 1.

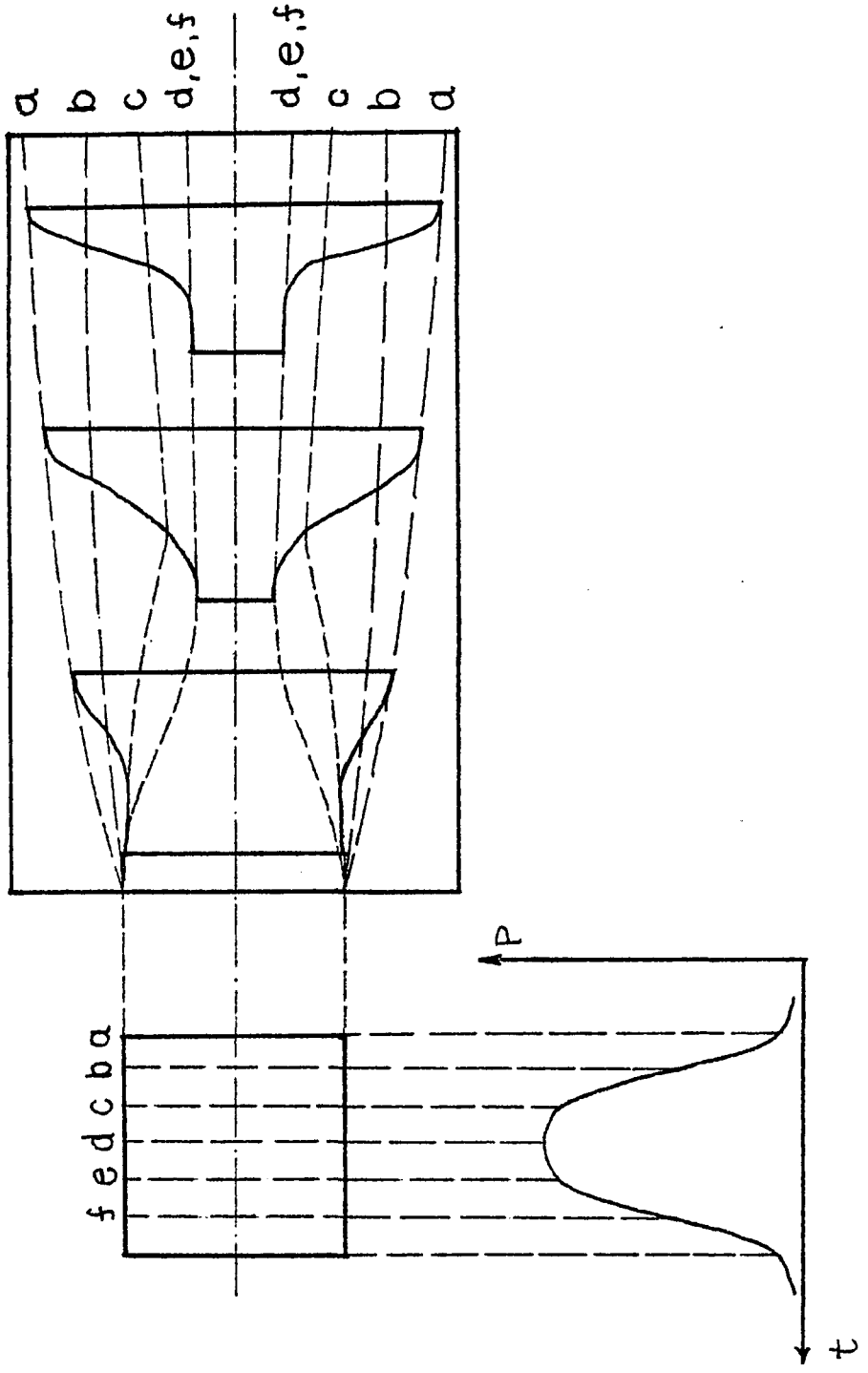


Figure 2.

References

1. S. L. Shapiro, J. A. Giordmaine and K. W. Wecht, PRL 19, 1093 (1967); G. G. Bret and H. P. Weber, IEEE J. Q. El. QE-4, 807 (1968); D. Van der Linde, M. Maier and W. Kaiser, Phys.Rev. 178,11(1969);O. Rahn, M. Maier and W. Kaiser Opt. Comm. 1, 109 (1969); M. Colles, Opt. Comm. 1, 169 (1969); R. Carman, M. Mack, F. Shimuzu and N. Bloembergen, Phys. Rev. Lett. 23, 1327 (1969).

2. R. L. Carman, F. Shimuzu, C. S. Wang and N. Bloembergen, Phys. Rev. A2, 60 (1970)

3. R. G. Brewer, Phys. Rev. Lett. 19, 8 (1967); F. Shimuzu, PRL 19, 1097 (1969); R. R. Alfano and S. L. Shapiro, Phys. Rev. Lett. 24, 584 (1970); 24, 592 (1970); 24, 1217 (1970); Phys. Rev. A6, 433 (1972).

4. Y. R. Shen and N. Bloembergen, Phys. Rev. 137, A1786 (1965) W. Kaiser, and M. Maier in "Laser Handbook", ed. by F. Arecchi and E. O. Schulz-Dubois (North-Holland, Amsterdam 1972); Y. R. Shen in "Raman Scattering in Solids", ed. by M. Cardona (Springer, Berlin 1975).

5. See, for example Heitler , "The Quantum Theory of Radiation", (Cambridge University, New York 1954), p. 192.

6. C. C. Wang, Phys. Rev. Lett. 16, 344 (1966); Phys. Rev. Lett. 152, 149 (1966); J. Appl. Phys. 37, 1943 (1966)

7. S. A. Akhmanov, A. P. Sukhorukov, and R. V. Khokhlov, Sov. Phys.-Usp. 93, 609 (1968); U. Lugovoi, and A. Prokhorov, Sov. Phys.-Usp. 16, 658 (1974).

8. R. R. Alfano, GTE Lab Technical Report TR72-230.1 (1972)

and references cited therein.

9. J. I. Gersten, R. R. Alfano, and M. Belic, to be published in Phys. Rev. A, (1980)

10. This approximation carries with it a price: energy conservation can no longer be applied to the system.

11. Damping terms may be introduced in a systematic manner through the Rayleigh dissipation function. If conductivity effects are neglected no damping occurs in Eq. (6) or (7).

12. The above set of equations is not adequate to discuss fluctuation phenomena due to the neglect of the Langevin terms on the right hand sides of Eqs. (6)-(8). We omit a discussion of fluctuations here.

13. This term has been studied by Maker and Terhune, Phys. Rev. 137, 807 (1965)

14. Alternatively, we can assume that $\mathcal{S}$ is of thermal origin. We choose the former possibility for the sake of simplicity.

15. A. Lotka, "Elements of Physical Biology",(Baltimore, 1925) V. Volterra, "Lecons sur la Theorie Mathematique de la Lutte pour la Vie", (Gauthier-Villars, Paris, 1937).

16. Of course higher Stokes processes then become important and energy is partitioned among many modes. This effect is omitted in our set of approximations.

CHAPTER 2. Unstable Resonators for High Power Lasers

2.1 Introduction

Optical resonators are needed for three principle (and related) purposes:

(a) to build up large field intensities at specified (resonance) frequencies,

(b) to act as efficient energy storage devices, and

(c) to act as spatial and frequency filters.

The chief difference with lower frequency resonators is that at longer wavelengths it is possible to restrict the number of resonances to one or few by choosing cavity dimensions comparable to the wavelength. At optical frequencies this is not feasible - an enclosed optical resonator would contain billions of modes per unit resonator volume. Thus special techniques must be invoked to suppress unwanted optical modes. The simplest one is the use of open resonators.

The earliest proposal for using open resonators seems to have been that of Dicke[1]. Schawlow and Townes[2] suggested in their original laser article that such a resonator will

discriminate heavily against higher-order transverse modes. The initial paper on unstable resonators by Siegman in 1965[3] asserted that "there are ... good reasons for believing that ... unstable resonators will be of practical utility in laser applications". A subsequent paper in 1967[4] listed three of these reasons as:

(a) large mode volumes,

(b) adjustible diffraction output coupling, and

(c) substantial discrimination against higher order transverse modes.

Typical examples of a stable and an unstable optical resonator configuration are illustrated in Fig. 1. The stable resonator, whose mirror configuration corresponds to a stable periodic focusing system has a long slender Gaussian-profile lowest-order mode whose diameter is generally less than the diameter of the laser mirrors themselves. Since the area ratio of the laser medium cross section to the lowest-mode cross section is of the same order as the Fresnel number N_F , the lowest-order mode will extract only a fraction $1/N_F$ of the energy available in the laser medium. Consequently, if the Fresnel number is much larger than unity, the laser must oscillate in a sizable number of higher-order modes in order to exctract a sizable portion of the energy from the laser medium. This is hardly desirable from the frequency selection point of view. The lowest-order mode in the unstable resonator, by contrast fills the entire cross section of at least one of the laser

mirrors, however large it may be.

Some predecessors of the unstable resonator concept include the ball mirror resonator of Okaya[5], a proposal by Mollenauer et al.[6] to use a divergent mirror at one end of a ruby rod, etc. LaTourrette and colleagues[7] first introduced the important concept of diffraction coupling. Fox and Li in one instance applied their Huygens integral method to calculate the losses of some unstable mirror configurations.[8]

The most widely used - confocal or telescopic unstable resonator was introduced by Anan'ev et al.[9] and by Krupke and Sooy[10]. Very significant contributions to unstable resonator understanding, both conceptual and experimental have been made by Soviet workers, notably Anan'ev, as summarized in his review article[11]. Another excellent review article on unstable resonators is written by Siegman.[12]

In this chapter a numerical analysis of unstable loaded optical resonators is presented. An analysis of this kind is very important in the design and construction of high-power lasers. We carried out our analysis on a resonator model which closely resembles the actual experimental situation. The model is depicted in Fig. 1. It is a positive-branch confocal three dimensional resonator with square mirrors. The laser output is taken as a diffraction-coupled beam passing around the smaller (output) mirror. Between the mirrors is the gain medium which we treated in the thin-

gain-sheet approximation. In this approximation the gain medium is broken into a number of gain stations. The propagating field as it passes through these stations is modified by the appropriate gain formula.

The propagating field is a solution to the wave equation in the paraxial approximation. Actually, a system of wave equations for the forward and backward fields is solved simultaneously using fast Fourier transform (FFT) methods. Both forward and backward fields figure in the gain formula, which couples the wave equations. We let the main, forward field repeatedly bounce off the mirrors, constantly modifying it for the gain, until it stabilized into a single-mode oscillation. This stabilized field distribution, as it passes the output mirror, is shown in Figs. 2, 3, and 4 for various sets of parameters.

In review , the propagation code which we have developed[13] contains the following new features to obtain an accurate resonator field distribution.

(a) A continuous Fourier transform (CFT) algorithm is used in place of FFT to propagate from the small mirror to the gain medium. This step is crucial for sharp edge mirrors.

(b) We have reduced the core requirements by an order of magnitude by simultaneous integration of forward and backward waves and by use of the symmetry of the problem.

(c) A modified gain formula is used to include interference between forward and backward waves. The numerical results indicate that the interference effects exceed 10%.

The propagation code is sufficiently modest in core and time requirements to be run on a DEC 10 computer.

This chapter is organized in the following way. Section 2 presents the theoretical background - paraxial wave optics of the laser medium with homogeneous broadening. Section 3 introduces the resonator model to be dealt with. In section 4 an earlier attempt on this problem is presented. Firstly we tried to use standard difference methods in solving the wave equations at hand, and we encountered problems we could not overcome. However, we kept this difference code as an independent check on the one-dimensional fast Fourier transform method which we used later. In section 5 the new FFT algorithm is developed and applied to the full three dimensional case. In this, central section all facets of the problem were addressed. Finally, in section 6 we discuss results.

CHAPTER 2. High Power Laser

2.2 Theory[14]

The conventional starting point is the Maxwell equations

$$\left.\begin{aligned} \partial_{\vec{r}} \times \vec{E} &= i\omega\mu_0 \vec{H} \\ \partial_{\vec{r}} \times \vec{H} &= -i\omega\epsilon_0 \varkappa \vec{E} \end{aligned}\right\} \tag{2.2.1}$$

$$\partial_{\vec{r}} \cdot (\varkappa \vec{E}) = \rho \quad , \quad \partial_{\vec{r}} \cdot \vec{H} = 0 \tag{2.2.2}$$

for a monochromatic wave varying as $\exp(-i\omega t)$. $\varkappa$ is the dielectric coefficient,

$$\varkappa = \epsilon/\epsilon_0 + i\,\sigma/\omega\epsilon_0 \tag{2.2.3}$$

From Eq. (1) and (2) it follows

$$\partial_{\vec{r}} \times (\partial_{\vec{r}} \times \vec{E}) = \left(\frac{\omega}{c}\right)^2 \varkappa \vec{E} \tag{2.2.4}$$

- the wave equation to be solved.

Let us represent $\vec{E}$ as

$$\vec{E} = \vec{E}_T + \hat{z} E_z \equiv e^{ikz} (\vec{F}_T + \hat{z} F_z) \tag{2.2.5}$$

so that the wave equation becomes:

$$\partial_T (\partial_T \cdot \vec{F}_T + \partial_z F_z + ikF_z) - \partial_T^2 \vec{F}_T - \partial_z^2 \vec{F}_T - 2ik\partial_z \vec{F}_T + k^2 \vec{F}_T = \left(\frac{\omega}{c}\right)^2 \varepsilon \vec{F}_T$$

$$\partial_z (\partial_T \cdot \vec{F}_T) + ik \partial_T \cdot \vec{F}_T - \partial_T^2 F_z = \left(\frac{\omega}{c}\right)^2 \varepsilon F_z \tag{2.2.6}$$

If we are concerned with Gaussian-like beams, then there is some characteristic width w_o in the transverse direction. For such a beam there is also a diffraction length l,

$$l = k w_o^2 \tag{2.2.7}$$

which is a characteristic length in the longitudinal direction. Accordingly, we proceed to scale Eqs. (6). Let

$$x = w_o \xi \quad , \quad y = w_o \eta \quad , \quad z = l \zeta \tag{2.2.8}$$

Call

$$f \equiv \frac{w_o}{l} = \frac{1}{k w_o} \tag{2.2.9}$$

Then

$$\left.\begin{aligned} &\partial_{\vec{\xi}}\left(f\partial_{\vec{\xi}}\cdot\vec{F}_\xi + f^2\partial_\xi F_\xi + iF_\xi\right) - f\partial_{\vec{\xi}}^2\vec{F}_\xi - f^3\partial_\xi^2\vec{F}_\xi - 2if\partial_\xi\vec{F}_\xi = \\ &= f\left[\left(\frac{\omega}{c}\right)^2\mathcal{K} - k^2\right]w_o^2\vec{F}_\xi \\ &f^3\partial_\xi\left(\partial_{\vec{\xi}}\cdot\vec{F}_\xi\right) + if\partial_{\vec{\xi}}\cdot\vec{F}_\xi - f^2\partial_{\vec{\xi}}^2 F_\xi = f^2\left(\frac{\omega}{c}\right)^2\mathcal{K}w_o^2 F_\xi \end{aligned}\right\} \qquad (2.2.10)$$

where

$$\vec{F}_T(\vec{r}_T, z) \to \vec{F}_\xi(\vec{\rho}, \xi) \quad , \quad F_z(\vec{r}_T, z) \to F_\xi(\vec{\rho}, \xi) \quad , \qquad (2.2.11)$$

and

$$\partial_{\vec{\xi}} = \hat{x}\partial_\xi + \hat{y}\partial_\eta \quad . \qquad (2.2.12)$$

Let us consider a homogeneously broadened two level laser amplifier medium with $\delta = 0$ and

$$\mathcal{K}/\mathcal{K}_L = 1 + gk^{-1}\frac{\Omega - i}{1+\Omega^2+I} \equiv 1 + gk^{-1}m(I) \qquad (2.2.13)$$

where $\mathcal{K}_L$ is the linear dielectric constant, $k = \sqrt{\mathcal{K}_L}\,\frac{\omega}{c}$, $\Omega = \frac{\omega - \omega_{ab}}{\gamma_{ab}}$ is the detuning parameter, γ_{ab} is the homogeneous linewidth, g is the on-resonance signal gain per meter and I is the field intensity. For this model Eqs. (10) reduce to

$$\partial_{\vec{\xi}}\left(f\partial_{\vec{\xi}}\cdot\vec{F}_\xi + f^2\partial_\xi F_\xi + iF_\xi\right) - f\partial_{\vec{\xi}}^2\vec{F}_\xi - f^3\partial_\xi^2\vec{F}_\xi - 2if\partial_\xi\vec{F}_\xi = fgLm\vec{F}_\xi$$

$$f^3\partial_\xi(\partial_{\vec{\xi}}\cdot\vec{F}_\xi)+if\partial_{\vec{\xi}}\cdot\vec{F}_\xi-f^2\partial^2_{\vec{\xi}}F_\xi=\left[1+f^2glm\right]F_\xi \qquad (2.2.14)$$

Since $f \ll 1$ always, let us expand the field in powers of f. Only alternate powers are found necessary :

$$\vec{F}_\xi=\vec{F}^{(0)}_\xi+f^2\vec{F}^{(2)}_\xi+\cdots \quad , \quad F_\xi=fF^{(1)}_\xi+f^3F^{(3)}_\xi+\cdots \qquad (2.2.15)$$

In the lowest order Eqs. (14) become

$$\left.\begin{aligned}\partial^2_{\vec{\xi}}\vec{F}^{(0)}_\xi+2i\partial_\xi\vec{F}^{(0)}_\xi&=-glm_0\vec{F}^{(0)}_\xi \ ,\\ F^{(1)}_\xi&=i\partial_{\vec{\xi}}\cdot\vec{F}^{(0)}_\xi \ ,\end{aligned}\right\} \qquad (2.2.16)$$

where

$$m_0=\frac{\Omega-i}{1+\Omega^2+I^{(0)}}=m\left(|\vec{F}^{(0)}_\xi|^2\right) \qquad (2.2.17)$$

is the form appropriate to homogeneous broadening. In the unscaled variables Eqs. (16) read

$$\left.\begin{aligned}\partial^2_{\vec{T}}\vec{F}^{(0)}_T+2ik\partial_z\vec{F}^{(0)}_T&=-gkm_0\vec{F}^{(0)}_T \ ,\\ kF^{(1)}_z&=i\partial_{\vec{T}}\cdot\vec{F}^{(0)}_T \ .\end{aligned}\right\} \qquad (2.2.18)$$

These are the wave equations in the paraxial approximation. With no loss of generality, let us look for plane polarized solutions of the form

$$\vec{F}_{\sigma}^{(o)} = e^{iS}\vec{\Psi} \quad . \tag{2.2.19}$$

When we use this in (16a), two equations follow, for the real and imaginary parts

$$\left.\begin{aligned} \partial_{\vec{\xi}}^2\vec{\Psi} - (\partial_{\vec{\xi}}S)^2\vec{\Psi} - 2\partial_\xi S\,\vec{\Psi} &= -g\ell(Re\,m_o)\vec{\Psi} \,, \\ 2(\partial_{\vec{\xi}}S\cdot\partial_{\vec{\xi}})\vec{\Psi} + \partial_{\vec{\xi}}^2 S\,\vec{\Psi} + 2\partial_\xi\vec{\Psi} &= -g\ell(Im\,m_o)\,\vec{\Psi} \,. \end{aligned}\right\} \tag{2.2.20}$$

If we take scalar products with $\vec{\Psi}$ we obtain

$$(\partial_{\vec{\xi}}S)^2 + 2\partial_\xi S = g\ell\, Re\, m_o + \vec{\Psi}\cdot\partial_{\vec{\xi}}^2\vec{\Psi}\;\Psi^{-2} \quad , \tag{2.2.21a}$$

which is the eikonal equation, and

$$\partial_{\vec{\xi}}\cdot(\Psi^2\partial_{\vec{\xi}}S) + \partial_\xi\Psi^2 = -g\ell\,(Im\,m_o)\,\Psi^2 \quad , \tag{2.2.21b}$$

which is the energy transport equation. Rays normal to surfaces of constant phase obey

$$\frac{d\vec{\rho}}{d\xi} = \partial_{\vec{\xi}}S \quad , \tag{2.2.22}$$

so that if U is any function of $\vec{\rho}$ and ξ, we have

$$\frac{dU}{d\xi} = \frac{d\vec{\rho}}{d\xi}\cdot\partial_{\vec{\xi}}U + \partial_\xi U = \partial_{\vec{\xi}}S\cdot\partial_{\vec{\xi}}U + \partial_\xi U \tag{2.2.23}$$

and also

$$\frac{d^2\vec{\xi}}{d\xi^2} = (\partial_{\vec{\xi}} S \cdot \partial_{\vec{\xi}})\partial_{\vec{\xi}} S + \partial_\xi(\partial_{\vec{\xi}} S) =$$

$$\frac{1}{2}\left[\partial_{\vec{\xi}}(\partial_{\vec{\xi}} S)^2 + 2\,\partial_{\vec{\xi}}(\partial_\xi S)\right] = \frac{1}{2}\partial_{\vec{\xi}}\left[(\partial_{\vec{\xi}} S)^2 + 2\partial_\xi S\right] \quad . \tag{2.2.24}$$

Using these formulae and $\Psi^2 = I$ and cylindrical coordinates (axial symmetry assumed), we obtain for the eikonal, energy, and ray equation

$$\left.\begin{aligned} &(\partial_\rho S)^2 + 2\partial_\xi S = \vartheta L \,\mathrm{Re}\, m_0(I) + \frac{1}{\rho\sqrt{I}}\partial_\rho\left(\rho\partial_\rho\sqrt{I}\right), \\ &\frac{1}{\rho}\partial_\rho\left(I\rho\partial_\rho S\right) + \partial_\xi I = -\vartheta L\, I\, \mathrm{Im}\, m_0(I) \quad , \\ &\frac{d\rho}{d\xi} = \partial_\rho S \quad . \end{aligned}\right\} \tag{2.2.25}$$

As an example, let us solve Eqs. (25a) and (25b) for an empty spherical resonator. Then $m_0(I)=0$, and assume $\vec{F}_{\vec{\xi}}^{(0)} = (\sqrt{I}e^{iS}, 0, 0)$, the field polarized in x-direction. Exact solutions may be written as

$$S_\ell = \frac{\rho^2}{R(\xi)} - (2\ell+1)\tan^{-1} 2\xi \quad , \tag{2.2.26}$$

$$I_\ell = \frac{1}{W^2(\xi)}\, e^{-\frac{2\rho^2}{W^2}}\, L_\ell^2\left(\frac{2\rho^2}{W^2}\right) \quad , \tag{2.2.27}$$

where

$$W^2(\xi) = 1+(2\xi)^2 \quad , \quad R(\xi) = \frac{W^2(\xi)}{2\xi} \tag{2.2.28}$$

and $L_\ell(x)$ are the Laguerre polynomials. The ray equation becomes

$$\frac{d\varrho}{d\xi} = \frac{4\xi\varrho}{1+(2\xi)^2} \quad , \tag{2.2.29}$$

which yields the hyperbolic rays

$$\varrho = \varrho_0 \sqrt{1+(2\xi)^2} \quad , \tag{2.2.30}$$

where ϱ_0 is the radius of the ray at $\xi = 0$.

For the resonator with an active medium we must go to the numerical investigation. Before going into details of numerical methods applied, let us cast Eqs. (25) into a more convenient form. Introduce a set of new variables:

$$K \equiv \ln I \quad , Z \equiv g l \xi \equiv \frac{1}{\epsilon}\xi \quad , Q \equiv \varrho^2 \quad , \sigma \equiv \frac{1}{\varrho}\partial_\varrho S \tag{2.2.31}$$

After some algebra, the equations become

$$\begin{aligned} \frac{dK}{dZ} &= -\mathrm{Im}\, m_0(K) - 2\epsilon\partial_Q(Q\sigma) \quad , \\ \frac{d\sigma}{dz} &= -\epsilon\sigma^2 + \partial_Q P \quad , \\ P &\equiv \mathrm{Re}\, m_0 + \epsilon\left[(Q\partial_Q K+2)\partial_Q K + 2Q\partial_Q^2 K\right] \end{aligned} \tag{2.2.32}$$

$$\frac{dQ}{dZ} = 2\epsilon Q\zeta$$

These are the equations in the moving frame, $Q=Q(Z)$. In the fixed frame Q is Z-independent and equations get the form

$$\left.\begin{aligned} \partial_z K &= -\mathrm{Im}\, m_0 - 2\epsilon\left[\partial_Q(Q\zeta) + Q\zeta\partial_Q K\right], \\ \partial_z \zeta &= \partial_Q \mathrm{Re}\, m_0 + \epsilon\partial_Q\left[(Q\partial_Q K+2)\partial_Q K + 2Q\partial_Q^2 K - Q\zeta^2\right]. \end{aligned}\right\} \quad (2.2.33)$$

CHAPTER 2. High Power Laser

2.3 Resonator Model

In this analysis of unstable resonators our purpose is twofold,

(a) to further the theoretical understanding and numerical handling of unstable resonators, and

(b) to explain, at least qualitatively some of the available experimenal data.

Naturally, to this end, the resonator model chosen here will resemble as closely as possible the actual experimental setup. We will consider an open three dimensional unstable optical resonator consisting of two spherical mirrors coaxially placed at a distance d appart. The mirrors are rectangular in shape and unequal in size. Arbitrarily, we set the smaller one to be on the left-hand side and convex, with the radius of curvature $R_1 < 0$, and the transverse dimensions a_1 and b_1. Correspondingly, the other, larger, RHS mirror is of the radius $R_2 > 0$ and transverse dimensions a_2 and b_2 . In between the mirrors there is

either an active, lasing medium or vacuum (bare resonator). The active medium is chosen to extend from the distance d_1 away from the small mirror to the distance d_2 away from the large mirror, its thickness being $d-(d_1 + d_2)$. If only one transverse dimension of the mirror is mentioned (eg. a_1 or a_2), we will have a square mirror in mind, without explicitly saying so.

The unstable resonator is a divergent periodic focusing system - the lowest order mode expands on repeated bounces to fill the entire cross section of at least one of the laser mirrors, however large it may be. The laser output is taken as a diffraction-coupled beam passing around rather than through the output (small) mirror. As of the present day situation, it seems that the unstable resonators have proven their usefulness and utility in laser applications, based on the following facts : Unstable resonators

(a) can have large mode volume even in very short resonators,

(b) can have very substantial discrimination against higher--order transverse modes,

(c) provide efficient energy extraction from a wide variety of high-energy and high-power, CW and pulsed lasers, and

(d) allow for adjustible (large) diffraction output couplings.

It now seems clear that for any type of laser system characterized by large gain per pass and by a large Fresnel number, the optimum laser cavity will be an unstable resonator - it provides the best practical method for nearly

complete energy extraction from the laser medium, combined with high beam quality and a practical resonator structure.

Let us define some basic parameters[15] for the unstable resonators. The g-parameters are given with

$$g_1 = 1 - \frac{d}{R_1} \quad , \quad g_2 = 1 - \frac{d}{R_2} \tag{2.3.1}$$

In order for resonator to be unstable it must be that

$$g_1 g_2 < 0 \quad \text{or} \quad g_1 g_2 > 1 \quad , \tag{2.3.2}$$

conditions arrived at by a simple geometrical argument.[16] Another important resonator parameter arising from the geometrical analysis is the round-trip magnification M,

$$M = \left| \frac{g_1}{g_2} \right| \tag{2.3.3}$$

Descriptively, M corresponds to the rate of increase of the (linear) beam-spot size in one round-trip cycle. The geometrically predicted loss of the unstable resonator arises from the output coupling of the magnified wave past one (or both) of the resonator mirrors. This geometrical coupling ratio is independent of the mirror sizes or Fresnel number and depends only on M. For three-dimensional spherical mirrors it is

$$C = 1 - \frac{1}{M^2} \tag{2.3.4}$$

Another universal parameter is the Fresnel number

$$2\pi N_F = \frac{L}{d} \tag{2.3.5}$$

where $l=ka_1^2$ is the diffraction length connected with some resonator mode. The most useful of all appears to be the confocal (or telescopic) unstable resonator, primarily because it automatically produces a collimated output beam. According to the position of the common focus, the confocal resonators are either negative-branch or positive-branch. For negative-branch the focus is inside the resonator, for positive-branch it is outside. The negative-branch confocal configuration has significant practical advantages in the form of more easily obtainable shorter-radius mirrors and easier mirror alignment tolerances. However, the positive branch resonator seems to be universally employed in practice because the internal focal point in the negative-branch case leads to unacceptible difficulties with optical breakdown. In this analysis we will be mainly interested in the positive-branch confocal unstable resonators.

Since for the confocal resonator with spherical mirrors

$$R_1 + R_2 = 2d \tag{2.3.6}$$

it follows

$$M = -\frac{R_2}{R_1} \quad , \tag{2.3.7}$$

and

$$g_1 = \frac{M+1}{2} \quad , \quad g_2 = \frac{M+1}{2M} \tag{2.3.8}$$

In order that all the radiation incident on the large mirror be returned to the resonator, its minimal size must be

$$a_2 = M a_1 \tag{2.3.9}$$

CHAPTER 2. High Power Laser

2.4 Modified Predictor-Corrector

The system of equations of interest (2.32) or (2.33) is a system of nonlinear partial differential equations. At present there is no mathematical method which can treat it analytically. Therefore we resorted to numerical methods. The first method we tried was the Hammings modified predictor-corrector method for solving systems of first order differential equations of the form

$$Y' = f(Y, Z) \tag{2.4.1}$$

with given initial value $Y_0 = Y(Z_0)$. Z is the independent variable and Y is the unknown function vector. The version we used was the fourth order integration procedure that requires the evaluation of the right-hand side of the system (1) only two times per step. This is a great advantage compared with other methods of the same order of accuracy, for example the Runge-Kutta method which requires the evaluation of the right-hand side four times per step.

Another advantage is that at each step the evaluation procedure gives an estimate for the local truncation error; thus the procedure is able, without a significant amount of calculation time to choose and change the step size h.

On the other hand, Hammings predictor-corrector method (HPC) itself is not self-starting; that is the functional values at a single point are not enough to get the functional values ahead. Therefore, to obtain the starting values a special Runge-Kutta procedure followed by one iteration step is added. Starting HPC method requires the function and derivative values at four equidistant points, that is Z_0, Z_0+h, Z_0+2h, Z_0+3h. The values Y_0 and $Y_0'=f(Z_0, Y_0)$ are specified by the input; the rest is found by this special procedure. Thus, knowing the result at the equidistant points Z_{j-3}, Z_{j-2}, Z_{j-1}, Z_j, the result at Z_{j+1} is first predicted, then modified, then corrected. The final result includes the predicted and corrected values. Then the whole procedure repeats for as many steps as needed. The difference between predicted and corrected value, which is available at each step is also the measure for the local truncation error.

The entire input of the procedure is

(a) Lower and upper bound of the integration interval, initial step size h of the independent variable, and upper bound for the local truncation error.

(b) Initial values of the dependent variables and the number

of equations N.

(c) As external subroutine subprograms, the computation of the RHS of the system is required; for the flexibility in output, an output subroutine is required.

For better accuracy we worked in double precision. In actual computations we interchangably used the system library version DHPCG and Professor Lax's modified version DHPCU which we found more convenient. The results were identical. In the modified version the starting procedure is simplified, resulting in a faster and less core space consuming code. The step size changing allowance has been dropped (often you dont want to change your step size) and the entire code has been rewritten in ratfor, making it easily readable and understandable.

As for the programming considerations, we assumed that the number of equations is a multiple of 3. Then we reserved the first third locations in the vector Y for K(Z,Q), the second third for Q(Z) and the last third for $\dot{Q}$(Z,Q). Thus we got N/3 transverse (Q) points, i.e. we considered N/3 rays. As it is seen from (2.32), the RHS of our system involves derivatives with respect to Q. We needed therefore a reasonable number of transverse points, for a reasonable accuracy in the first and second derivatives. It turned out that even 22 transverse points was not unreasonable. To evaluate the derivatives we used DDGT3 subroutine from the system library, which employs the Lagrangian interpolation

polynomial of degree 2. It consequently uses three succesive points to evaluate the derivative at the middle point. In the beginning and at the end the formula is modified. Since later on we encountered stability problems, we decided to smooth out all arrays which are input to the derivative routine. As the smoothing procedure we applied a five point least square fitting to a parabola. We placed the center of the parabola at the mean of these five points

$$\bar{z} = \frac{1}{5}(z_1 + z_2 + z_3 + z_4 + z_5) \tag{2.4.2}$$

Let

$$f(z) \equiv a(z-\bar{z})^2 + b(z-\bar{z}) + c\ , \tag{2.4.3}$$

and

$$f_i = a(z_i-\bar{z})^2 + b(z_i-\bar{z}) + c\ . \tag{2.4.4}$$

The least square fitting method requires that

$$P \equiv \sum_{i=1}^{5} (f_i - Y_i)^2 \tag{2.4.5}$$

be a minimum. Y_i are the functional points which correspond to Z_i. The minimum conditions

$$\partial_a P = \partial_b P = \partial_c P = 0 \tag{2.4.6}$$

are then used to determine a, b and c. If we define

$$\mu_n \equiv \frac{1}{5}\sum_{i=1}^{5}(z_i-\bar{z})^n \tag{2.4.7}$$

and

$$\nu_n \equiv \frac{1}{5}\sum_{i=1}^{5}Y_i(z_i-\bar{z})^n \tag{2.4.8}$$

then the conditions are

$$\begin{pmatrix} \mu_4 & \mu_3 & \mu_2 \\ \mu_3 & \mu_2 & \mu_1 \\ \mu_2 & \mu_1 & \mu_0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \nu_2 \\ \nu_1 \\ \nu_0 \end{pmatrix}, \tag{2.4.9}$$

which is easily solved for a, b, c.

The effect of smoothing was favorable, although the stability problem was not removed. In the end it was this problem as well as the difficulties in applying the algorithm to the three dimensional case that made us change the whole numerical method. However, we kept the predictor-corrector method as an independent check on the new, FFT method. It turned out that they agree excellently for identical sets of parameters.

Another related problem that we succesfully overcame with the smoothing was the Nyquist criterion. As it is well known, when solving a differential equation which contains, say, a first order derivative in z-variable and a second order derivative in x-variable, then there is a condition (called the Nyquist criterion) on the grid increments in x and z directions of the form

$$C(\Delta x)^2 \geq \Delta z \qquad (2.4.10)$$

where C is a constant of order unity. In other words, for a reasonable accuracy (not too few transverse points) the step in z-direction is supposed to be small. This, of course makes the code time consuming, which is hardly desirable. A suppression of instabilities introduced by the smoothing procedure allowed us to increase the constant C to the order of hundred (for a similar number of steps) and enabled us to propagate much further along the z-axis.

We wrote two sets of programs to accomodate both the Lagrangian (moving frame) and Eulerian (fixed frame) sets of equations. We run those programs both forward and backward. As we ran from Z=0 to some Z and then backwards to Z=0, the Lagrangian programs reproduced the input at z=0 exactly, while the Eulerian did not. In general, this is not hard to explain: as the wave propagates, it expands while (in the Eulerian case) the transverse grid extension remains fixed -

some information is being lost. Therefore the initial information cannot be retreived completely.

CHAPTER 2. High Power Laser

2.5 Fast Fourier Transform

Fourier transform (FT) is an alternative method to difference systems in solving differential equations. By way of an example, let us solve the bare resonator paraxial wave equation (free space propagation) using FT. The wave equation for one transverse field component is given by

$$(2i\partial_\xi + \partial^2_{\vec{\rho}})\,\psi_0(\vec{\rho},\xi) = 0 \qquad (2.5.1)$$

Let us take the Fourier transform of this (linear) equation with respect to transverse coordinates

$$(2i\partial_\xi - q^2)\,\bar{\psi}_0(\vec{q},\xi) = 0 \qquad (2.5.2)$$

where

$$\bar{\psi}(\vec{q},\xi) = \int \frac{d^2\rho}{(2\pi)^2}\,\psi(\vec{\rho},\xi)\,e^{-i\vec{\rho}\cdot\vec{q}} \qquad (2.5.3)$$

is the direct FT, and

$$\Psi(\vec{\rho},\xi) = \int d^2q\ \bar{\Psi}(\vec{q},\xi)\, e^{i\vec{\rho}\cdot\vec{q}} \tag{2.5.4}$$

is the inverse FT. We can now solve Eq. (2) for $\bar{\Psi}_0$,

$$\bar{\Psi}_0(\vec{q},\xi) = e^{-\frac{i}{2}q^2\xi}\,\bar{\Psi}_0(\vec{q},0) \quad , \tag{2.5.5}$$

where $\bar{\Psi}_0(\vec{q},0)$ is the FT of the initial field $\Psi_0(\vec{\rho},0)$ of Eq. (1). Symbolically

$$\Psi_0(\vec{\rho},\xi) = (FT)^{-1} e^{-\frac{i}{2}q^2\xi} (FT)\, \Psi_0(\vec{\rho},0) \quad , \tag{2.5.6}$$

and if, for example, the initial field is Gaussian-like

$$\Psi_0(\vec{\rho},0) = e^{-\rho^2} \quad , \tag{2.5.7}$$

then

$$\bar{\Psi}_0(\vec{q},0) = \frac{1}{4\pi}\, e^{-\frac{q^2}{4}} \quad , \tag{2.5.8}$$

so that

$$\Psi_0(\vec{\rho},\xi) = (FT)^{-1}\, \frac{1}{4\pi}\, e^{-\frac{q^2}{4}(1+2i\xi)}$$

$$= \frac{1}{1+2i\xi}\, e^{-\frac{\rho^2}{1+2i\xi}} \quad . \tag{2.5.9}$$

In order to recognize the l=0 solution from Sec. 2, write

$$1+2i\xi = \sqrt{1+(2\xi)^2}\; e^{i\tan^{-1}2\xi} \quad , \tag{2.5.9}$$

and

$$\psi_0(\vec{\rho},\xi)=\frac{1}{\sqrt{1+(2\xi)^2}}\, e^{-\frac{\rho^2}{1+(2\xi)^2}+i\left[\frac{\rho^2}{1+(2\xi)^2}2\xi-\tan^{-1}2\xi\right]} \quad , \tag{2.5.11}$$

which the same as Eqs. (2.26) and (2.27).

Proceeding along the same lines, let us consider the full nonlinear paraxial equation (active medium propagation)[17]:

$$(2i\partial_\xi+\partial^2_{\vec{\rho}})\psi = -gLm_0(|\psi|^2)\,\psi(\vec{\rho},\xi) \quad , \tag{2.5.12}$$

and write the solution symbolically as

$$\psi(\vec{\rho},\xi)=T_\xi\, e^{\frac{i}{2}\left[\xi\partial^2_{\vec{\rho}}+gL\int_0^\xi m_0 d\xi'\right]}\psi(\vec{\rho},0)\equiv U\psi(\vec{\rho},0) \quad , \tag{2.5.13}$$

where T_ξ is the ξ-ordering operator, and U satisfies

$$(2i\partial_\xi+\partial^2_{\vec{\rho}})U=-gLm_0U \quad . \tag{2.5.14}$$

Let us call

$$A\equiv\frac{i}{2}\partial^2_{\vec{\rho}} \quad , \tag{2.5.15}$$

$$B(\vec{\rho},\xi) \equiv \frac{i\mathcal{G}\mathcal{L}m_o}{2} \quad . \tag{2.5.16}$$

If we look for a solution in the form

$$U = e^{\frac{1}{2}A\xi}\,\Gamma\, e^{\frac{1}{2}A\xi} \quad , \tag{2.5.17}$$

then

$$\frac{d\Gamma}{d\xi} = C(\vec{\rho},\xi)\,\Gamma \tag{2.5.18}$$

and

$$C(\vec{\rho},\xi) = e^{-\frac{1}{2}A\xi}\,B\,e^{\frac{1}{2}A\xi} + \frac{1}{2}A - \frac{1}{2}\Gamma A \Gamma^{-1} \quad . \tag{2.5.19}$$

Iterative solution to this pair of equations leads to

$$C = B + \frac{1}{2}\left[A, \int_0^{\xi} B\,d\xi' - \xi B\right] + \cdots \quad . \tag{2.5.20}$$

If B is a constant, the second term vanishes and the error resides in the double commutators. However, this is not the case, so assume that $B(\vec{\rho},\xi)$ is expandable in ξ :

$$B = B_o(\vec{\rho}) + B_1(\vec{\rho})\,\xi + \cdots \quad . \tag{2.5.21}$$

Then to first order

$$C = B - [A, B_1]\frac{\xi^2}{4} \quad . \tag{2.5.22}$$

Thus, the corrections in C are of order ξ^2 and for Γ are of order ξ^3. For this particular choice of U the second order errors cancel, which is why this particular U has been chosen. If we neglect the third order corrections, Eq. (13) simplifies to

$$\psi(\vec{\rho},\xi)=e^{\frac{i\xi}{4}\partial_{\vec{\rho}}^2}e^{\frac{igL}{2}\int_0^{\xi}m_0 d\xi'}e^{\frac{i\xi}{4}\partial_{\vec{\rho}}^2}\psi(\vec{\rho},0) \quad . \tag{2.5.23}$$

More generally, if we propagate for $\Delta\xi$, from ξ_n to ξ_{n+1}, the appropriate solution is

$$\psi(\rho,\xi_{n+1})=e^{\frac{i\Delta\xi}{4}\partial_{\vec{\rho}}^2}e^{\frac{igL}{2}\int_{\xi_n}^{\xi_{n+1}}m_0 d\xi'}e^{\frac{i\Delta\xi}{4}\partial_{\vec{\rho}}^2}\psi(\vec{\rho},\xi_n) \quad . \tag{2.5.24}$$

This is the main working formula. It may be visualized as if a sequence of three operators is acting on the initial field $\psi(\vec{\rho},\xi_n)$ producing propagation of the field for $\Delta\xi$ in the active medium. The first operator is simply the free space propagation for $\Delta\xi/2$, as it might be seen by inspecting Eq. (1). The second one is the gain correction over the whole interval $\Delta\xi$, and the third one is again the free space propagation for the remaining $\Delta\xi/2$. The whole active medium region can be divided into some number of $\Delta\xi$ steps, each of those steps having a gain-sheet in the middle. Then the active medium propagation is broken into a number of free space propagations for $\Delta\xi/2$ and gain corrections at the gain

stations[18]. The accuracy of the method presumably is better when there are more gain stations.

Let us look into the gain correction in more detail. Consider some n-th step, so that the appropriate gain factor is

$$G \equiv \exp \frac{igl}{2} \int_{\xi_n}^{\xi_{n+1}} m_0 d\xi' \quad . \tag{2.5.25}$$

Since $\Delta\xi$ is assumed small, let us apply a simple trapezoidal formula to the integral in the exponent

$$\int_{\xi_n}^{\xi_{n+1}} m_0(\psi) d\xi' \approx m_0(\psi(\xi_n))\Delta\xi + \frac{1}{2}\left[m_0(\psi(\xi_{n+1})) - m_0(\psi(\xi_n))\right]\Delta\xi =$$

$$= \frac{1}{2}\left[m_0(\psi(\xi_n)) + m_0(\psi(\xi_{n+1}))\right]\Delta\xi \quad . \tag{2.5.26}$$

However, $\psi(\xi_{n+1})$ is not known as yet; only $\psi(\xi_n)$ is known. Therefore, at the gain sheet some iteration procedure must be devised. The simplest one is the one point iteration: assume in the first approximation that $\psi^{(1)}(\xi_{n+1}) = \psi(\xi_n)$, then correct for the gain, propagate freely for $\Delta\xi/2$ and evaluate a new, presumably more correct value for $\psi(\xi_{n+1}) = \psi^{(2)}$.

Then repeat the whole procedure as many times as it is felt necessary. In the actual computations, however, we found more convenient not to do the iteration but to approximate the integral simply by $m_0(\xi_n)\Delta\xi$ and take a smaller $\Delta\xi$ step. The results were very close and the appeal

of the simplified procedure is in the saving of computer core space .

The next part of the resonator problem is mirror reflection. It consists of two parts, one being the multiplication of the propagating field by the reflection coefficient and the other multiplication of the field by the cutoff function. The order of these two operations doesn't matter. The reflection coefficient takes into account the change in the phase and intensity of the field by reflection and the cutoff function takes into account the finite size of mirrors and introduces so-called guard bands . A suitable choice of the cutoff function also suppresses the Gibbs overshoot[19]. The notion and reason for the existance of guard bands comes from the following argument.

The fast Fourier transform (FFT) procedure actually solves the problem of an infinite array of identical cells located side by side with spacing equal to the full width of the transverse grid. If close to the edge of the cell the field is not zero, by doing FFT the diffracted field will spill into adjacent cells (i.e. into itself), making the result inaccurate and unreliable. Thus each cell must be separated from the adjoining cells by an adequate guard band where the field is essentially zero.

Another problem connected with the FFT method is the Gibbs overshoot. When a wave pattern with a sharply cutoff

edge is sampled at discrete intervals, FF transformed and then reconstructed by inverse transformation,

the sampled points are reproduced exactly. However, propagation of the original function by FFT methods will usually produce a significant overshoot on either side of the step discontinuity. Part of this overshoot is real, since the diffracted wave pattern includes edge-wave contributions also. However, part of the overshoot also represents a Gibbs phenomenon associated with the truncation of the FT at the maximum frequency of the FFT. Straightforward application of the FFT to the FT is equivalent to multiplying the FT of the input wave front by an ideal lowpass filter or a data window. To reduce the Gibbs-phenomenon overshoot it is desirable to multiply the FT by a data window having more gradually tapered high-frequency edges, or equivalently to soften the sharp edge of the original wave-front.

Let us now derive the appropriate reflection coefficient. Using the standard boundary condition for the tangential components of E field, it is

$$E_r(\rho_s, z_s) + E_i(\rho_s, z_s) = E_t(\rho_s, z_s) \tag{2.5.27}$$

at the mirror surface. E_r is the reflected field, E_i is the incident field and E_t is the transmitted field. By definition,

$$E_r(\rho, z_m) = R_E E_i(\rho, z_m) \quad , \tag{2.5.28}$$

$$E_t(\rho, z_m) = T_E E_i(\rho, z_m) \quad , \tag{2.5.29}$$

where z_m is the mirror site, and R_E and T_E are the reflection and transition coefficients. Note that T_E is nearly real, E_i and E_t being almost in phase. R_E is clearly complex. As assumed in the section 2 on theory,

$$E_i^{\pm} = e^{\pm ikz} \psi_i^{\pm}(\rho, z) \tag{2.5.30}$$

where $\psi_i^{\pm}(\rho, z)$ is the envelope function, slowly varying in the propagation ($\pm z$) direction. The mirrors are assumed spherical, and let us also assume that the left-hand side, small mirror is placed at the coordinate origin $z=0$. Then the right-hand side, large mirror is placed at $z=d$, where d is the length of the laser cavity. Let us first consider reflection from the large mirror. Then the equation of the mirror surface is

$$R^2 = \rho_s^2 + [z_s - (d-R)]^2 \quad , \tag{2.5.31}$$

where, according to our resonator model we chose R positive (a concave mirror). In order to obtain the reflection coefficient we must interpolate Eq. (27) to the $z_m = d$ plane where the coefficients are defined. Therefore

$$z_m = z_s + \delta z \quad , \quad \varrho_m = \varrho_s + \delta\varrho \tag{2.5.32}$$

where $\delta z > 0$ for concave, and $\delta z < 0$ for convex mirrors. Furthermore , $\delta\varrho$ is an order smaller than δz , since we are in the paraxial approximation, and hence it will be neglected. Rewriting Eq. (27)

$$E_r(\varrho_m, z_m - \delta z) + E_i(\varrho_m, z_m - \delta_2) = E_t(\varrho_m, z_m - \delta z) \quad , \tag{2.5.32}$$

or

$$e^{-\delta z \partial_{z_m}} E_r(\varrho_m, z_m) + e^{-\delta z \partial_{z_m}} E_i(\varrho_m, z_m) = e^{-\delta z \partial_{z_m}} E_t(\varrho_m, z_m) \quad . \tag{2.5.33}$$

Since the ψ-functions are slowly varying in z,

$$\partial_{z_m} e^{\pm ikz_m} \psi^{\pm}(\varrho_m, z_m) \approx \pm ik\, e^{\pm ikz_m} \psi^{\pm}(\varrho_m, z_m) \tag{2.5.34}$$

and also

$$f(\partial_{z_m})\, E^{\pm}(\varrho_m, z_m) \approx f(\pm ik)\, E^{\pm}(\varrho_m, z_m) \quad , \tag{2.5.35}$$

for the forward (+) and the backward (-) propagating fields. Therfore, at the large mirror

$$e^{ik\delta z} E_r(\varrho_m, z_m) + e^{-ik\delta z} E_i(\varrho_m, z_m) = e^{-ik\delta z} E_t(\varrho_m, z_m) \quad , \tag{2.5.36}$$

or

$$E_r(\rho_m, z_m) = -e^{-2ik\delta z}(1-T_E)E_i(\rho_m, z_m) \quad . \qquad (2.5.37)$$

Using Eqs. (31) and (32) it is possible to represent δz via R and ρ_m, namely :

$$R^2 = \rho_m^2 + (\delta z - R)^2 \quad , \qquad (2.5.38)$$

and

$$\delta z = \frac{\rho_m^2}{2R} \qquad (2.5.39)$$

Finally

$$R_E = -|R_E| e^{-\frac{ik\rho_m^2}{R}} \quad , \qquad (2.5.40)$$

where

$$|R_E| = 1 - T_E \quad . \qquad (2.5.41)$$

Since $0 \leqslant T_E \leqslant 1$ it follows $0 \leqslant |R_E| \leqslant 1$ also. Note that the reflection coefficient for ψ-fields, R_ψ is given by

$$R_\psi = R_E e^{2ikz_m} = R_E e^{2ikd} \quad . \qquad (2.5.42)$$

However, the round-trip condition $2kd = 2\pi$ is satisfied for confocal resonators, so that in the propagation of ψ-fields we can use the same reflection coefficient

$$\mathcal{R}_\psi = \mathcal{R}_E \equiv \mathcal{R} \tag{2.5.43}$$

Analogously, for the small mirror

$$z_m = 0 \;, \quad R^2 = (z_s - R)^2 + \varrho_s^2 \;, \tag{2.5.44}$$

where we assumed R negative (a convex mirror). Since now

$$\delta z = -z_s > 0 \;, \quad \varrho_m \approx \varrho_s \tag{2.5.45}$$

it is

$$\delta z = -\frac{\varrho_s^2}{2R} \;. \tag{2.5.46}$$

Also

$$E_r(\varrho_m, z_m) = -e^{2ik\delta z}(1 - T_E)\, E_i(\varrho_m, z_m) \tag{2.5.47}$$

since the sense of the propagation of fields is now opposite. Therefore the same result (except for the sign of R) follows for R_E.

As the wave gets reflected, a new problem arises. We

plan to move the wave back and forth until it eventually stabilizes. We have in mind here the wave which satisfies Eq. (12) , which we will call forward or main wave. However, the wave which moves backward is different, it satisfies a different equation, in the same way as the bare resonator allows two different solutions,

$$\psi_o^f = \frac{1}{1+2i\xi} e^{-\frac{\rho^2}{1+2i\xi}} \quad , \tag{2.5.9}$$

and

$$\psi_o^b = \frac{1}{1-2i\xi} e^{-\frac{\rho^2}{1-2i\xi}} \quad , \tag{2.5.9'}$$

which satisfy two different equations:

$$(2i\partial_\xi + \partial_{\vec{\rho}}^2)\psi_o^f = 0 \quad , \tag{2.5.1}$$

and

$$(-2i\partial_\xi + \partial_{\vec{\rho}}^2)\psi_o^b = 0 \quad . \tag{2.5.1'}$$

(Admittedly, this situation is somewhat simpler because the equations and therefore the solutions are simply complex conjugate of each other). Moreover, forward and backward waves do interfere and the stable, settled solution is a sort of standing wave. Actually, the gain function involves both forward and backward intensities, so that the full

problem[20] requires a solution to the system of coupled equations

$$(2i\partial_\xi+\partial_\rho^2)\psi^f = -g\ell m_o^f(I_f,I_b)\,\psi^f \quad , \tag{2.5.12}$$

$$(-2i\partial_\xi+\partial_\rho^2)\psi^b = -g\ell m_o^b(I_f,I_b)\,\psi^b \quad , \tag{2.5.12'}$$

where

$$m_o^{f/b} = \frac{\Omega - i}{\sqrt{a^2-b^2}}\left(1 - \frac{a-\sqrt{a^2-b^2}}{2I_{f/b}}\right) \quad , \tag{2.5.48}$$

$$a = 1+\Omega^2+I_f+I_b \quad , \qquad b = 2\sqrt{I_f I_b} \quad , \tag{2.5.49}$$

and I_f and I_b are the forward and backward intensities. This general problem is a bit too formidable for our resources. So we devised an equivalent procedure which is slightly simpler but also converges to an equivalent steady-state solution. The philosophy is as follows.

Consider some point inside the resonator. Two waves are collapsing on that point, the forward coming from the left, and the backward coming from the right. Each wave satisfies its own equation, (12) or (12'), subject to some boundary condition at the LHS and RHS mirror. In order to integrate these equations or to propagate these waves to adjacent points we need know the fields at our point. As a matter of fact we need know both fields simultaneously at any point in

the resonator. Therefore, we must integrate Eqs. (12) and (12') simultaneously starting, say, from the LHS mirror. Naturally, the active medium plays different roles for forward and backward waves - it is a gain-medium for the forward and loss-medium for the backward wave. We start at the LHS mirror with some initial fields and as we arrive at the RHS mirror we are tempted to simply reflect the fields and let the thing evolve by itself. However, in such a process the backward wave would quickly diminish and we would be left with a one-field problem which, as we mentioned does not correspond to the actual situation. To the actual, steady-state situation both forward and backward waves in some sense "equally" contribute. So, instead of reflecting the backward wave, we will set it equal to the forward wave at that place. In a sense we are restating the boundary condition to one which is presumably more close to the correct, steady-state condition. Now the waves propagate again, and the backward wave repeats values of the forward wave of the previous run. This is exactly what we wanted it to do, because this is what happens in the steady state (when any run is the same as the previous run) and because it will help the whole procedure converge to the steady state. It is our experience that this procedure of simultaneous integration and continuous improvement of the boundary condition converges efficiently and rapidly to the steady solution. Moreover, it seems to us that even the physics of the problem is imitated closely by the procedure.

Furthermore, the computer core requirements for this procedure are by far less then for the methods of similar accuracy advocated by others[19].

As mentioned earlier, a way to reduce the Gibbs overshoot is to soften the sharp edge of the wave front. We did that whenever we found it necessary. However, at the most sensitive place - small mirror reflection - we decided to eliminate the overshoot altogether. Instead of using FFT which causes the overshoot, we used an alternative procedure, devised by Lax, Agrawal and Louisell[21], named the continuous Fourier transform (CFT).

An ordinary FFT program is doing a very simple thing: you supply an array of (complex) numbers a_j , $j = 1,\ldots,N$ as an input, and as the output you obtain an array of numbers $\bar{a}_l = \sum_{j=1}^{N} e^{\frac{2\pi i}{N} jl} a_j$, $l = 1,\ldots,N$. It is then relatively easy to incorporate this fact into the approximation of the one-dimensional or two-dimensional FT integrals. In CFT, however, the Fourier integrals of the type

$$\bar{f}(k) = \int_{-\infty}^{\infty} f(z)\, e^{ikz}\, dz \qquad (2.5.50)$$

are dealt with differently. Firstly, a spline fit to the input function f(z) is obtained,

$$f(z) = \sum_{j=1}^{N_m - l} a_j B_{j,l}(z) \qquad (2.5.51)$$

where 1-th order B spline (basic spline) $B_{j,l}(z)$ is a polynomial of degree 1-1 in each (z_j, z_{j+1}) interval and zero outside. N_m is the number of mesh points $z_j = jh$, $j = 1,\ldots,N_m$ which may have nothing to do with the grid of sample points at which the function $f(z)$ is supplied. B splines form a numerically stable set of basis functions and can be obtained from recurrence relations or from the table of splines. On substituting Eq. (51) into Eq. (50) we obtain

$$\bar{f}(k) = \sum_j a_j \int_{z_j}^{z_{j+1}} B_{j,l}(z) e^{ikz} dz \quad . \tag{2.5.52}$$

Integral in Eq. (52) can be done analytically, by doing a partial integration 1-l times and taking into account a recurrence relation between the B splines. The final answer is of the form

$$\bar{f}(k) = h \sum_{j=1}^{N_m - l} a_j e^{ikjh} \phi_j(k) \tag{2.5.53}$$

where $\phi_j(k)$ acts as a window factor in the frequency domain and is given by

$$\phi_j(k) = \left(\frac{e^{ikh} - 1}{ikh} \right)^l \equiv \phi(k) \tag{2.5.54}$$

We should mention that, at both ends, because of multiple mesh points, the mesh is not uniform and proper end

corrections should be applied. We checked the procedure by propagating a square pulse in the bare resonator firstly using FFT and then CFT for identical sets of parameters. The CFT results were more accurate.

CHAPTER 2. High Power Laser

2.6 Programming Considerations and Results

When all parts of the problem were assembled, a fairly long and complicated code resulted. It included main program and 8 subroutines, totalling well over one thousand lines. All the subroutines perform some function which is repeatedly met in the main code. So, there are three kinds of propagation routines - the free-space-propagation (FSP) routine, which propagates the field for some distance in the empty space, the active-medium-propagation (AMPR) routine, which propagates the field for some distance in the active medium and the free-space-CFT (FSC) routine which propagates the field with a sharp edge. Furthermore, there is the initialization routine, the reflection routine, the gain-factor routine, the output routine and the plotting routine. All the propagating routines use a FFT program and an auxiliary program called the symmetric-half-point-FFT (SHPFFT) which enables us to obtain a full FFT by using only one quarter of the available data points. Namely, because of the $x \rightarrow -x$, $x \rightarrow y$, and $y \rightarrow -y$ symmetry of the problem

only one eighth of the grid points carries independent information. We succeeded in using only twice this minimal number, effectively cutting the core space requirement by a factor of 4. And it was not so easy as it sounds!

The program code is as follows: Starting from the LHS mirror at which initial fields are supplied, the fields are propagated for the distance d_1 using FSC for the forward and ordinary FSP for the backward wave. At d_1 the active medium begins and AMPR is being used for both waves. In any $\Delta\xi$ step each of the waves is propagated freely for $\Delta\xi/2$ then corrected for the gain and finally propagated freely for the remaining $\Delta\xi/2$. In order to save CPU time two adjacent $\Delta\xi/2$ propagations are lumped together. Since the gain medium ends somewhere in the transverse directions (most probably it does not extend beyond the large mirror) we cut it near the grid edge. This insures that the field at the grid edge does not grow uncontrolably and unrealisticly. From $d-d_2$, where the active medium ends to the RHS mirror at d we again propagate freely using FSP. At the mirror site the forward wave is reflected and the backward wave is reset to the forward wave at that place. And then the whole procedure repeats itself in the reverse order, completing a full cycle. In the second leg of the cycle, however no FSC is being used. And then the whole procedure repeats itself for certain preset number of cycles or until a stable solution is reached. We note that the code can be applied to one

transverse dimension as well - then we have a two dimensional amplifier.

As for the results we present some in Figs. 2, 3, and 4. The stable intensity distribution at the output mirror is presented for various sets of parameters. One striking fact is apparent immediately: intensity distribution for the three dimensional case is proportional to the product of the two dimensional distributions,

$$I^{(3)}(x,y,z) = I^{(2)}(x,z)\, I^{(2)}(y,z) \qquad (1)$$

We do not have an adequate explanation for this numerical result. However, it suggests a possibility of obtaining three dimensional distributions from the two dimensional ones, in which case all storage problems are superficial. We also note apparent qualitative simmilarity between the experimental intensity distribution along one of the main axes, given in Fig. 5 and the two dimensional distribution with many points given in Fig. 4.

Figure captions

Figure 1.(a) Typical stable and unstable resonators
(b) Positive- and negative-branch confocal or telescopic unstable resonators.
(c) The round-trip magnification M characterizes the geometrical properties of an unstable resonator.

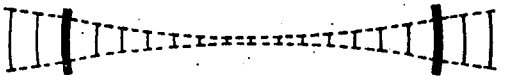
Figure 2. A plot of 2-dim. intensity distribution I(X,Z=0) in units of the saturation intensity, where $X=\sqrt{2\pi N_F}\frac{x}{a}$, $Z=\frac{z}{d}$, a is the small mirror half-width, d is the mirror separation and N_F is the Fresnel number. A confocal configuration was used with a round-trip magnification of 2.5, and the Fresnel number obeyed $\sqrt{2\pi N_F} = 7.5$. The small one way signal gain is $gd = 10$ and the active region length is divided into six gain sheets. Roundtrips 8,9 and 10 are not distinguishable on the plot.

Figure 3. A plot of 3-dim. intensity distribution I(X,Y,Z=0) for six values of Y. Parameters are the same as in Fig. 2. A perfect symmetry and striking resemblence with the two dimensional case is obvious.

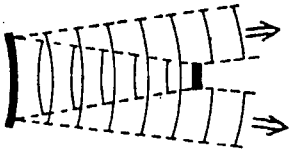
Figure 4. Same as in Fig. 2, only values for the parameters are different: $\sqrt{2\pi N_F} = 12.8$, gd=17.25, number of transverse points is 1025.

Figure 5. An experimental, actual laser intensity output at the small mirror. A number of qualitative agreements with the predicted outputs in Figs. 3 and 4 is evident: (a) Within the

small mirror the intensity distribution is fairly uniform, (b) Right after the small mirror edge the intensity drops rapidly and then peaks rapidly, (c) Right before the large mirror edge the intensity peaks rapidly and then drops rapidly, (d) Outside corners of the small mirror show low intensity, and inside corners of the large mirror show high intensity. Admittedly, the agreement is not very pronaunced, because neither of the theoretical cavities corresponds exactly to the experimental one. Also the small mirror holders (which are not provided for in the theoretical case) are evidently producing a large effect.

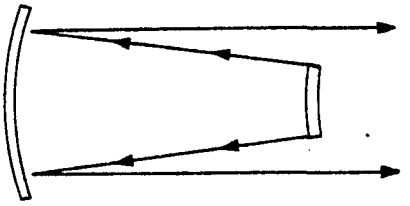


STABLE

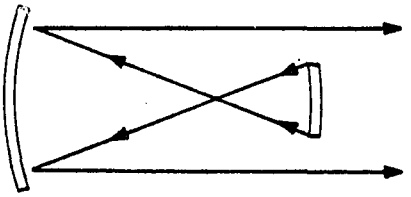


UNSTABLE

a)

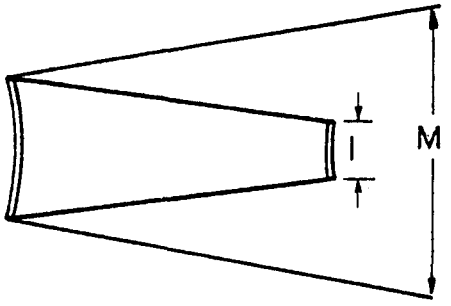


POSITIVE BRANCH



NEGATIVE BRANCH

b)



$$\frac{\text{LOSS}}{\text{BOUNCE}} = 1 - \left(\frac{l}{M}\right)^2$$

c)

Figure 1.

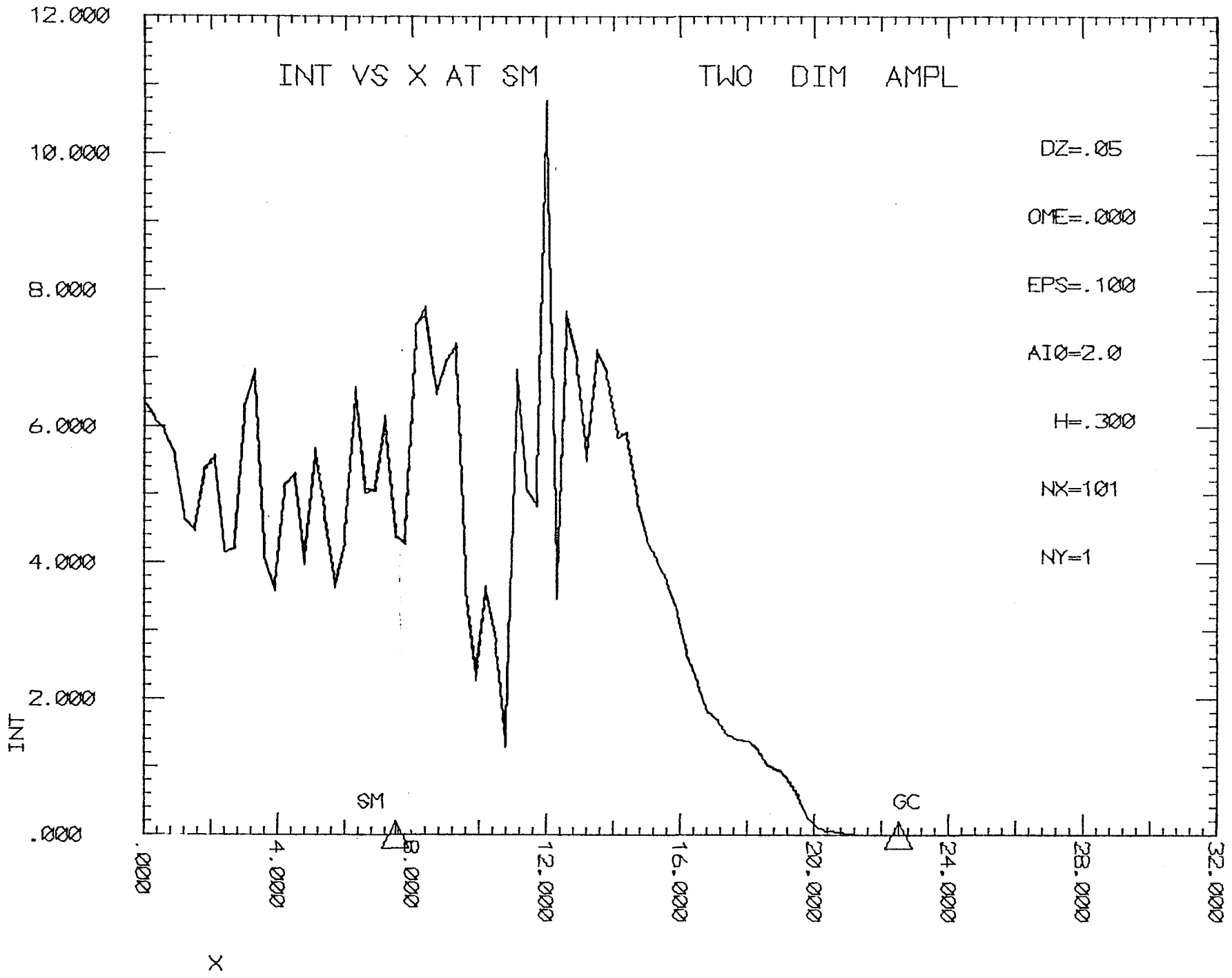

INT VS X AT SM
TWO DIM AMPL
DZ=.05
OME=.000
EPS=.100
AI0=2.0
H=.300
NX=101
NY=1
INT
X
SM
GC
12.000
10.000
8.000
6.000
4.000
2.000
.000
.000
4.000
8.000
12.000
16.000
20.000
24.000
28.000
32.000

Figure 2

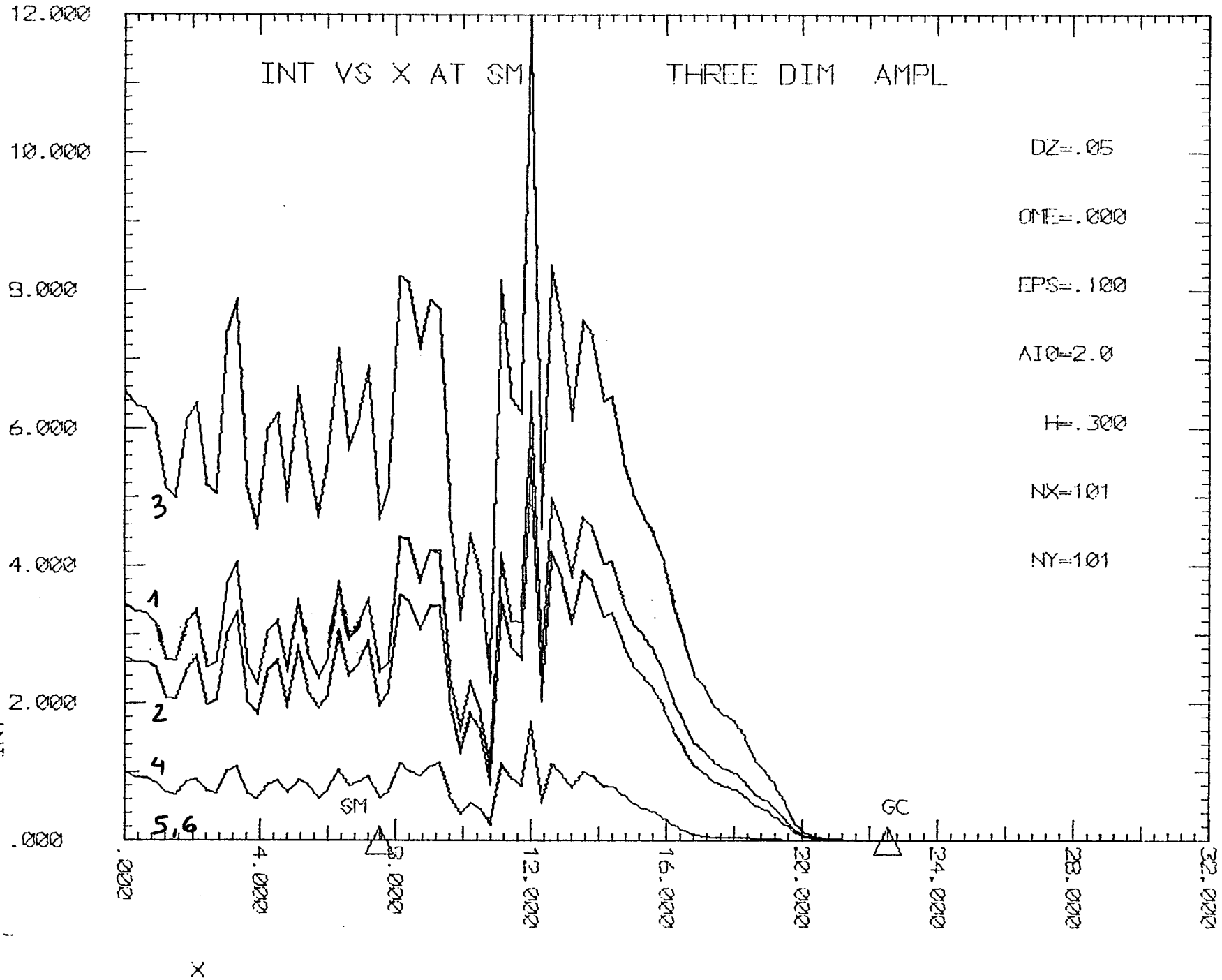

INT VS X AT SM
THREE DIM AMPL
DZ=.05
DME=.000
EPS=.100
AI0=2.0
H=.300
NX=101
NY=101
12.000
10.000
8.000
6.000
4.000
2.000
.000
INT
X
SM
GC
3
1
2
4
5,6

Figure 3

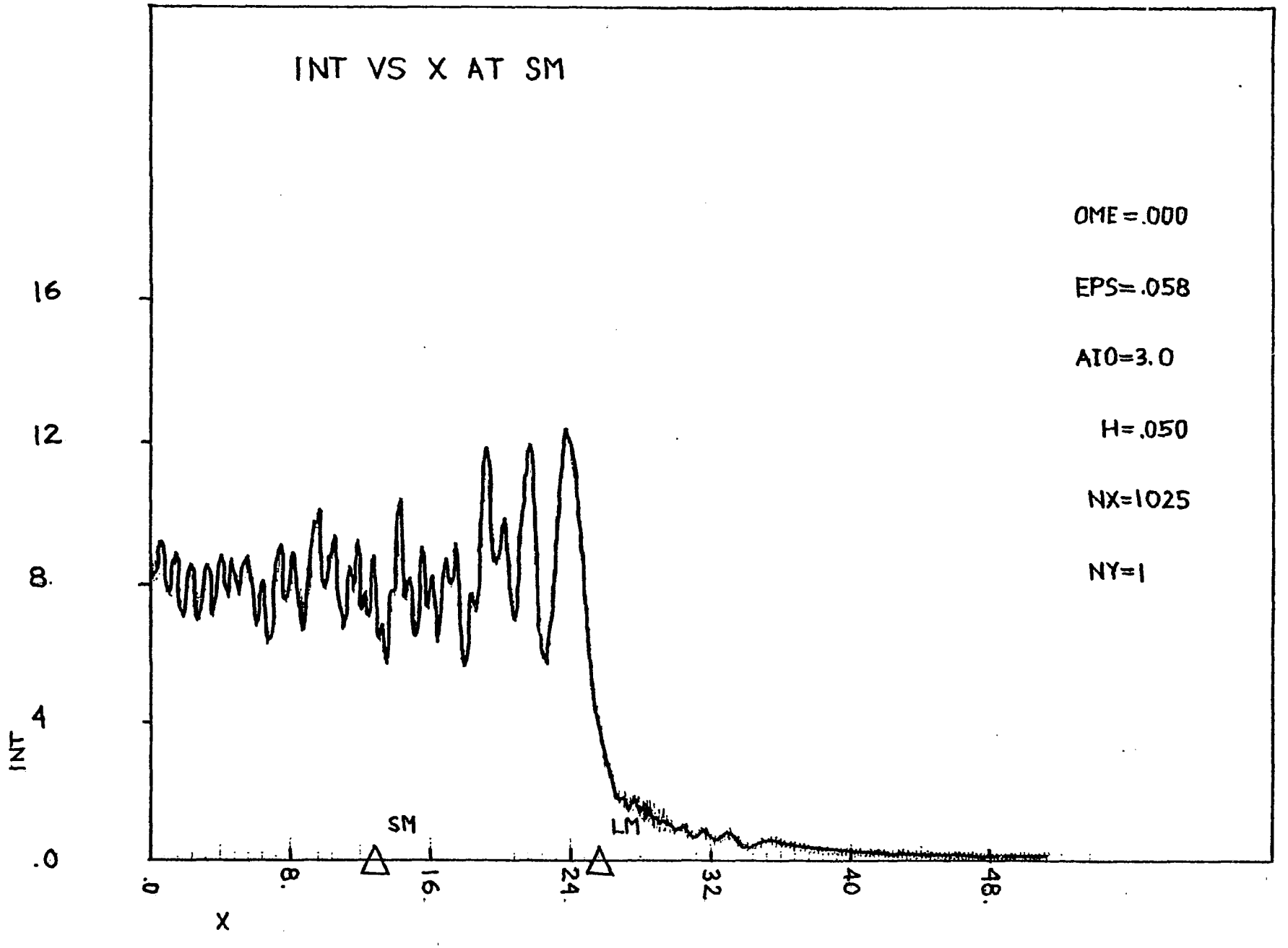

INT VS X AT SM
OME=.000
EPS=.058
AIO=3.0
H=.050
NX=1025
NY=1
INT
16
12
8.
4
.0
SM
LM
0
8.
16.
24.
32
40
48.
X

Figure 4.

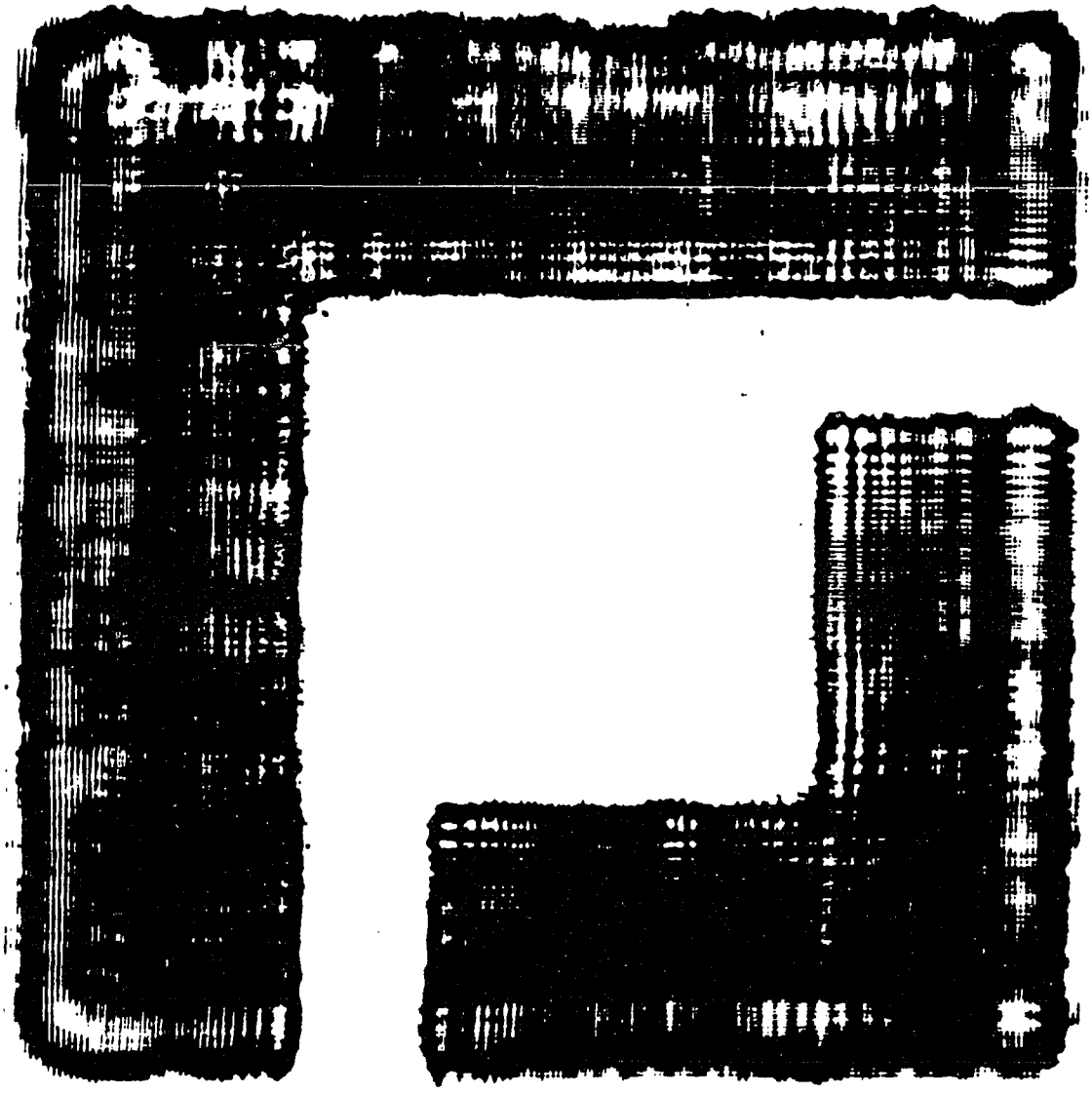


Figure 5.

References

1. R. H. Dicke, US Patent No. 2,851,652 (1958)
2. A. L. Schawlow and C. H. Townes, Phys.Rev. 112, 1940 (1958)
3. A. E. Siegman, Proc. IEEE 53, 277 (1965)
4. A. E. Siegman and R. W. Arrathon, IEEE J. Quantum Electron. QE-3, 156 (1967)
5. A. Okaya, Proc. IEEE 51, 1033 (1963)
6. L. F. Mollenauer, G. F. Imbush, H. W. Moos, A. L. Schawlow, and A. D. May in "Optical Masers", J. Fox ed., vol. 13 (Polytechnic Press, Brooklyn 1964)
7. J. T. LaTourrette, S. Jacobs, and P. Rabinowitz, Appl. Opt. 3, 981 (1964)
8. A. G. Fox and T. Li in "Quantum Electronics III", P. Grivet and N. Bloembergen eds., (Columbia U.P. New York 1963)
9. Yu. A. Anan'ev, G. Vinokurov, L. Kovalchuk, N. Sventsitskaya and V. E. Sherstobitov, Sov. Phys.-JETP 31, 420 (1970)
10. W. F. Krupke and W. R. Sooy, IEEE J. Quantum Electron. QE-5 575 (1969)
11. Yu. A. Anan'ev, Sov. J. Quantum Electron. 1, 565 (1972)
12. A. E. Siegman, Appl. Optics 13, 353 (1974)
13. M. Lax, G. P. Agrawal, M. Belic, and W. H. Louisell, Submitted for XI International Quantum Electronics Conference in Boston, June 1980
14. M. Lax, W. H. Louisell and W. B. McKnight, Phys. Rev. A11, 1365 (1975)
15. D. B. Rensch and A. N. Chester, Appl. Opt. 12, 997 (1973)
16. See, for example in A. Yariv's "Quantum Electronics"

(John Wiley & Sons, New York 1975), p. 135

17. M. Lax , J. H. Batteh and G. P. Agrawal, "Channeling of Intense Electromagnetic Beams", submitted for publ.

18. A. E. Siegman and E. A. Sziklas, Appl. Opt. 13, 2775 (1974)

19. E. A. Sziklas and A. E. Siegman, Appl. Opt. 14, 1874 (1975)

20. G. P. Agrawal and M. Lax, JOSAA, (Dec. 1979)

21. M. Lax, G. P. Agrawal and W. H. Louisell, Opt. Lett. 4, 303 (1979)

CHAPTER 3. Solar Heat Pump

3.1 Introduction

This section addresses the question: If we shine light on a physical system, is it capable of acting as a refrigerator. If the answer is in the affirmative, some rather attractive and far reaching consequences for solar energy utilization may be forseen. Some insight into this question may be gained by asking what limitations, if any, are imposed by thermodynamics. The device we will ultimately propose consists of an array of semiconductor and metal elements. Part of the device acts as a solar battery and part of the device acts as a Peltier junction. The electric current produced by the "solar battery" is delivered to the Peltier junction and the net result is a light-driven refrigerator.

Not much, even conceptually has been done on this subject. In some papers on optical refrigeration[1,2] a different mechanism is being proposed, a tunneling-assisted photon emission[3], in which an electron makes transition

across a forward-biased p-n junction by a two-step process: tunneling and photon-emission. It has been found that in a number of diodes the energy of the luminescent photons exceeds the voltage drop across the junction, the energy difference coming from the lattice. In the result, light is being emitted, the junction region gets cooled and the battery does the work. In a sense we envision a completely reverse process: sunlight is being absorbed, the junction region gets heated and the sun does the work.

Consider the diagram illustrated in Fig. 1. The upper and lower rectangles depict two thermal reservoirs at temperatures T_1 and T_2 respectively, where $T_1 > T_2$. The circle labeled "A" is symbolic of the photovoltaic junction, whereas the circle labeled "B" denotes the Peltier junction. Let E be the optical energy delivered to the cell S in some time interval. A fraction f of this energy is converted to electrical energy and is delivered to the Peltier junction. Let Q_2 denote the heat extracted from the system and Q_1 the heat delivered to the environment. Then

$$Q_1 = fE + Q_2 \quad . \tag{1}$$

In order for the system to be cooled, we must have

$$(1-f)E < Q_2 \quad . \tag{2}$$

The coefficient of performance of the Peltier cell acting as a refrigerator is

$$\omega = \frac{Q_2}{fE} \quad . \tag{3}$$

An upper limit on this coefficient is provided by the Carnot expression for a reversible refrigerator:

$$\omega < \frac{T_2}{T_1 - T_2} \quad . \tag{4}$$

Combining the above relations leads to the expression

$$\frac{1-f}{f} < \frac{T_2}{T_1 - T_2} \tag{5}$$

or

$$f > 1 - \frac{T_2}{T_1} \quad . \tag{6}$$

Typically we might expect f to be as large as 0.1, so for $T_1 = 300°K$ we may have $T_2 = 270°K$.

This estimate appears quite encouraging from a practical point of view. If the device is operated in the reverse direction it can act as a heat pump to drive a thermal current from hot to cold - a current above and beyond the normal conduction current. This would be desirable in situations where solar heating is to be envisaged.

This chapter is organized in the following way. Section 2 concerns itself with the kinetic theory basis of the physical equations needed to describe the conversion of light energy into the driving forces for thermal currents. In section 3 we develop a set of semi-phenomenological equations to describe the effect. In section 4 these equations are analyzed under some simplifying assumptions. In section 5 the thermal currents are studied in some detail and in section 6 a device is proposed which would act as a light-driven refrigerator.

CHAPTER 3. Solar Heat Pump

3.2 Kinetic Theory

We are interested in describing a physical system which is subjected to external radiation and which exhibits thermal and electrical current flows. In order to obtain a complete theory it is perhaps wise to choose a fundamental starting point and systematically develop the formalism from that point. This is the goal of the present section.

Our primary interest will be in semiconductor systems with two types of carriers - electrons and holes. Let f_n be the phase space distribution function for the electrons and f_p the corresponding function for the holes. The Boltzmann equation for f_n is

$$\partial_t f_n + \vec{v}\cdot\partial_{\vec{r}} f_n + \frac{e}{m_n}\partial_{\vec{r}}\psi\cdot\partial_{\vec{v}} f_n = -\frac{f_n - f_n^0}{\tau_n} + g_n \quad . \tag{3.2.1}$$

Here m_n denotes the electron effective mass, which will be assumed scalar. The electrostatic potential is denoted by ψ.

The distribution function relaxes to the quasi-Fermi distribution f_n° in a relaxation time τ_n . The generation rate of electrons due to radiation and recombination is denoted by g_n . The quasi-Fermi distribution is given by

$$f_n^\circ = \left[1 + \exp \beta(E - e\Psi - \phi_n)\right]^{-1} , \tag{3.2.2}$$

where E is the electronic energy, β is $1/kT$, and ϕ_n is electrons electro-chemical potential referred to the bottom of the conduction band. An equation similar to Eq. (1) will apply to the holes but with e replaced by -e, n replaced by p, and ϕ_p referred to the top of the valence band. In the thermal equilibrium situation we have

$$\phi_n + \phi_p = E_c - E_v = E_g , \tag{3.2.3}$$

where E_g is the gap energy. We shall be interested in the non-equilibrium case, however.

The electron density is given by the integral over velocity space of f_n :

$$n = \int f_n \, d^3v , \tag{3.2.4}$$

and the electron current density is

$$\vec{J}_n = \int \vec{v} f_n \, d^3v . \tag{3.2.5}$$

The zeroth moment of the Boltzmann equation is

$$\partial_t n + \partial_{\vec{r}} \cdot \vec{J}_n = G \quad , \tag{3.2.6}$$

where G is the net generation rate per unit volume:

$$G = \int g_n \, d^3v \quad . \tag{3.2.7}$$

The effect of lattice collisions is simply to redistribute population within a band, so

$$\int \frac{f_n - f_n^\circ}{\tau_n} \, d^3v = 0 \quad . \tag{3.2.8}$$

The first moment of the Boltzmann equation with respect to energy is

$$\int \left[E - e\psi\right] \left[\partial_t f_n + \vec{v} \cdot \partial_{\vec{r}} f_n + \frac{e}{m_n} \partial_{\vec{r}} \psi \cdot \partial_{\vec{v}} f_n + \frac{f_n - f_n^\circ}{\tau_n} - g_n\right] d^3v = 0 \quad . \tag{3.2.9}$$

Assume henceforth a steady state , so $\partial_t f_n = 0$. Let R_n denote the energy transfer rate to the lattice (per unit volume) due to collisions:

$$R_n = \int \left[E - e\psi\right] \frac{f_n - f_n^\circ}{\tau_n} \, d^3v \quad . \tag{3.2.10}$$

The energy generation rate is

$$GQ_n \equiv \int (E - e\psi) g_n \, d^3v \quad . \tag{3.2.11}$$

Hence Eq.(9) may be written as

$$\partial_{\vec{r}} \cdot \vec{J}_{U_n} = -R_n + GQ_n \quad , \tag{3.2.12}$$

where $\vec{J}_{U_n}$ denotes the energy current associated with the electrons. Now let us expand f_n around the quasi-Fermi distribution. Let $f_n = f_n^0 + F_n$, and define a linear operator

$$L_n \equiv \vec{v} \cdot \partial_{\vec{r}} + \frac{e}{m_n} \partial_{\vec{r}} \psi \cdot \partial_{\vec{v}} \quad . \tag{3.2.13}$$

The Boltzmann equation becomes

$$L_n f_n^0 + L_n F_n = -\frac{F_n}{\tau_n} + g_n \quad . \tag{3.2.14}$$

We will make the local approximation and assume that $|L_n F_n| << |F_n/\tau_n|$. This is equivalent to saying that the collision time is sufficiently short so that only local thermodynamic properties are of importance. Then

$$F_n = g_n \tau_n - \tau_n L_n f_n^0 \tag{3.2.15}$$

so

$$f_n = f_n^0 + g_n \tau_n - \tau_n L_n f_n^0 \tag{3.2.16}$$

Define

$$n_0 \equiv \int f_n^0 \, d^3v \quad , \tag{3.2.17}$$

$$\delta n \equiv \int g_n \tau_n \, d^3v \quad , \tag{3.2.18}$$

and note that

$$\int \tau_n L_n f_n^0 \, d^3v = 0 \, . \tag{3.2.19}$$

Then $n = n_0 + \delta n$, and

$$L_n f_n^0 = \frac{1}{\beta} \partial_E f_n^0 \left[(E - e\psi) \vec{v} \cdot \partial_{\vec{r}} \beta - \vec{v} \cdot \partial_{\vec{r}} (\beta \phi_n) \right] \, , \tag{3.2.20}$$

so

$$\vec{J}_n = - \int \frac{\tau_n \vec{v}}{\beta} \partial_E f_n^0 \left[(E - e\psi) \vec{v} \cdot \partial_{\vec{r}} \beta - \vec{v} \cdot \partial_{\vec{r}} (\beta \phi_n) \right] d^3v \, . \tag{3.2.21}$$

We will assume an effective isotropic semiconductor with parabolic bands, so that we may make the replacement

$$\langle v_i v_j \rangle = \frac{2E}{3m_n} \delta_{ij} \, , \quad i,j = 1,2,3 \tag{3.2.22}$$

Let us define a set of collision integrals

$$\tau_n^{(j)} = -\frac{\beta^{j-1}}{n_o}\int \tau_n E^j \partial_E f_n^o \, d^3v \qquad (3.2.23)$$

so

$$\vec{J}_n = \frac{2n_o}{3m_n}\left[\left(\frac{\tau_n^{(2)}}{\beta^2} - e\psi\frac{\tau_n^{(1)}}{\beta}\right)\partial_{\vec{r}}\beta + \frac{\tau_n^{(1)}}{\beta}\partial_{\vec{r}}(-\beta\phi_n)\right] . \qquad (3.2.24)$$

Next we look at second moments and find

$$\vec{J}_{v_n} = \frac{2n_o}{3m_n}\left[\left(\frac{\tau_n^{(3)}}{\beta^3} - 2e\psi\frac{\tau_n^{(2)}}{\beta^2} + (e\psi)^2\frac{\tau_n^{(1)}}{\beta}\right]\partial_{\vec{r}}\beta\right.$$

$$\left. + \left(\frac{\tau_n^{(2)}}{\beta^2} - e\psi\frac{\tau_n^{(1)}}{\beta}\right)\partial_{\vec{r}}(-\beta\phi_n)\right) . \qquad (3.2.25)$$

A similar equation may be written for the energy current transported by the holes. We shall assume that energy is transported by the lattice (phonons) according to the usual Fourier law

$$\vec{J}_{v_L} = -K_L\partial_{\vec{r}}T = \frac{K_L}{\beta}T\,\partial_{\vec{r}}\beta . \qquad (3.2.26)$$

The total energy current is the sum of the contributions from the electrons, holes, and phonons

$$\vec{J}_v = \vec{J}_{v_n} + \vec{J}_{v_p} + \vec{J}_{v_L} . \qquad (3.2.27)$$

Energy may be transfered from the electrons or holes to the phonons so $\vec{J}_v$, rather than the individual energy

currents is of primary concern. We are neglecting any direct coupling of the electrons to the holes in this calculation.

In the following we will be interested in three types of currents. These are the electric current density

$$\vec{J}_E = e(\vec{J}_p - \vec{J}_n) \quad , \tag{3.2.28}$$

the heat current density

$$\vec{J}_q = \vec{J}_U - \phi_n \vec{J}_n - \phi_p \vec{J}_p \quad , \tag{3.2.29}$$

and the particle current density

$$\vec{J} = \vec{J}_n + \vec{J}_p \quad . \tag{3.2.30}$$

It is perhaps interesting to note that these currents are gauge- invariant. That is if we make the replacements

$$\psi \to \psi + \text{const.}$$

$$\phi_n \to \phi_n - e \cdot \text{const}, \quad \phi_p \to \phi_p + e \cdot \text{const} \quad ,$$

then the equations and the above currents are invariant. The energy current, on the other hand is not gauge invariant. These equations will be analyzed in more detail in the next section.

CHAPTER 3. Solar Heat Pump

3.3 Phenomenological Equations

Let us summarize the results derived so far and regard them as a set of phenomenological equations. We will augument them by whatever additional assumptions are needed to close the set. From Eqs (2.24) and (2.25) we have

$$\begin{pmatrix} \vec{J}_n \\ \vec{J}_p \\ \vec{J}_u \end{pmatrix} = \begin{pmatrix} L_{11} & 0 & L_{13} \\ 0 & L_{22} & L_{23} \\ L_{13} & L_{23} & L_{33} \end{pmatrix} \begin{pmatrix} \partial\vec{r}(-\beta\phi_n) \\ \partial\vec{r}(-\beta\phi_p) \\ \partial\vec{r}\,\beta \end{pmatrix} , \qquad (3.3.1)$$

where we have introduced a set of L_{ij} coefficients

$$L_{11} = \frac{2n_o}{3m_n}\frac{\tau_n^{(1)}}{\beta} , \qquad (3.3.2)$$

$$L_{13} = \frac{2n_o}{3m_n}\left[\frac{\tau_n^{(2)}}{\beta^2} - e\psi\frac{\tau_n^{(1)}}{\beta}\right] , \qquad (3.3.3)$$

$$L_{22} = \frac{2p_o}{3m_p}\frac{\tau_p^{(1)}}{\beta} , \qquad (3.3.4)$$

$$L_{23} = \frac{2p_0}{3m_p}\left[\frac{\tau_p^{(2)}}{\beta^2} + e\psi\frac{\tau_p^{(1)}}{\beta}\right] , \tag{3.3.5}$$

$$L_{33} = \frac{2n_0}{3m_n}\left[\frac{\tau_n^{(3)}}{\beta^3} - 2e\psi\frac{\tau_n^{(2)}}{\beta^2} + (e\psi)^2\frac{\tau_n^{(1)}}{\beta}\right] +$$

$$+ \frac{2p_0}{3m_p}\left[\frac{\tau_p^{(3)}}{\beta^3} + 2e\psi\frac{\tau_p^{(2)}}{\beta^2} + (e\psi)^2\frac{\tau_p^{(1)}}{\beta}\right] + \frac{k_L}{k\beta^2} . \tag{3.3.6}$$

The L_{ij} matrix is verified to be symmetric, in agreement with the Onsager relations.

From Eqs. (2.6) and (2.12) we obtain the continuity relations for the electron, hole, and energy currents

$$\partial_{\vec{r}} \cdot \vec{J}_n = G , \tag{3.3.7}$$

$$\partial_{\vec{r}} \cdot \vec{J}_p = G , \tag{3.3.8}$$

$$\partial_{\vec{r}} \cdot \vec{J}_U = G(\phi_n + \phi_p) , \tag{3.3.9}$$

where we have set Q_n+Q_p of Eq. (2.11) equal to $\phi_n+\phi_p$. This replacement is made on physical grounds. The effect of absorbing radiation is assumed to be to create an electron on the top of the electron Fermi sea and a hole at the bottom of the hole Fermi sea. The electron and hole densities are given now by

$$n = \int f_n \, d^3v \tag{3.3.10}$$

and

$$p = \int f_p \, d^3v \quad . \tag{3.3.11}$$

Poisson's equation is

$$\nabla_{\vec{r}}^2 \psi = - \frac{4\pi e}{\epsilon_o} \left[\eta(\vec{r}) + p - n - N_{tr} \right] \quad , \tag{3.3.12}$$

where ϵ_o is the lattice dielectric constant, $\eta(\vec{r})$ is the doping density, which may be an inhomogeneous function of space, and N_{tr} denotes the contribution due to trapped carriers. Equations (1), (7) - (9) and (10) -(12) provide us with a complete set of equations to solve for the unknown variables β, ϕ_n, ϕ_p, n , p , ψ, $\vec{J}_n$, $\vec{J}_p$, and $\vec{J}_u$. A solution to this system as applied to a certain model with appropriate boundary conditions is the ultimate goal; it would provide us with the quantities that describe the system: particle currents, temperature, electro-chemical potentials, and densities. However, it is evident that a rigorous solution presents insuperable difficulties. A simplified procedure must be devised, based on physically reasonable approximations and reasonably powerful computers.

If we define an entropy current $\vec{J}_s$,

$$\vec{J}_s = -\beta\phi_n \vec{J}_n - \beta\phi_p \vec{J}_p + \beta \vec{J}_u \quad , \tag{3.3.13}$$

we may derive an expression for the entropy production rate per unit volume

$$\partial_{\vec{r}} \cdot \vec{J}_S = \vec{J}_n \cdot \partial_{\vec{r}}(-\beta\phi_n) + \vec{J}_p \cdot \partial_{\vec{r}}(-\beta\phi_p) + \vec{J}_U \cdot \partial_{\vec{r}}\beta \quad . \tag{3.3.14}$$

The terms involving $\partial_{\vec{r}}(-\beta\phi_n)$ and $\partial_{\vec{r}}(-\beta\phi_p)$ describe both diffusion and electrical conduction. Since the entropy production rate is expressible as a sum of products of currents with generalized forces one has an aposteriory justification of the Onsager relations.

Let us now consider the generation process in more detail. In order to describe the recombination process we employ the Schokley-Read mechanism[4]. Let N_t be the density of recombination centers, N_1 be the density of empty recombination centers and N_2 the density of recombination centers that have trapped a carrier. For the sake of definitness assume the sequence to be first the capture of an electron followed by the capture of a hole, leading to recombination. The process involved may be symbolized as the first recombination step

$$T_1 + n \rightarrow T_2 \quad , \tag{3.3.15}$$

the second recombination step

$$T_2 + p \rightarrow T_1 \quad , \tag{3.3.16}$$

and the photoproduction process

$$\gamma \longrightarrow n + p \quad . \tag{3.3.17}$$

The rate equation for N_2 is

$$\dot{N}_2 = C_n N_1 n - C_p N_2 p - \gamma_n N_2 + \gamma_p N_1 \quad , \tag{3.3.18}$$

and $\dot{N}_1 = -\dot{N}_2$ since $N_1 + N_2 = N_t$.

In particular, for thermal equilibrium the rate of change of N_2 must vanish, so

$$0 = C_n N_1^\circ n_\circ - C_p N_2^\circ p_\circ - \gamma_n N_2^\circ + \gamma_p N_1^\circ \quad , \tag{3.3.19}$$

where the superscript zero denotes thermal equilibrium. The net rate of change of electrons is

$$\frac{\delta n}{\delta t} = -C_n N_1 n + \gamma_n N_2 + b\varrho - b'\varrho n p - a n p \tag{3.3.20}$$

where ϱ denotes the radiation intensity. Again at equilibrium

$$0 = -C_n N_1^\circ n_\circ + \gamma_n N_2^\circ + b\varrho_\circ - b'\varrho_\circ n_\circ p_\circ - a n_\circ p_\circ \quad . \tag{3.3.21}$$

The rate of change of intensity is governed by

$$\frac{\delta \varrho}{\delta t} = -b\varrho + b'\varrho n p + a n p \quad , \tag{3.3.22}$$

and at equilibrium

$$0 = -b\varrho_0 + b'\varrho_0 n_0 p_0 + a n_0 p_0 \quad . \tag{3.3.23}$$

Hence

$$\frac{b}{n_0 p_0} = b' + \frac{a}{\varrho_0} \quad . \tag{3.3.24}$$

The black body radiation formula is

$$\frac{1}{\varrho_0} = \frac{\pi^2 c^3}{\hbar\omega^3 \Delta\omega}\left[e^{\beta\hbar\omega} - 1\right] \equiv Q\left[e^{\beta\hbar\omega} - 1\right] \quad , \tag{3.3.25}$$

Where $\Delta\omega$ is a typical bandwidth over which absorption takes place. Furthermore

$$n_0 p_0 = N_c N_v e^{-\beta E_g} \tag{3.3.26}$$

and $E_g = \hbar\omega$, so $b' = aQ$, and

$$\frac{b}{N_c N_v} = aQ \quad . \tag{3.3.27}$$

Hence

$$\frac{\delta \varrho}{\delta t} = aQ\varrho(np - n_0 p_0) + a(np - \frac{\varrho}{\varrho_0} n_0 p_0) \quad . \tag{3.3.28}$$

Also

$$\gamma_n = C_n n_o \frac{N_1^o}{N_2^o} \quad , \tag{3.3.29}$$

and

$$\gamma_p = C_p p_o \frac{N_2^o}{N_1^o} \quad , \tag{3.3.29'}$$

so

$$\frac{\delta n}{\delta t} = -C_n N_1 n + C_n n_o \frac{N_1^o}{N_2^o} + a n_o p_o \left(Q + \frac{1}{\rho_o}\right) - aQ\rho np - anp \quad . \tag{3.3.30}$$

For steady state, $\dot{N}_1 = 0$, so

$$0 = N_1 \left[C_n n + C_p p_o \frac{N_2^o}{N_1^o}\right] - N_2 \left[C_p p + C_n n_o \frac{N_1^o}{N_2^o}\right] \quad . \tag{3.3.31}$$

Hence

$$\frac{\delta n}{\delta t} = a\left[Q\rho(n_o p_o - np) + n_o p_o \frac{\rho}{\rho_o} - np\right] + \frac{N_t C_n C_p (n_o p_o - np)}{C_n\left(n + \frac{N_1^o n_o}{N_2^o}\right) + C_p\left(p + \frac{N_2^o p_o}{N_1^o}\right)} \quad . \tag{3.3.32}$$

Note that the populations are given by

$$\frac{N_1^o n_o}{N_2^o} = N_c e^{\beta(E_t - E_c)} \quad , \tag{3.3.33}$$

and

$$\frac{N_2^\circ P_0}{N_1^\circ} = N_v e^{\beta(E_v - E_t)} \quad , \tag{3.3.34}$$

where E_t denotes a trap level. Using the fact that

$$\frac{\delta n}{\delta t} = \frac{\delta p}{\delta t} = G \quad , \tag{3.3.35}$$

we obtain

$$G = a\left[N_c N_v \frac{\pi^2 c^3}{\hbar^2 \omega^2} \rho - np\left(1 + \frac{\pi^2 c^3}{\hbar^2 \omega^2}\rho\right)\right] + \frac{N_t C_n C_p (n_0 p_0 - np)}{C_n\left(n + n_0 \frac{N_1^\circ}{N_2^\circ}\right) + C_p\left(p + p_0 \frac{N_2^\circ}{N_1^\circ}\right)} \quad . \tag{3.3.36}$$

The first term represents the photogeneration and spontaneous and stimulated radiative recombination processes. The second term represents the Schockley-Read generation and recombination processes. If we neglect radiative recombination we may simplify the expression for G to

$$G = \frac{\alpha I}{\hbar\omega} + \frac{N_t\left[N_c N_v e^{-\beta\hbar\omega} - np\right]}{\frac{n+\Delta n}{C_p} + \frac{p+\Delta p}{C_n}} \quad . \tag{3.3.37}$$

Here α is defined as the optical absorption coefficient, I is the incident light intensity, and $\hbar\omega$ is the incident photon energy. We have defined

$$\Delta n \equiv N_c e^{\beta(E_t - E_c)} \tag{3.3.38}$$

and

$$\Delta p = N_v e^{\beta(E_v - E_t)} \tag{3.3.39}$$

If we identify N_{tr} of Eq. (12) with N_2 we have

$$N_2 = \frac{N_t(C_n \Delta n + C_p \Delta p)}{C_n(n+\Delta n) + C_p(p+\Delta p)} \tag{3.3.40}$$

This expression can be suitably modified if the recombination process involves capture of holes followed by capture of electrons.

CHAPTER 3. Solar Heat Pump

3.4 Analysis of the Equations

One of the systems of interest will be p-n junction. We will limit our attention to the case of extrinsic doping. Rather than attempt to solve the three dimensional equations we will restrict our attention to a one dimensional geometry. The analysis will be valid even for three dimensional systems when the transverse dimensions of the system vary slowly in space.

Since the size of the junction region is likely to be small we will analyze the phenomenological equations subject to the condition $\beta = \text{const}$. This presents a slight complication because Eq. (3.9) cannot then be strictly valid. Keeping a constant temperature is equivalent to having the junction in contact with a thermal reservoir. The reservoir will provide a source of energy to the system so, strictly speaking, one must generalize Eq. (3.9) to

$$\partial_{\vec{r}} \cdot \vec{J}_U = G(\phi_n + \phi_p) + S \qquad (1)$$

Since the reservoir, in actuality, does not exist, this is just a device for eliminating Eq. (1). It has been replaced by the equation

$$\beta = \text{const.} \tag{2}$$

We make further assumption of neglecting the charge in the recombination traps and of neglecting specific surface recombination mechanisms. The phenomenological equations then reduce to

$$J_n = -n D_n \beta \frac{d\phi_n}{dx} \quad , \tag{3}$$

$$J_p = -p D_p \beta \frac{d\phi_p}{dx} \quad , \tag{4}$$

$$J_U = -n D_n \left[\frac{\sigma_n^{(2)}}{\sigma_n^{(1)}} - e\beta\psi \right] \frac{d\phi_n}{dx} - p D_p \left[\frac{\sigma_p^{(2)}}{\sigma_p^{(1)}} + e\beta\psi \right] \frac{d\phi_p}{dx} \quad , \tag{5}$$

$$\frac{dJ_n}{dx} = G \quad , \tag{6}$$

$$\frac{dJ_p}{dx} = G \quad , \tag{7}$$

$$n = N_c e^{\beta(\phi_n + e\psi)} \quad , \tag{8}$$

$$p = N_v e^{\beta(\phi_p - e\psi)} \quad , \tag{9}$$

$$\frac{d\psi^2}{dx^2} = -\frac{4\pi e}{\epsilon_0}\left[\eta(x) + p - n\right] \quad . \tag{10}$$

We have dropped the subscripts of n and p for convenience. The diffusion constants $D_{n/p}$ are defined by

$$D_{n/p} \equiv \frac{2\bar{\epsilon}^{(1)}_{n/p}}{3 m_{n/p} \beta} \quad . \tag{11}$$

Using Eqs. (8) and (9) we may eliminate ϕ_n and ϕ_p to obtain a set of six coupled first order differential equations

$$J_n = -D_n \frac{dn}{dx} - e\beta D_n n E \quad , \tag{12}$$

$$J_p = -D_p \frac{dp}{dx} + e\beta D_p p E \quad , \tag{13}$$

$$\frac{dJ_n}{dx} = G \quad , \tag{14}$$

$$\frac{dJ_p}{dx} = G \quad , \tag{15}$$

$$\frac{d\psi}{dx} = -E \quad , \tag{16}$$

$$\frac{dE}{dx} = \frac{4\pi}{\epsilon_0} Q \quad , \tag{17}$$

where E is the electric field and Q is the charge

$$Q = e(\eta(x) + p - n) \quad . \tag{18}$$

With the exception of the term G , these equations are familiar from the theory of the p-n junction rectifier. There are no known analytic solutions to this system. We could not obtain one either. Since the first order systems are suitable for numerical integration, there are many papers which treat the problem numerically[5]. There are also many qualitative and/or asymptotic treatments in literature and textbooks[6,7,8] which provide a good understanding of the processes involved. We also created and ran our own programs and, not suprisingly, we discovered that for our purposes a qualitative analysis is best suited. Since we are rather interested in conceptual problems of existence of the effect (sun-powered heat transfer) and do not need accurate quantitative relations, we will drop the numerical analysis here and pursue more qualitative arguments in the coming sections.

CHAPTER 3. Solar Heat Pump

3.5 Thermal Currents

Since our goal is to propose a device that would pump heat, it is appropriate to review Phow the heat current $\vec{J}_q$ is connected with the quantities that describe the system,

$$\vec{J}_q = \vec{J}_U - \phi_n \vec{J}_n - \phi_p \vec{J}_p \quad . \tag{3.5.1}$$

In order to simplify this expression, let us back up a step and clarify the expressions for the Onsager coefficients. First assume that the relaxation time $\tau_{n/p}(E)$ is a power function of energy, $\tau_{n/p}(E) = \tau_{n/p} E^{\lambda}$, so that for example, in the case of electrons:

$$\frac{n_0 \tau_n^{(j)}}{\beta^j} = -\int_{E_c}^{\infty} \tau_n (E-E_c)(E-E_c)^j \frac{1}{\beta} \partial_E f_n^o (E-E_c) D_n (E-E_c)\, dE$$

$$= \tau_n D_n \frac{\lambda+j+\frac{1}{2}}{\beta^{\lambda+j+3/2}} \Phi_{\lambda+j-\frac{1}{2}} (\beta\phi_n + e\beta\psi) \quad , \tag{3.5.2}$$

where

$$D_n(E) = D_n \sqrt{E} \tag{3.5.3}$$

is the electronic density of states, f_n^o is the quasi-Fermi distribution, and

$$\Phi_\alpha(\beta X) \equiv \int_0^\infty \frac{(\beta E)^\alpha d(\beta E)}{1+e^{\beta(E-X)}} \tag{3.5.4}$$

is the Fermi integral of order α. Since

$$n_0 = \int_0^\infty \frac{D_n(E)dE}{1+e^{\beta(E-\phi_n-e\psi)}} = \frac{D_n}{\beta^{3/2}} \Phi_{1/2}(\beta\phi_n + e\beta\psi) \quad , \tag{3.5.5}$$

it is

$$\mathcal{T}_n^{(j)} = \mathcal{T}_n \frac{\lambda+j+1/2}{\beta^\lambda} \frac{\Phi_{\lambda+j-1/2}(\beta\phi_n+e\beta\psi)}{\Phi_{1/2}(\beta\phi_n+e\beta\psi)} \quad . \tag{3.5.6}$$

Using degenerate and nondegenerate approximations to the Fermi integral,

$$\Phi_\alpha^d(\beta X) = \frac{(\beta X)^{\alpha+1}}{\alpha+1} \quad , \tag{3.5.7}$$

and

$$\Phi_\alpha^{nd}(\beta X) = \Gamma(\alpha+1)\, e^{\beta X} \quad , \tag{3.5.8}$$

it is easy to recover the familiar expressions for

degenerate and classical kinetic coefficients:

$$\sigma_n^{(j)d} = \frac{3}{2}\frac{\sigma_n}{\beta^\lambda}\left(\beta\phi_n + e\beta\Psi\right)^{\lambda+j-1} , \tag{3.5.9}$$

$$\sigma_n^{(j)nd} = \frac{\sigma_n}{\beta^\lambda}\,\frac{\Gamma(\lambda+j+3/2)}{\Gamma(3/2)} . \tag{3.5.10}$$

Similar expressions follow for the case of holes:

$$p_o = \frac{D_p}{\beta^{3/2}}\,\Phi_{1/2}(\beta\phi_p - e\beta\Psi) , \tag{3.5.11}$$

$$\sigma_p^{(j)} = \sigma_p\,\frac{\lambda+j+1/2}{\beta^\lambda}\,\frac{\Phi_{\lambda+j-1/2}(\beta\phi_p - e\beta\Psi)}{\Phi_{1/2}(\beta\phi_p - e\beta\Psi)} . \tag{3.5.12}$$

To simplify notation, let us introduce

$$\Phi_\alpha^n \equiv \Phi_\alpha(\beta\phi_n + e\beta\Psi) , \tag{3.5.13}$$

$$\Phi_\alpha^p \equiv \Phi_\alpha(\beta\phi_p - e\beta\Psi) , \tag{3.5.14}$$

and

$$F_{n/p} \equiv \frac{2\,D_{n/p}\,\sigma_{n/p}}{3\,m_{n/p}\,\beta} . \tag{3.5.15}$$

Now rewrite the expressions for the Onsager coefficients

$$L_{11} = F_n\,\frac{\lambda+3/2}{\beta^{\lambda+3/2}}\,\Phi_{\lambda+1/2}^n , \tag{3.5.16}$$

$$L_{22} = F_p\,\frac{\lambda+3/2}{\beta^{\lambda+3/2}}\,\Phi_{\lambda+1/2}^p , \tag{3.5.17}$$

$$L_{13} = F_n\left[\frac{\lambda+5/2}{\beta^{\lambda+5/2}}\Phi^n_{\lambda+3/2} - e\psi\,\frac{\lambda+3/2}{\beta^{\lambda+3/2}}\Phi^n_{\lambda+1/2}\right], \tag{3.5.18}$$

$$L_{23} = F_p\left[\frac{\lambda+5/2}{\beta^{\lambda+5/2}}\Phi^p_{\lambda+3/2} + e\psi\,\frac{\lambda+3/2}{\beta^{\lambda+3/2}}\Phi^p_{\lambda+1/2}\right], \tag{3.5.19}$$

$$L_{33} = F_n\left[\frac{\lambda+7/2}{\beta^{\lambda+7/2}}\Phi^n_{\lambda+5/2} - 2e\psi\,\frac{\lambda+5/2}{\beta^{\lambda+5/2}}\Phi^n_{\lambda+3/2} + (e\psi)^2\frac{\lambda+3/2}{\beta^{\lambda+3/2}}\Phi^n_{\lambda+1/2}\right] +$$

$$F_p\left[\frac{\lambda+7/2}{\beta^{\lambda+7/2}}\Phi^p_{\lambda+5/2} + 2e\psi\,\frac{\lambda+5/2}{\beta^{\lambda+5/2}}\Phi^p_{\lambda+3/2} + (e\psi)^2\frac{\lambda+3/2}{\beta^{\lambda+3/2}}\Phi^p_{\lambda+1/2}\right] + \frac{k_L T}{\beta}. \tag{3.5.20}$$

In order to simplify further calculations, define

$$K^{n/p}_{j1} \equiv F_{n/p}\,\frac{\lambda+j+1/2}{\beta^{\lambda+j+1/2}}\Phi^{n/p}_{\lambda+j-1/2}. \tag{3.5.21}$$

Then

$$\vec{J}_n = -K^n_{11}\beta\,\partial_{\vec{r}}\phi_n + \left[K^n_{21} - \psi_n K^n_{11}\right]\partial_{\vec{r}}\beta, \tag{3.5.22}$$

$$\vec{J}_p = -K^p_{11}\beta\,\partial_{\vec{r}}\phi_p + \left[K^p_{21} - \psi_p K^p_{11}\right]\partial_{\vec{r}}\beta, \tag{3.5.23}$$

$$\vec{J}_U = \left[K^n_{31} + K^p_{31} + \frac{k_L T}{\beta}\right]\partial_{\vec{r}}\beta - K^n_{21}\left[2\psi_n\,\partial_{\vec{r}}\beta + \beta\,\partial_{\vec{r}}\phi_n\right] +$$

$$K^n_{11}\psi_n\left[\psi_n\partial_{\vec{r}}\beta + \beta\,\partial_{\vec{r}}\phi_n\right] - K^p_{21}\left[2\psi_p\partial_{\vec{r}}\beta + \beta\,\partial_{\vec{r}}\phi_p\right] + K^p_{11}\psi_p\left[\psi_p\partial_{\vec{r}}\beta + \beta\,\partial_{\vec{r}}\phi_p\right], \tag{3.5.24}$$

with

$$\psi_n \equiv \phi_n + e\psi, \qquad \psi_p \equiv \phi_p - e\psi. \tag{3.5.25}$$

After some rearranging,

$$\vec{J}_2 = \left[K_{31}^n - \psi_n K_{21}^n\right]\partial_{\vec{r}}\beta - \left[K_{21}^n - \psi_n K_{11}^n\right]\left[\psi_n \partial_{\vec{r}}\beta + \beta\partial_{\vec{r}}\phi_n\right] +$$

$$+\left[K_{31}^p - \psi_p K_{21}^p\right]\partial_{\vec{r}}\beta - \left[K_{21}^p - \psi_p K_{11}^p\right]\left[\psi_p \partial_{\vec{r}}\beta + \beta\partial_{\vec{r}}\phi_p\right] + \frac{k_L T}{\beta}\partial_{\vec{r}}\beta \quad . \qquad (3.5.26)$$

Since

$$\psi_n \partial_{\vec{r}}\beta + \beta\partial_{\vec{r}}\phi_n = \frac{1}{K_{11}^n}\left[K_{21}^n \partial_{\vec{r}}\beta - \vec{J}_n\right] , \qquad (3.5.27)$$

and

$$\psi_p \partial_{\vec{r}}\beta + \beta\partial_{\vec{r}}\phi_p = \frac{1}{K_{11}^p}\left[K_{21}^p \partial_{\vec{r}}\beta - \vec{J}_p\right] , \qquad (3.5.28)$$

we can write

$$\vec{J}_2 = \frac{K_{31}^n K_{11}^n - K_{21}^{n\,2}}{K_{11}^n}\partial_{\vec{r}}\beta + \frac{K_{21}^n - \psi_n K_{11}^n}{K_{11}^n}\vec{J}_n +$$

$$\frac{K_{31}^p K_{11}^p - K_{21}^{p\,2}}{K_{11}^p}\partial_{\vec{r}}\beta + \frac{K_{21}^p - \psi_p K_{11}^p}{K_{11}^p}\vec{J}_p + \frac{k_L T}{\beta}\partial_{\vec{r}}\beta \qquad (3.5.29)$$

The last equation, written in the form

$$\vec{J}_2 = \frac{k_{TOT} T}{\beta}\partial_{\vec{r}}\beta + P_n \vec{J}_n + P_p \vec{J}_p \qquad (3.5.29')$$

with

$$K_{TOT} = K_n + K_p + K_L \tag{3.5.30}$$

is a familiar equation for the heat current expressed as the sum (actually, the difference) of heat currents caused by the temperature gradient and particle currents. $k_{n/p}$ and $P_{n/p}$ are the carrier thermal conductivities and Peltier coefficients,

$$k_{n/p} \equiv \beta \frac{K_{31}^{n/p} K_{11}^{n/p} - K_{21}^{n/p\,2}}{T K_{11}^{n/p}} , \tag{3.5.31}$$

$$P_{n/p} \equiv \frac{K_{21}^{n/p} - \psi_{n/p} K_{11}^{n/p}}{K_{11}^{n/p}} = \frac{\lambda+5/2}{\beta(\lambda+3/2)} \frac{\Phi_{\lambda+3/2}^{n/p}}{\Phi_{\lambda+1/2}^{n/p}} - \psi_{n/p} . \tag{3.5.32}$$

In Fig. 2 we have provided a graph of the function $\beta P_p(\beta x)$. Note that the large negative βx corresponds to the non-degenerate case and large positive βx to the degenerate case. The theoretical discussion provided in these past few sections provides the basic equations necessary to discuss a general semiconductor device which is to act as a solar powered heat pump. We now turn to the more practical question of analyzing a particular device which we shall propose.

CHAPTER 3. Solar Heat Pump

3.6 The Device

Most of the content of the present section is taken from the Ref. 9. A cross-sectional view of the proposed device is presented in Fig. 3. Light (1) impinges on the top of the device, passes through a metallic mesh (2) and through a semiconductor layer (3) onto a semiconductor substrate (4) whose conductivity type is different from the layer above it. At the sides of the aforementioned semiconductors is a region of thermal and electrical insulating material (6). On the other side of this insulating material is semiconductor material (5) whose impurity doping density is different than in (4). At the base of the device is a metallic layer (7).

By way of example, consider the case where one semiconductor (3) is of n type, another (4) is of p type with doping density denoted by N_A , and the third semiconductor

(5) is of p type with doping density denoted by N'_A . The photocurrent that is generated at the interface between regions (3) and (4) circulates through the inner p-type semiconductor (4), through the metallic base (7), up through the outer p'-type semiconductor (5), through the metallic mesh (2), and back through the n-type semiconductor (3). In regions (4) and (5) the majority carriers are holes. In addition to carrying electrical charge, these carriers transport thermal energy. Let us apply expression for the thermal current to the device model. For the p-type region

$$J_q^p = -\frac{k_p T}{\beta}\partial_x \beta + P J_p \tag{3.6.1}$$

whereas for the p' region

$$J_q^{p'} = \frac{k'_p T}{\beta}\partial_x \beta + P' J'_p \tag{3.6.2}$$

If the doping densities are arranged properly, more heat may be transported by the carriers upward through (5) than are transported downward through (4). This is the physical origin of the Peltier effects that occur at the junctions between regions (4) and (7) and between regions (5) and (7).

In addition to the above thermal flows, there exists a thermal flow associated with the phonons of semiconductors (3)-(5). Material (6) is chosen to have a sufficiently low

thermal conductivity so its contribution to the heat flux may be neglected.

Let the cross-sectional area in a direction perpendicular to the symmetry axis of Fig. (3) be denoted by A_i , $i = 2,\ldots,7$ and let k_i denote the total thermal conductivities of the various materials. Assume for the moment that β = const.; The whole device is in contact with a reservoir and just starting to operate. Then the net rate of heat transported up is

$$\Delta J_2 = P' J_p' - P J_p = (P' - P) J_p \qquad (3.6.3)$$

since the particle current is the same. We want $\Delta J_2 > 0$, and it is seen that the effect is largest when the p material is in the nondegenerate region and p' is in the degenerate region. Then

$$P' = \Psi_p' \qquad (3.6.4)$$

while βP is constant in the same order of approximation. It is unlikely that $\beta P = \lambda + 5/2$ is very meaningful in a real situation, however we will use it as an estimate. Now that we are in the degenerate region, we may solve for using the appropriate limit

$$p' = N_V' (\beta \Psi_p')^{3/2} \qquad (3.6.5)$$

Effectively, in the bulk p' material the hole density equals the doping density N_A' , so that

$$\beta\psi_p' = \left(\frac{N_A'}{N_V'}\right)^{2/3} \quad . \tag{3.6.6}$$

Thus the pumping rate is given with

$$\Delta J_2 = \Delta P J_p = \frac{\left(N_A'/N_V'\right)^{2/3} - (\lambda+5/2)}{\beta} J_p \tag{3.6.7}$$

As the device starts to operate, a temperature difference develops and both terms in J_2 are supposed to be taken into account. The net heat current flowing up through (5) is then

$$I_2' = P' I_p - \frac{K_5 A_5}{H_5}(T_1 - T_2) \quad . \tag{3.6.8}$$

The net heat current flowing down through the lower part of (4) is

$$I_2 = P I_p + \frac{K_4 A_4}{H_4}(T_1 - T_2) \tag{3.6.9}$$

Where T_1 is the temperature of the metallic mesh , and T_2 is the temperature of the metallic base . The net heat current upward through the device is

$$\Delta I_2 = (P'-P) I_p - \left(\frac{K_4 A_4}{H_4} + \frac{K_5 A_5}{H_5}\right)(T_1 - T_2) \tag{3.6.10}$$

For refrigeration of the base to occur, We must have $\Delta I_2 > 0$. Let us now estimate the maximum temperature difference this device might produce. When operating at this maximum temperature difference, no heat is removed from the cold end, so that

$$\Delta I_2 = \Delta P I_p - \left(\frac{k_4 A_4}{H_4} + \frac{k_5 A_r}{H_5} \right) \Delta T_{max} = 0 \qquad (3.6.11)$$

To get a quantitative estimate we need an expression for I_p.

To this end let us review the theory of the photovoltaic effect in p-n junctions[10]. We consider the junction with no external voltage and ask what effects will take place when light produces electron-hole pairs. The pairs within a diffusion length on the two sides diffuse to the barrier, where they are separated by the electrostatic field of the junction. The holes float into the p-region, the electrons sink to the n-region, and in doing so they produce a voltage which tends to lower the junction barrier, just as if a forward voltage were to be applied. In principle if there were no dark current, the charge separation could continue until the barrier were completely wiped out. Since the lowering of the barrier tends to produce a forward current in the junction, a balance of two effects is soon reached.

Let us consider an equivalent circuit diagram for the device, presented in Fig. 4 . The photocurrent is

represented by a current generator I_{ph} and is opposite in direction to the forward bias current of the junction I_d. R is the electrical resistance of the circuit and is approximately

$$R = \frac{1}{e}\left(\frac{H_4}{N_A u_4 A_4} + \frac{H_5}{N_A' u_5 A_5}\right) \qquad (3.6.12)$$

where u_4 and u_5 are the majority-carrier mobilities of semiconductors (4) and (5). The current I_d is the injection current which would flow through the junction under the influence of a forward bias

$$I_d = I_o\left(e^{e\beta V} - 1\right) \quad . \qquad (3.6.13)$$

Here I_0 is the saturation current, representing free carriers which can flow through the junction, overcoming the junction barrier under the influence of thermal activation.

Let g denote the number of hole-electron pairs generated per unit volume per unit time in the junction region. It is related to the incident flux I through the relation

$$g = \frac{q I}{hf H_3} \qquad (3.6.14)$$

where hf is the mean energy of an incident photon and q is the quantum efficiency for electron-hole production. The

photocurrent is then

$$I_{ph} = g e L A_4 \tag{3.6.15}$$

where L is the diffusion length of the junction. The output current is given by (see Fig. 4)

$$I_{OUT} = I_{ph} - I_o (e^{e\beta V} - 1) \ . \tag{3.6.16}$$

The short circuit current is given simply by

$$I_{sc} = I_{ph} \tag{3.6.17}$$

and the open circuit voltage is obtained when there is no output current

$$V_{oc} = \frac{1}{e\beta} \ln \left(\frac{I_{sc}}{I_o} + 1 \right) \ . \tag{3.6.18}$$

Since the power output is $V_{OUT} I_{OUT}$, the maximum power output can be found by differentiating the product and setting the result equal to zero

$$P_{OUT}^{MAX} = I_m V_m \tag{3.6.19}$$

where

$$I_m = (I_{sc} + I_o) \frac{e\beta V_m}{1 + e\beta V_m} \tag{3.6.20}$$

is the current output at the maximum power and

$$e^{e\beta V_m}(1+e\beta V_m) = 1 + \frac{I_{sc}}{I_o} = e^{e\beta V_{oc}} \qquad (3.6.21)$$

allows the voltage at maximum power output to be calculated. Referring back to Eq. (11), for the highest power output we should use I_m as I_{OUT} . However, for the highest temperature difference it is best to be as close as possible to I_{sc} . As an estimate we will use I_{sc} for I_{OUT} . Then

$$\Delta T_{MAX} = \frac{k_B T_2^2 I L A_4 \left[\left(N_A'/N_D' \right)^{2/3} - (\lambda + 5/2) \right]}{h f H_3 \left(\frac{k_4 A_4}{H_4} + \frac{k_5 A_5}{H_5} \right)} \quad . \qquad (3.6.22)$$

A criterion for the device to be non-Ohmic is that the gap energy for semiconductor (4) be larger than the Ohmic voltage drop

$$E_g \gtrsim eRI_{ph} \quad . \qquad (3.6.23)$$

Now let us consider the operation of the device in the heat-pump mode. Referring back to Eq. (3), assume that the doping in the semiconductors is such that $P' < P$. The heat current will now be negative, indicating a flow away from the iluminated side - a heat pumping action.

In summary, we have proposed an integrated solid-state structure which acts as a solar-powered heat transfer

device. Depending on the exact design specifications, the device may act as a refrigerator or a heat pump.

Figure captions

Figure 1. Schematic thermodynamic diagram depicting the flows of energy. The upper rectangle denotes a thermal reservoir at temperature T_1 and the lower rectangle symbolizes a reservoir at temperature T_2 . The circle labeled A represents a photovoltaic junction and the circle labeled B represents a set of Peltier junctions.

Figure 2. A plot of the Peltier coefficients for some λ . In heat transfer processes only Π is to be used. Note that the large negative βx corresponds to the nondegenerate case and large positive βx to the degenerate case.

Figure 3. A cross-sectional view of the proposed device. Light (1) passes through a metallic mesh (2), through a thin layer of semiconductor (3), and illuminates the photovoltaic junction (the interface between (3) and (4)). On the sides is a semiconducting shaft (5). At the bottom is a metallic base (7). A thermal and electric insulator (6) separates the semiconductor columns.

Figure 4. (a) An equivalent circuit of an ideal solar cell, (b) A realistic equivalent circuit, with series and shunt resistances. (c) Volt-ampere characteristic of an ideal solar cell - (1) in the dark, (2) when illuminated.

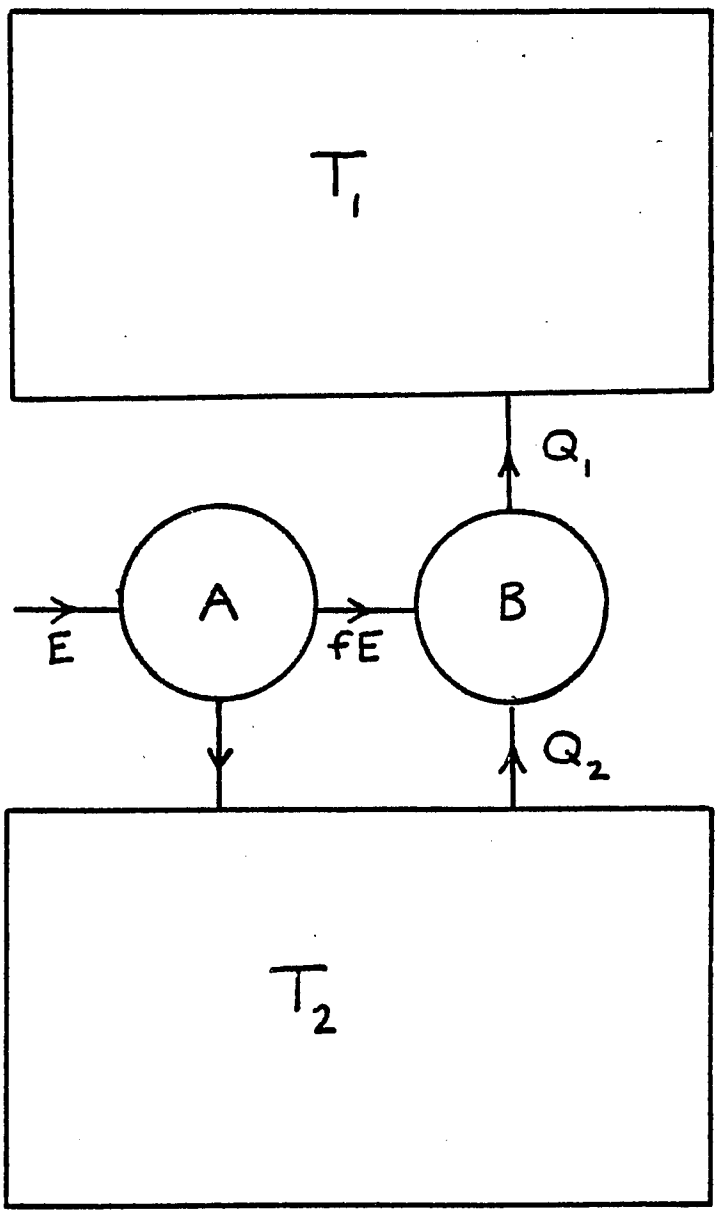


Figure 1.

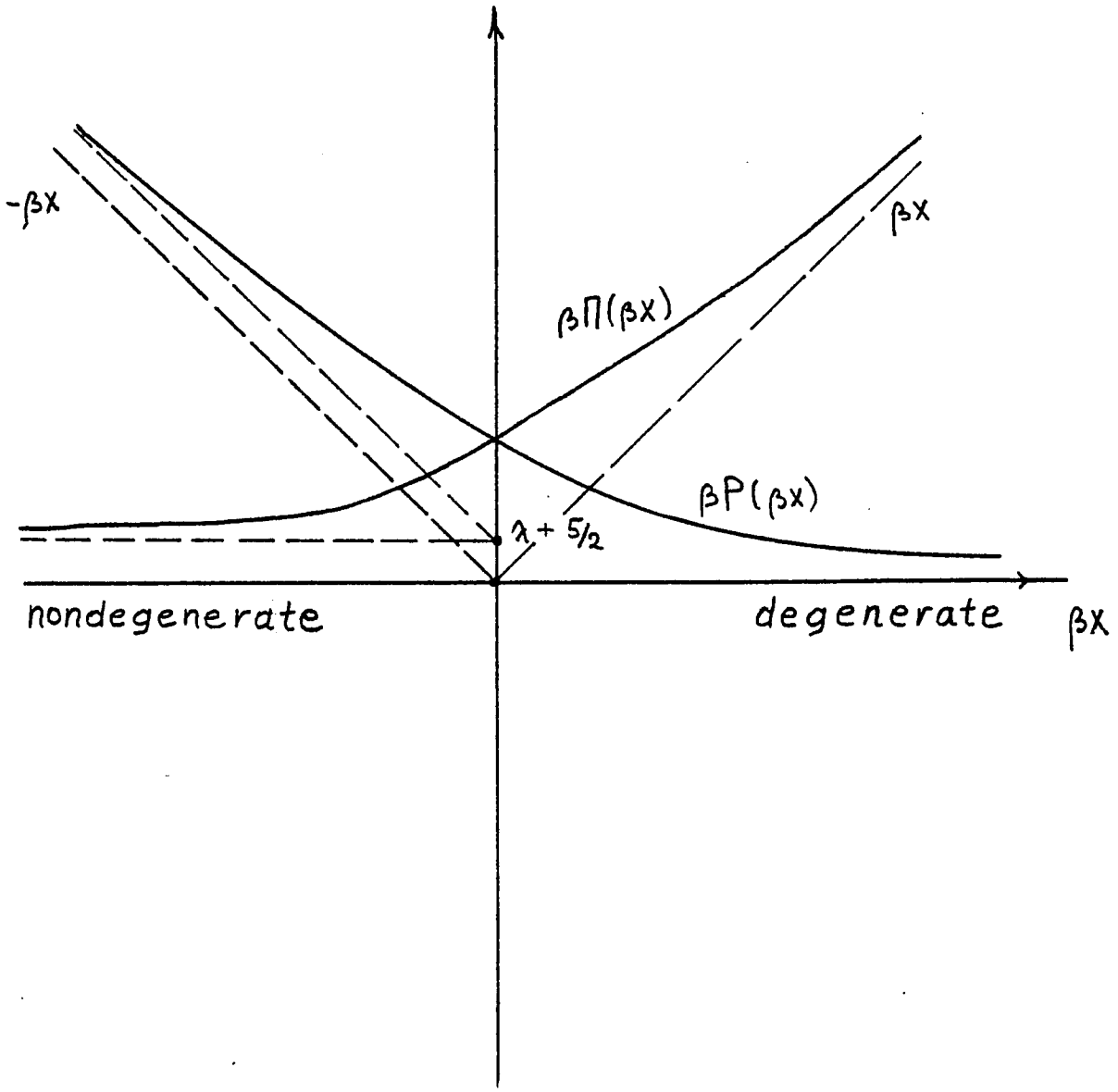


Figure 2

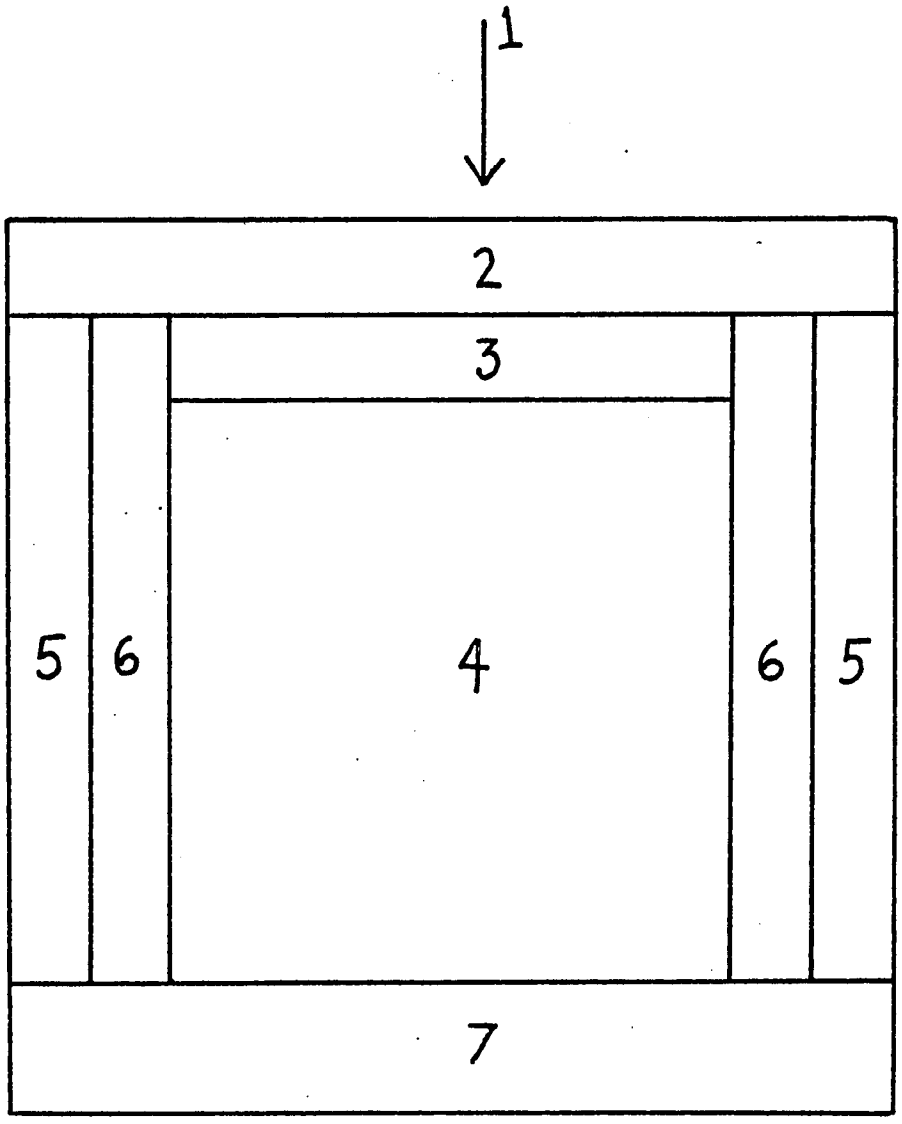
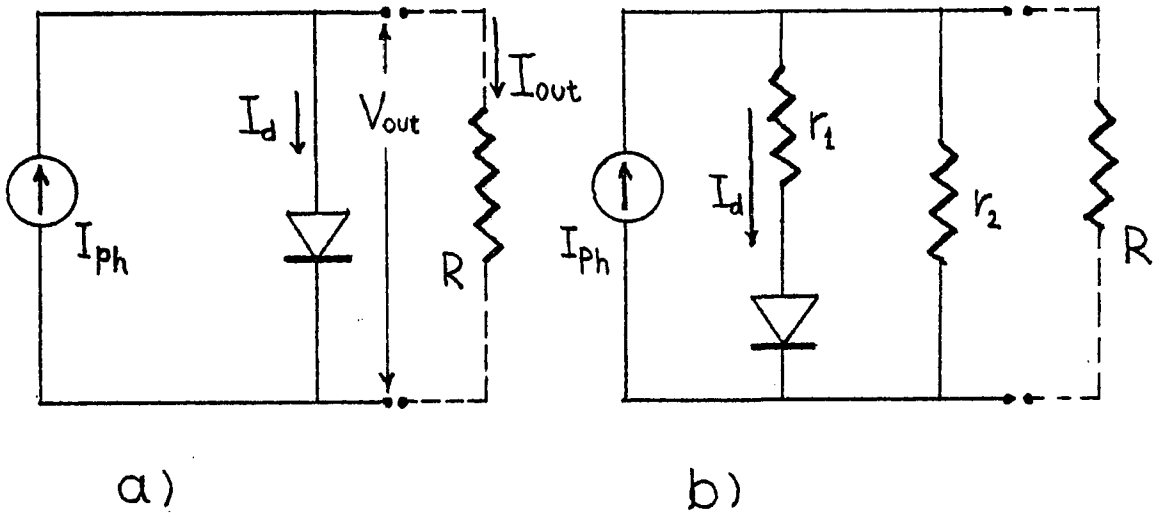


Figure 3.



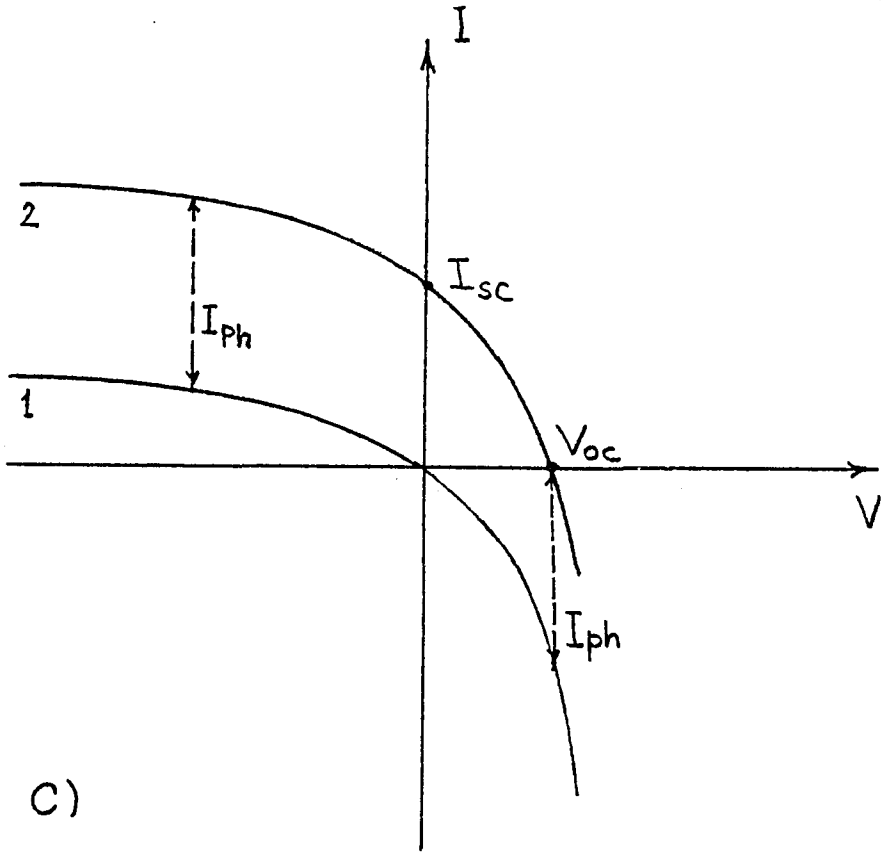


Figure 4.

References

1. R. T. Keyes and T. M. Quist, Proc. IRE 50, 1822 (1962)

2. G. S. Dousmanis, C. W. Mueller, H. Nelson, and K. G. Petzinger, Phys. Rev. 133, A316 (1964)

3. J. I. Pankove in "Optical Processes in Semiconductors" (Dover Publications, New York 1971), p. 177

4. W. Shockley and W. T. Read, Jr., Phys. Rev. 87, 835 (1952)

5. W. M. Bullis and W. R. Runyan, IEEE Trans. El. Dev. ED-14, 75 (1967); S. Kaye and G. P. Rolik, IEEE Trans. El. Dev. ED-13, 563 (1966); R. van Overstraeten and W. Nuyts, IEEE Trans. El. Dev. ED-16, 632 (1969)

6. M. Wolf, Proc. IEEE 51, 674 (1963); B. Ellis and T. S. Moss, Solid State Electron. 13, 1 (1970); R. L. Cummerow, Phys. Rev. 95, 16 (1954)

7. See, for example J. Dunlap, "An Introduction to Semiconductors" or R. R. Heikes and R. Ure, Jr., "Thermoelectricity" (Intersc. Publishers, New York 1961), or A. C. Smith and R. L. Longini, "Introduction to Semiconductor Physics" (John Wiley & Sons, New York 1969), etc.

8. H. J. Hovel, in "Solar Cells", ed. by R. K. Willardson and A. C. Beer (Academic, New York 1975)

9. M. Belic and J. I. Gersten, J. Appl. Phys 50, 5682 (1979)

10. J. Tauc, "Photo and Thermoelectric Effects in Semiconductors" (Oxford, New York 1962)

CHAPTER 4. Band Spectrum of a Solid in a Laser

4.1 Introduction

The system consisting of a solid in an external laser field presents a rather interesting challenge to the theoretical physicist. On the one hand the spatial periodicity of the solid leads to the conventional band theory of solids. The one electron states are characterized by the crystal momentum whose value is restricted to the first Brillouin zone and a band index (as well as a spin index). On the other hand the temporal periodicity imposed by the laser field combined with a Floquet analysis of the problem leads one to a picture in which one need only consider the band structure in a range $0 \leq \epsilon \leq \hbar\omega$. A preliminary analysis of this problem has been presented in the literature[1].

A basic question remains to be answered about this problem. What is the mathematical nature of the spectrum? Since the band structure problem has been reduced to a knowledge of the spectrum in the hyper-rectangle defined by

$$\vec{k} \in 1^{st} \text{ Brillouin Zone} \; ; \quad 0 \leq \epsilon \leq \hbar\omega \tag{4.1.1}$$

how are the eigenvalues distributed? Are they dense (i.e. do they completely fill the hyper-rectangle)? Do gaps exist? Are there discrete eigenvalues? Are the residues of the Green's function dense? These possibilities are illustrated in Fig.1, where hypothetical band structures are depicted. A rigorous answer to these questions is difficult to obtain, so we set out on a more modest course. We studied a simple model which displays all the difficulties of the general problem, and tried to give some answers. Then we addressed the general case, stated the problems and outlined some avenues for future research.

CHAPTER 4. Band spectrum

4.2 Modified Kronig-Penney Model

The model to be considered is a simple generalization of the one dimensional Kronig-Penney model :

$$H = \frac{p^2}{2M} - \frac{2\pi V}{b}\sum_{m=-\infty}^{\infty}\delta(z-ma) - \frac{2\pi AP}{\omega}\sum_{n=-\infty}^{\infty}\delta(t-n\tau) \quad ,$$

$$ab = \omega\tau = 2\pi \tag{4.2.1}$$

This Hamiltonian describes a one dimensional solid which is periodically perturbed by a train of electric field pulses. The strength of the electric field is proportional to A and the strength of the crystalline field is proportional to V. We look for solutions to the Schrodinger equation

$$H\psi = i\hbar\partial_t\psi \tag{4.2.2}$$

which are of the Floquet-Bloch form:[2]

$$\psi(z,t) = U(z,t)e^{i(kz-\epsilon t)} \tag{4.2.3}$$

The function U is doubly periodic, i.e.

$$U(z,t)=U(z+a,t)=U(z,t+\tau) \quad . \tag{4.2.4}$$

Before going into details of the calculations, it might be appropriate to review the ordinary Kronig-Penney model, which we may treat as one extreme case ($A\to 0$) of our model. The Hamiltonian of the ordinary KP model is given by

$$H_0=\frac{p^2}{2M}-\frac{2\pi V}{b}\sum_m \delta(z-ma) \tag{4.2.5}$$

The solution is sought in form of Bloch-waves

$$\psi(z)=u(z)e^{ikz} \quad , \tag{4.2.6}$$

where u is periodic, and can be Fourier-decomposed

$$u(z)=\sum_n e^{inbz}u_n \quad . \tag{4.2.7}$$

When this solution is substituted, a system of equations for the Fourier coefficients follows

$$\sum_n \left\{\delta_{mn}\left[\frac{\hbar^2}{2M}(k+mb)^2-E\right]-V\right\}u_n=0 \tag{4.2.8}$$

and the allowed energies come from the secular equation

$$\det \Sigma(E,k) = 0 \quad , \tag{4.2.9}$$

where

$$\Sigma_{mn}(E,k) = \delta_{mn} - \frac{V}{\frac{\hbar}{2M}(k+mb)^2 - E} \quad . \tag{4.2.10}$$

In a customary fashion these allowed energies are grouped into a series of energy bands separated by a series of energy gaps. However, by a slight manipulation, a more explicit dispersion relation can be obtained. Denote

$$L \equiv \sum_n u_n \tag{4.2.11}$$

Then from Eq. (8) it immediately follows

$$L - VL \sum_m \frac{1}{\frac{\hbar^2}{2M}(k+mb)^2 - E} = 0 \quad , \tag{4.2.12}$$

and the dispersion relation takes the form

$$1 - V\,\zeta(E,k) = 0 \tag{4.2.13}$$

where

$$\zeta(E,k) \equiv \sum_m \frac{1}{\frac{\hbar^2}{2M}(k+mb)^2 - E}$$

$$= \frac{M\pi}{\hbar b\sqrt{2ME}}\left[\cot\frac{\pi}{b}\left(k - \frac{1}{\hbar}\sqrt{2ME}\right) - \cot\frac{\pi}{b}\left(k + \frac{1}{\hbar}\sqrt{2ME}\right)\right] . \tag{4.2.14}$$

The series summation is easily accomplished by the residue method. This form is equivalent to the more customary form

$$\cos ka = \cos \frac{a}{\hbar}\sqrt{2ME} - \frac{MV}{\hbar\sqrt{2ME}} \sin \frac{a}{\hbar}\sqrt{2ME} \qquad (4.2.15)$$

usually derived in textbooks.[3]

The other extreme case ($V \to 0$) of our model is treated similarly. The allowed energies come from

$$1 - \frac{Ap\pi}{\hbar\omega} \cot \frac{\pi}{\hbar\omega}\left(\frac{p^2}{2M} - \hbar\epsilon\right) = 0 \qquad (4.2.16)$$

and we have a continuous spectrum. Note that ϵ now stands for quasi-energy.

Now let us turn to the general MKP. The doubly periodic function U in Eq. (4) can be Fourier decomposed in z and t variables.

$$U(z,t) = \sum_{m'n'} u_{m'n'} e^{i(m'bz - n'\omega t)} \,. \qquad (4.2.17)$$

When this solution is substituted into Eq. (1) a doubly infinite set of algebraic equations follows for the coefficients u_{mn}.

$$\Delta_{mn} u_{mn} = V \sum_{m'} u_{m'n} + \hbar(k+mb)A \sum_{n'} u_{mn'} \,, \qquad (4.2.18)$$

where

$$\Delta_{mn} \equiv \frac{\hbar^2}{2M}(k+mb)^2 - \hbar(\epsilon+n\omega) \quad . \tag{4.2.19}$$

A further simplification is achieved by letting

$$L_n \equiv \sum_{m'} u_{m'n} \quad , \tag{4.2.20}$$

$$R_m \equiv \sum_{n'} u_{mn'} \quad . \tag{4.2.21}$$

After some algebra one finds a system of equations for the singly subscripted variable R_m

$$\sum_m A_{mm'} R_{m'} = 0 \quad , \tag{4.2.22}$$

where

$$A_{mm'} = \delta_{mm'} - \frac{V}{1-\hbar A(k+mb)\eta_m} \sum_n \frac{\hbar(k+m'b)}{\Delta_{mn}\Delta_{m'n}(1-V\zeta_n)} \tag{4.2.23}$$

and

$$\zeta_n(\epsilon,k) \equiv \sum_m \frac{1}{\Delta_{mn}} =$$

$$\frac{M\pi}{\hbar b\sqrt{2M\hbar(\epsilon+n\omega)}}\left[\cot\frac{\pi}{b}\left(k-\sqrt{\tfrac{2M}{\hbar^2}(\epsilon+n\omega)}\right) - \cot\frac{\pi}{b}\left(k+\sqrt{\tfrac{2M}{\hbar^2}(\epsilon+n\omega)}\right)\right] \quad , \tag{4.2.24}$$

and

$$\eta_m(\epsilon,k) \equiv \sum_n \frac{1}{\Delta_{mn}} = \frac{\pi}{\hbar\omega} \cot \frac{\pi}{\hbar\omega} \left[\frac{\hbar^2}{2M} (k+mb)^2 - \hbar\epsilon \right] \qquad (4.2.25)$$

which is evaluated as before. An analogous system for the variable L_n is obtained in a similar manner

$$\sum_{n'} \mathcal{B}_{nn'} L_{n'} = 0 \quad , \qquad (4.2.26)$$

where

$$\mathcal{B}_{nn'} = \delta_{nn'} - \frac{V}{1 - V\xi_n} \sum_m \frac{\hbar(k+mb)A}{1 - \hbar(k+mb)A\eta_m} \frac{1}{\Delta_{mn}\Delta_{mn'}} \quad . \qquad (4.2.27)$$

Spectra of A and B, defined by

$$\det \mathcal{A} = 0 \quad , \quad \det \mathcal{B} = 0 \qquad (4.2.28)$$

are identical. An exact evaluation of these spectra, however turned out to be very difficult and not within our reach. Therefore, from now on our analysis will rather be qualitative and/or numerical in nature.

Firstly, let us consider the system without interaction - in some units the electron's energy is simply $E = k^2$. Its quasi-energy is depicted in Fig. 2, where we assumed $b = 1$ and $\omega = 1$ for simplicity. The quasi-energy is obtained in a standard manner - by slicing the parabola and bringing its pieces down to the principal hyper-rectangle.

The pieces cross each other at many points and the density of crossing points is not uniform , although the parabola segments cover the whole hyper-rectangle fairly uniformly, producing a dense spectrum. In other words, the electron can find itself arbitrarily close to any point in the hyper-rectangle. The number of crossings is at most denumerable. As the interaction is turned on, gaps open at the crossings. However, since most of the crossings are multiple (i.e. more than two segments cross at the same point), it is not clear in which way the gaps open, or in which way the band splitting takes place. Therefore we did not try to present a more detailed band structure in Fig. 2. Also it is not clear in which way the gaps change the appearence of the spectrum. It seems unreasonable that they would eject band lines from any (wide) area in the hyper-rectangle.

In order to support these qualitative statements we have developed a numerical truncation algorithm to solve for the roots of Eqs. (22) and (23). The program computes the determinant as a function of the quasi-energy for a given k . One may vary the size of the matrix whose determinant is to be evaluated. In the limit of large matrix size one approaches the solid-in-laser problem. Our goal is to study the distribution of the zeros of the determinant. In Fig. 2 a low order case is presented, where three space cells and three time cells were involved. Small circles on the figure represent the numerical points. We ran higher order cases as

well. It appears that roots occur for very many quasi-energy values, lending credence to the conjecture that the spectrum may, in fact, be dense in the hyper-rectangle. In order for this fact to be fully established, however, one must await a rigorous mathematical proof.

CHAPTER 4. Band spectrum

4.3 General Theory

Consider a solid interacting with an electromagnetic field. The one electron Hamiltonian in the Coulomb gauge and MKS units is of the form

$$H(\vec{p},\vec{r},t)=\frac{1}{2M}\left(\vec{p}-e\vec{A}(\vec{r},t)\right)^2+V(\vec{r}) \quad , \tag{4.3.1}$$

where

$$V(\vec{r}+\vec{R})=V(\vec{r}) \tag{4.3.2}$$

is the crystal potential, $\vec{R}$ being any lattice vector, and e is the electronic charge. We will consider here the case of the space-homogeneous time-periodic vector potential

$$\vec{A}(t+\tau)=\vec{A}(t) \tag{4.3.3}$$

In this case the problem of solving the time-dependent Schrodinger equation

$$i\hbar\partial_t \Psi_t\rangle = H(\vec{p},\vec{r},t)\,\Psi_t\rangle \tag{4.3.4}$$

is simplified by the use of the Floquet theory[2] of differential equations with periodic coefficients. Accordingly, the state vector may be chosen as

$$\Psi_t\rangle = e^{-i\epsilon t}\, U_t\rangle \ , \tag{4.3.5}$$

where

$$U_{t+\tau}\rangle = U_t\rangle \ , \tag{4.3.6}$$

and upon performing the Fourier transformation in t and the Bloch analysis in $\vec{r}$, the problem is reduced to the evaluation of quasi-energy bands

$$\epsilon_n = \epsilon_n(\vec{k},\omega) \ . \tag{4.3.7}$$

Before going to details of the evaluation, let us use the general statement of gauge invariance

$$H'(\vec{p},\vec{r},t) \equiv e^{i\lambda(\vec{p},\vec{r},t)}\, H(\vec{p},\vec{r},t)\, e^{-i\lambda(\vec{p},\vec{r},t)}$$

$$= H(\vec{p}-\hbar\partial_{\vec{r}}\lambda\ ,\vec{r}+\hbar\partial_{\vec{p}}\lambda\ ,t) - \hbar\partial_t\lambda(\vec{p},\vec{r},t) \tag{4.3.8}$$

to pick the gauge most convenient for calculations. By way

of an example, just note that the choice

$$\lambda_1(\vec{r},t) = -\frac{e}{\hbar}\vec{A}\cdot\vec{r} \tag{4.3.9}$$

brings up the $\vec{E}\cdot\vec{r}$ gauge

$$H'(\vec{p},\vec{r},t) = \frac{1}{2M}p^2 + V(\vec{r}) - e\vec{E}\cdot\vec{r} \quad , \tag{4.3.10}$$

another choice

$$\lambda_2(t) = \frac{e^2}{2M\hbar}\int^t A^2(t')\,dt' \tag{4.3.11}$$

brings up the $\vec{A}\cdot\vec{p}$ gauge, and the choice

$$\lambda_3(\vec{p},t) = -\int\frac{e}{M\hbar}\vec{A}\cdot\vec{p}\,dt' + \frac{e^2}{2M\hbar}\int A^2\,dt' \tag{4.3.12}$$

brings up the Kramers gauge. We will work in the $\vec{A}\cdot\vec{p}$ gauge, in which

$$H'(\vec{p},\vec{r},t) = \frac{p^2}{2M} + V(\vec{r}) - \frac{e}{M}\vec{A}\cdot\vec{p} \quad . \tag{4.3.13}$$

Upon applying the Floquet theorem and Fourier transformation, an infinite set of operator equations is obtained

$$\left[\frac{p^2}{2M} + V(\vec{r}) - \hbar(\epsilon+n\omega)\right]|u_n\rangle = \frac{e}{M}\vec{p}\cdot\sum_m \vec{A}_{n-m}|u_m\rangle \tag{4.3.14}$$

where

$$u_n\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{in\omega t}\, U_t\rangle\, d\omega t \quad , \tag{4.3.15}$$

$$\vec{A}_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{in\omega t}\, \vec{A}(t)\, d\omega t \quad , \tag{4.3.16}$$

and

$$\omega\delta = 2\pi \quad . \tag{4.3.17}$$

Let us call

$$H_0(\vec{p},\vec{r}) \equiv \frac{p^2}{2M} + V(\vec{r}) \quad . \tag{4.3.18}$$

One way to proceed is to go to the Schrodinger picture

$$\left[H_0(-i\hbar\partial_{\vec{r}},\vec{r}) - \hbar(\epsilon+n\omega)\right] u_n(\vec{r}) = \frac{e}{M}(-i\hbar\partial_{\vec{r}})\cdot\sum_m \vec{A}_{n-m}\, u_m(\vec{r}) \tag{4.3.19}$$

and then perform the Bloch analysis, i.e. pick the solution in the form

$$u_n(\vec{r}) = e^{i\vec{k}\cdot\vec{r}}\, v_n(\vec{r}) \quad , \tag{4.3.20}$$

where

$$v_n(\vec{r}+\vec{R}) = v_n(\vec{r}) \quad , \tag{4.3.21}$$

and then Fourier analyze in $\vec{r}$,

$$u_n(\vec{r}) = \sum_{\vec{K}} e^{i\vec{K}\cdot\vec{r}}\, u_{n\vec{K}} \quad , \tag{4.3.22}$$

$$V(\vec{r}) = \sum_{\vec{K}} e^{i\vec{K}\cdot\vec{r}}\, V_{\vec{K}} \quad , \tag{4.3.23}$$

where $\vec{K}$ is any reciprocal lattice vector. The result is a doubly infinite system

$$\left[\frac{\hbar^2}{2M}(\vec{k}+\vec{K})^2 - \hbar(\epsilon+n\omega)\right] u_{n\vec{K}} + \sum_{\vec{K}'} V_{\vec{K}-\vec{K}'}\, u_{n\vec{K}'} =$$

$$= \frac{e\hbar}{M}(\vec{k}+\vec{K})\cdot\sum_m \vec{A}_{n-m}\, u_{m\vec{K}} \tag{4.3.24}$$

and zeros of its Hill determinant

$$\Delta(\vec{k},E) \equiv$$

$$\det\left\{\delta_{\vec{K}\vec{K}'}\,\delta_{nn'} + \frac{1}{\frac{\hbar^2}{2M}(\vec{k}+\vec{K})^2 - \hbar(\epsilon+n\omega)}\left[V_{\vec{K}-\vec{K}'}\,\delta_{nn'} - \frac{e\hbar}{M}(\vec{k}+\vec{K})\cdot\vec{A}_{n-m}\,\delta_{\vec{K}\vec{K}'}\right]\right\} \tag{4.3.25}$$

determine the "eigenvalues" $\epsilon_n(\vec{k},\omega)$. Although this form of is not convenient for calculations, some general properties of the spectrum $\epsilon_n(\vec{k},\omega)$ follow from it[1].

Firstly, as in the conventional solid-state case, the Hill determinant is invariant under translations of $\vec{k}$ by a reciprocal lattice vector

$$\Delta(\vec{k}+\vec{K},\epsilon) = \Delta(\vec{k},\epsilon) \tag{4.3.26}$$

Hence, it follows that one need only study the first Brillouin zone. Secondly, the Hill determinant is invariant under translations of ϵ by any number of ω:

$$\Delta(\vec{k},\epsilon+m\omega) = \Delta(\vec{k},\epsilon) \tag{4.3.27}$$

Hence, one need only study the interval $[0,\omega]$ in ϵ. All the band structure associated with this problem is contained in a fundamental hyper-rectangle in the $(\epsilon,\vec{k})$ plane.

The other way to procede is to project the system (14) on the (ideal solid) Bloch states $\langle\vec{r}|n'\vec{k}\rangle = e^{i\vec{k}\cdot\vec{r}}\,\omega_{n'}(\vec{r})$,

$$\langle n'\vec{k}\left|\left[H_0(\vec{p},\vec{r})-\hbar(\epsilon+n\omega)\right]\right|u_n\rangle = \frac{e}{M}\langle n'\vec{k}\,|\,\vec{p}\cdot\sum_m \vec{A}_{n-m}\,u_m\rangle \quad . \tag{4.3.28}$$

However,

$$H_0|n'\vec{k}\rangle = E_{n'}(\vec{k})|n'\vec{k}\rangle \tag{4.3.29}$$

where $E_{n'}(\vec{k})$ is the band energy of the unperturbed solid. Also

$$\langle n'\vec{k}'|\vec{p}|n''\vec{k}''\rangle = \int\langle n'\vec{k}'|\vec{r}\rangle\, d^3r\,(-i\hbar\partial_{\vec{r}})\langle\vec{r}|n''\vec{k}''\rangle =$$

$$= \delta_{\vec{k}'\vec{k}''}\left[\hbar\vec{k}'\delta_{n'n''} - i\hbar\int d^3r\, \omega^*_{n'}(\vec{r})\,\partial_{\vec{r}}\,\omega_{n''}(\vec{r})\right] , \qquad (4.3.30)$$

and the only way in which energy bands couple is via the momentum matrix elements

$$\left[E_{n'}(\vec{k}) - \hbar(\epsilon + n\omega)\right]\langle n'\vec{k}|u_n\rangle = \frac{e\hbar}{M}\sum_m \vec{k}\cdot\vec{A}_{n-m}\langle n'\vec{k}|u_m\rangle -$$

$$\frac{ie\hbar}{M}\sum_m\sum_{n''}\int d^3r\, \omega^*_{n'}(\vec{r})\,\partial_{\vec{r}}\,\omega_{n''}(\vec{r})\cdot\vec{A}_{n-m}\langle n''\vec{k}|u_m\rangle \quad . \qquad (4.3.31)$$

This place is usually a branching point for any research. From now on usually more specific aspects of the problem are addressed and various approximations introduced - like two-band (or one-band) models, a monochromatic laser, various approximations for Bloch states, etc.[4] A common reason is self-evident - intractability of the general problem. This area of solid state physics is still very much alive and in a sense controversial. We hope to be able to set aside more time in the future and continue research in this interesting and important field.

We will finish this chapter with an illustrative example : a linear chain subjected to a monochromatic laser, in the tight-binding approximation. An analysis based on the system (31) then leads to a coupled set of equations

$$C_{mj}\left[\epsilon - \lambda - m\omega\right] = T\left[C_{mj+1} + C_{mj-1}\right] +$$

$$+ U\left[\quad {}_{j+1} + C_{m-1\,j+1}\right] + U\left[C_{m+1\,j-1} + C_{m-1\,j-1}\right] \qquad (4.3.32)$$

where λ is a site energy, T is a hopping integral and U is the strength of the external field. The form of the interaction with the external field is special in that absorption or emission of a "photon" causes hopping from one site to a neighboring site. Eqs. (32) may be solved by letting

$$C_{mj} = \gamma_m e^{ijb} \quad , \qquad (4.3.33)$$

so

$$\gamma_m\left[\epsilon - \lambda - m\omega\right] = 2T\gamma_m \cos b + U\left[\gamma_{m+1}e^{ib} + \gamma_{m-1}e^{ib} + \gamma_{m+1}e^{-ib} + \gamma_{m-1}e^{-ib}\right] \qquad (4.3.34)$$

In the limit where $U \to 0$ (the external interaction is turned off) we have

$$\epsilon = \lambda + m\omega + 2T\cos b \qquad (4.3.35)$$

which, aside from the $m\omega$ term, is the usual tight-binding result. The $m\omega$ term just accounts for the fact that ϵ is the quasi-energy and not the actual energy. For finite U let

$$\beta \equiv \frac{2U\cos b}{\omega} \quad , \qquad (4.3.36)$$

$$\alpha \equiv \frac{\epsilon - \lambda - 2T\cos b}{\omega} \quad , \tag{4.3.37}$$

so

$$(\alpha - m)\gamma_m = \beta(\gamma_{m+1} + \gamma_{m-1}) \quad . \tag{4.3.38}$$

A solution to this difference equation is

$$\gamma_m = J_{\alpha - m}(2\beta) \quad , \tag{4.3.39}$$

as may readily be verified by the recurrence relation for the Bessel functions:

$$J_{n-1}(z) + J_{n+1}(z) = \frac{2n}{z} J_n(z) \tag{4.3.40}$$

(another possible solution involves Bessel functions of the second kind Y_n). For $U \to 0$, $\beta \to 0$ and we only have a non-vanishing γ when $\alpha = m$. This just reduces to the previously obtained expression for the spectrum. For $U \neq 0$ it would appear that any ϵ is an acceptable solution (for a given value of b). However near $b = \pm\frac{\pi}{2}$, β gets to be small and we must ask if γ_m remains finite at these points. Any physically acceptable solution should remain finite. For small values of the argument

$$J_\nu(z) = \left(\frac{z}{2}\right)^\nu \frac{1}{\Gamma(\nu+1)} \tag{4.3.41}$$

where $\nu \neq -1,-2,-3,\ldots$ Hence if ν is not an integer we will have divergent behaviour for negative ν . Thus $\alpha \geq m$ or

$$\epsilon - \lambda - 2T\cos b \geq m\omega \tag{4.3.42}$$

Inspection of the above asymptotic form shows that J_ν has a cusp behaviour at the zone edge - again an unphysical behaviour. To avoid this we must have ν an integer, so

$$\epsilon - \lambda - 2T\cos b = m\omega \tag{4.3.43}$$

is the dispersion relation. A typical band structure is sketched in Fig. 3. Note that, aside from the dispersion relation being taken modulus itself, no distortion of the band structure occurs. This is characteristic of a one-band model.

Figure captions

Figure 1. Sketch of hypothetical band structures in a laser field.

Figure 2. Band structure for the modified Kronig - Penney model. Circles represent some of the numerical values for not too strong external field.

Figure 3. A typical band structure for one-band model in the tight-binding approximation. We have taken $\lambda = -2T$ and $T < 0$ for simplicity.

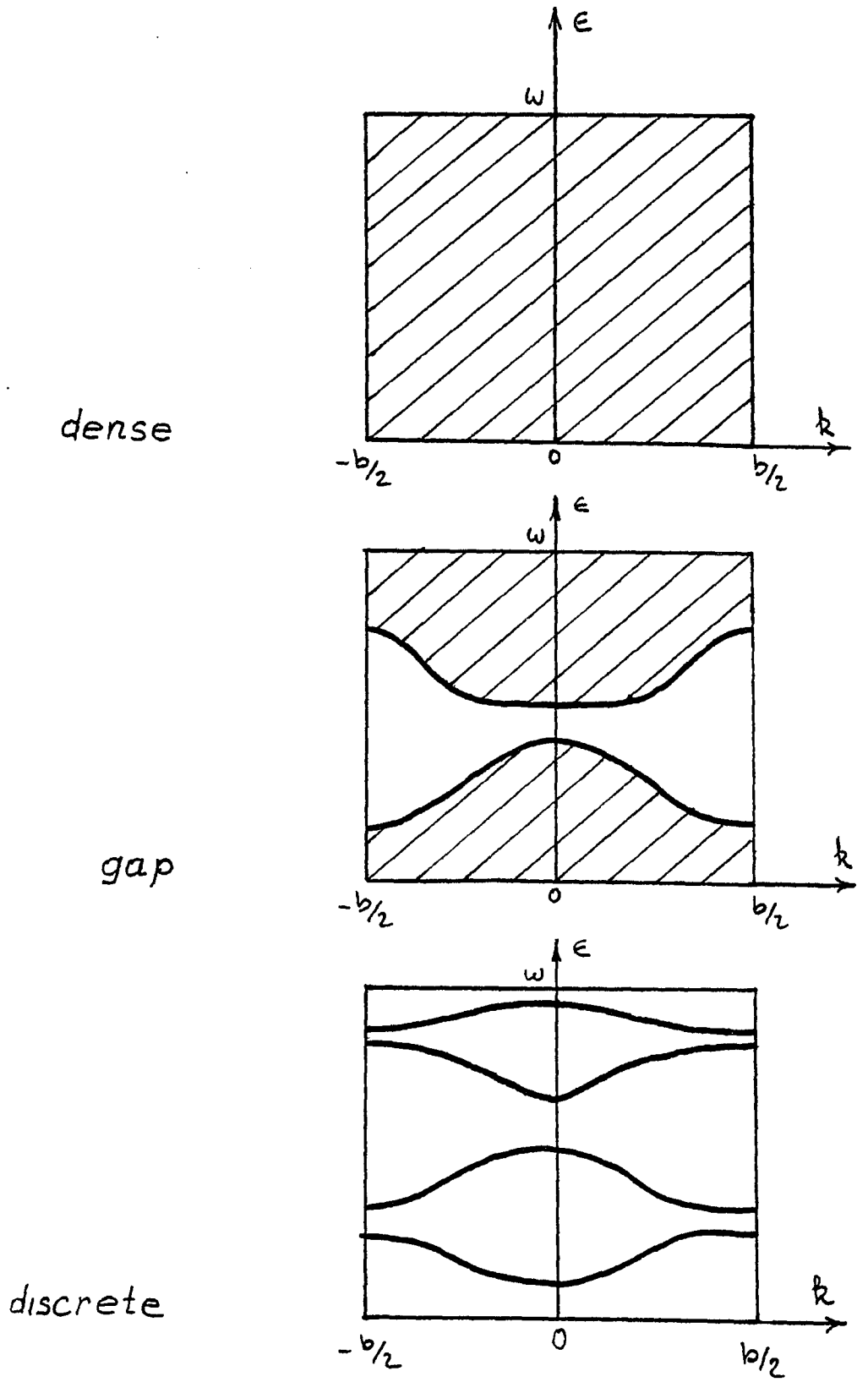


Figure 1

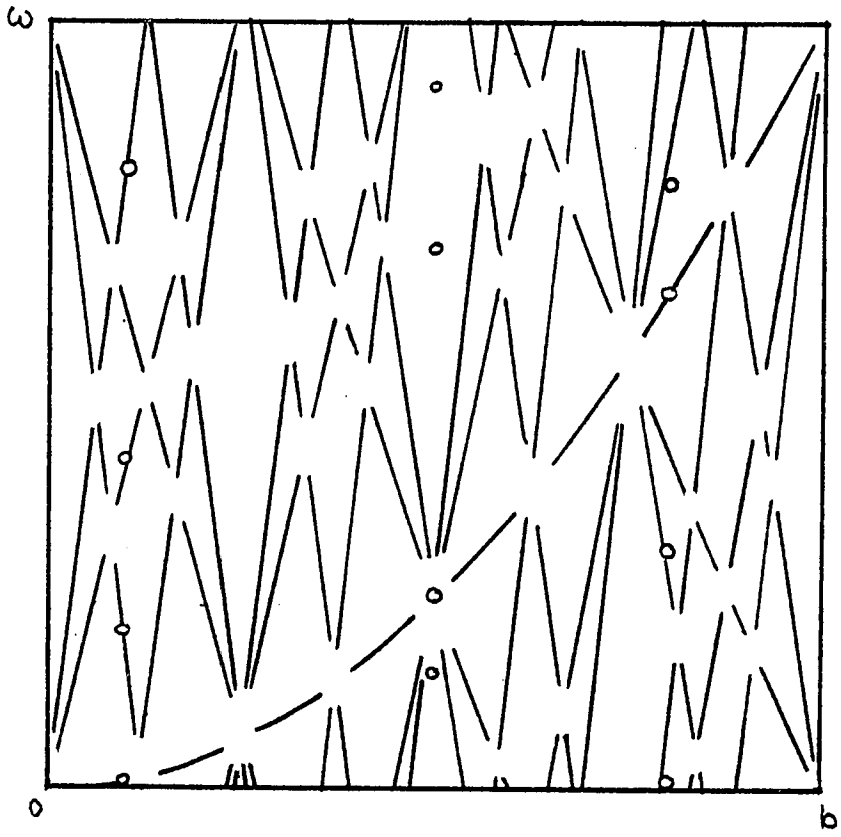


Figure 2

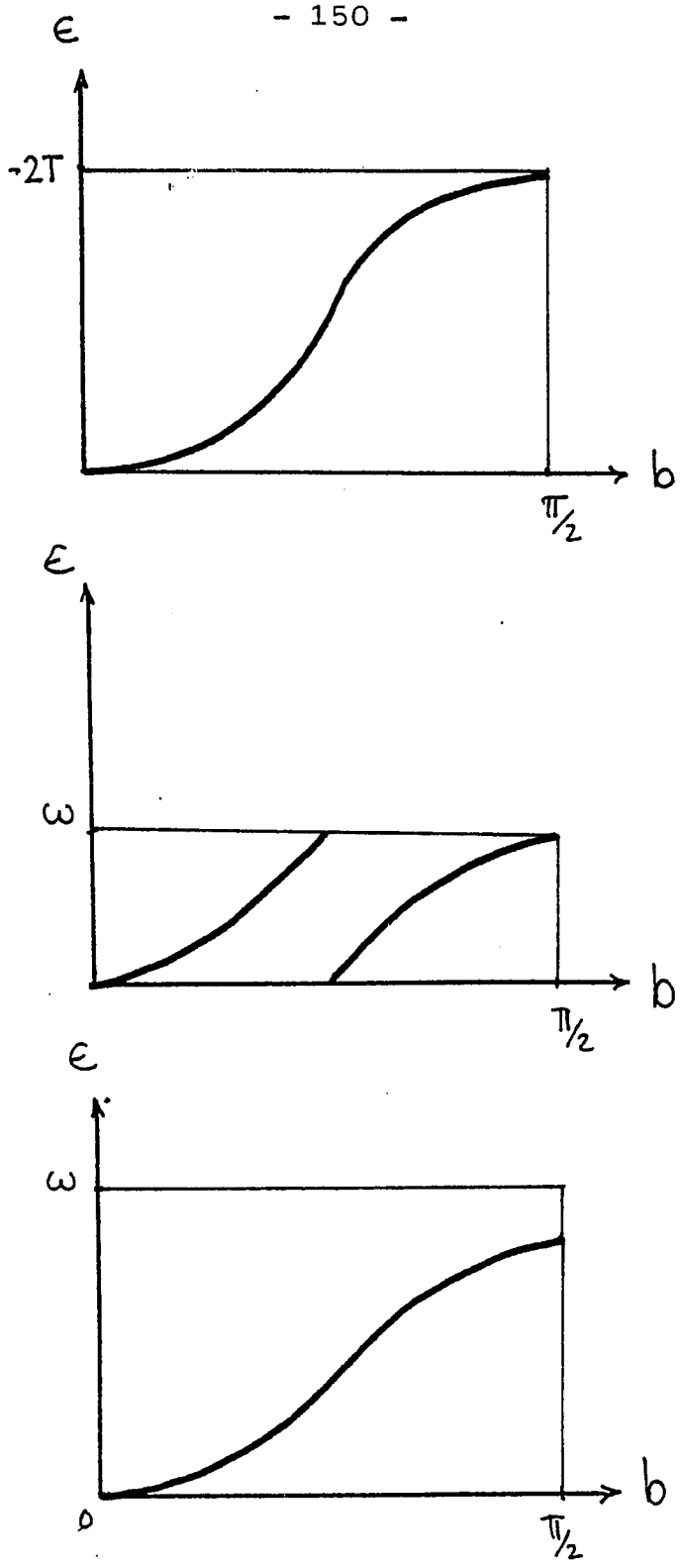


Figure 3

References

1. N. Tzoar and J. I. Gersten, Phys. Rev. B, 12, 1132 (1975)

2. J. H. Shirley, Phys. Rev. 138, B979 (1965)

3. See, for example J. Patterson, "Introduction to the Theory of Solid State Physics", (Adison-Wesley, Reading Mass.,1971)

4. A good review of the field with many references is provided in Callaway's "Quantum Theory of Solid State" (Academic Press New York, 1976)

5. W. V. Houston, Phys. Rev. 57, 184 (1940); E. I. Blount, Solid State Phys. 13 (1961); G. H. Wannier, Rev. Mod. Phys. 34, 645 (1962); J. Zak, Solid State Phys. 27 (1973)