Pump-probe Spectroscopy

Nguyen A. Tuan

Department of Physics
University of Cincinnati
Cincinnati, Ohio 45221

Abstract

Ultrafast spectroscopy of semiconductors is currently one of the most exciting areas of research in condensed-matter physics. Remarkable progress in the generation of tunable femtosecond pulses has allowed direct investigation of the most fundamental dynamical processes in semiconductors. This paper presents a brief overview of the pump-probe technique, the most common form of ultrafast spectroscopy, and several its variations which are widely used in practice.

INTRODUCTION

Pump-probe spectroscopy is a powerful technique for investigating electronic and vibrational properties of variety of systems, and has brought out extensive information into the properties of atoms, molecules and solids. Its strengths have been exploited for many years and recently have been strengthen when combined with the picosecond and femtosecond laser pulses to provide new insights into completely different aspects of semiconductors.

GENERAL PRINCIPLES

As the most common form of ultrafast optical spectroscopy, the pump-probe method needs at least two ultrashort light pulses. The first one, the so-called "pump" pulse, perturbs the sample at time $t = 0$. The second one, the so-called "probe" pulse, delayed with respect to the "pump" pulse, crosses the perturbed sample and acts as a probe [1, 2]. In the simplest form of pump-probe spectroscopy (so-called *degenerate* pump-probe spectroscopy), the output pulse train from an ultrafast laser is divided into two. The semiconductor sample under investigation is excited by the pump pulse and the changes induced in the sample are probed by the probe pulse which is suitably delayed with respect to the pump by introducing an optical delay in its path. Some property related to the probe (e.g. reflectivity, absorption, Raman scattering, luminescence) is then monitored to investigate the changes in the sample produced by the pump. In *non-degenerate* pump-probe spectroscopy, one uses two synchronized lasers at different wavelengths, or a laser and a synchronized white light continuum. This increases the versatility of the technique enormously by allowing a determination of the changes induced at photon energy different from that of the pump. Figure 1 represents a typical schematic diagram of a pump-probe experiment.

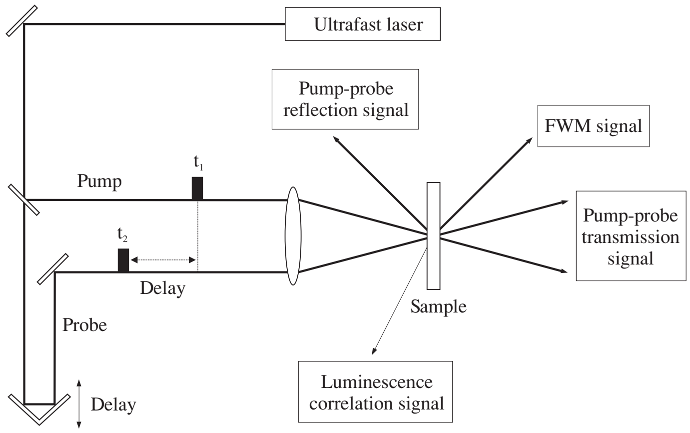

Figure 1: *Schematic for a generic setup for a pump-probe experiment. This setup can be used for pump-probe transmission or reflection spectroscopy, FWM spectroscopy, or luminescence correlation spectroscopy as shown [1].*

VARIATIONS OF PUMP-PROBE SPECTROSCOPY

Many different pump-probe spectroscopy techniques have been developed based on different techniques of linear optical spectroscopy. The following sections will briefly discuss them.

1. Transmission/Reflection Spectroscopy.

Perhaps the simplest of pump-probe methods is measurement of transmitted or reflected probe. The probe is typically much weaker than the pump, and the spot diameter of the probe at the sample is considerably smaller than that of the pump so that it probes a region of uniform photoexcited density. For the transmission spectroscopy (similarly to reflection spectroscopy), one generally measures the *change* in the transmitted probe pulse energy induced by the pump as a function of the time delay between the two pulses by chopping the pump beam and using lock-in amplifier. Some experiments may require that the pump and probe beams are chopped at different frequencies, and the change in transmitted probe is detected at, for instance, the sum of the two chopping frequencies.
The data are typically presented in the form of normalized Differential Transmission (DT) given by $\Delta T/T_0 = (T-T_0)/T_0$, the change in transmission $\Delta T = T-T_0$ induced by the pump pulse divided by the transmission of the probe in the absent of the pump, T_0. The Differential Transmission Spectrum (DTS) can be obtained either by varying the wavelength of the probe laser or by spectrally analyzing the transmitted probe energy with a spectrometer if a broadband ultrafast source (e.g. white light continuum) is used as the probe. The use of a multichannel as diode array or CCD camera is generally preferable in the latter case. Although one usually measures the total transmitted probe energy, the transmitted pulse shape can be measured by a suitable fast detector or a nonlinear gating technique [1].

2. Raman Spectroscopy.
In this form of pump-probe spectroscopy, the changes induced by the pump are measured by the change in the Raman scattering signal generated by the probe. In degenerate pump-probe Raman spectroscopy, the two beam are derived from the same laser and are generally cross-polarized to distinguish the Raman signals generated by the pump and the probe. If the two lasers have different wavelength, the signals can be distinguished spectrally. The time-integrated Raman signal of the probe, or the total energy of the Raman-scattered probe pulse, is measured as a function of time delay between the pump and the probe. Both vibrational and electronic Raman signals have been used in this form of time-resolved Raman spectroscopy to monitor the dynamic of photoexcited phonons and electronic excitations.

3. Luminescence Spectroscopy
Luminescence dynamics can be measured directly with fast detector and fast electronics. This includes the use of fast photodiodes which can provide time resolution in the picosecond for intense pulses, but are not well-suited for low-level signal detection. Therefore the correlation and upconversion spectroscopy are used alternatively.

(a) Correlation Spectroscopy. In this really form of pump-probe technique, two pulse trains delayed by τ_d excite the semiconductor and the time-integrated luminescence spectra are measured as a function of τ_d. If the luminescence signal is a nonlinear function of the excitation density, then the signal with two excitation pulses is not simply the sum of the signal with each beam exciting the sample separately. In practice, one chops the two beams at two different frequencies, and the nonlinear correlation signal is measured at the sum frequency. The advantages of the technique are its simplicity, and high time resolution limited only by the laser pulse width. The disadvantage is the complexity in the interpretation of the results.

(b) Upconversion Spectroscopy. This technique provides a means for measuring low-level luminescence signal with time resolution limited by laser pulse width and has been aplied to a wide variety of investigations in semiconductors. The ultrafast laser is divided into two beams. The first part excites the sample. The induced sample emission is collected in the usual manner and focused on a nonlinear crystal. The second laser beam is suitably delayed and also focused on the nonlinear crystal to overlap the luminescence spot. When the laser beam and the emission signal are present simultaneously in the nonlinear crystal, frequency mixing occurs, which results in the generation of an up-converted signal. For a given detection axis, the up-converted frequency is determined both by the angle between two incoming beams and the angle of the nonlinear crystal, and by the optical frequencies of these beams. The nonlinear crystal thus acts as an optical gate which is opened only when the laser pulse is present in it. By varying the optical path of the laser pulse which opens the gate, the photoluminescence signal is sampled at different time delays. Figure 2 illustrates a schematic diagram of luminescence upconversion spectroscopy.

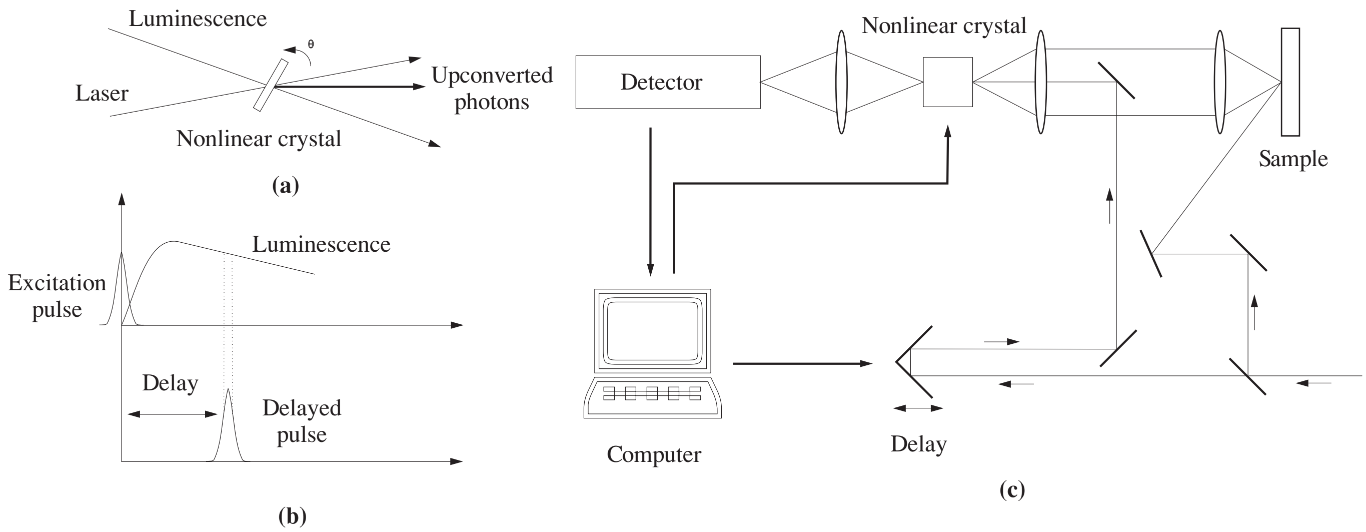

Figure 2: *Upconversion spectroscopy: (a) the principle, (b) the temporal and spectral sampling of the luminescence signal, and (c) an experimental setup [2].*

FOUR WAVE MIXING (FWM) SPECTROSCOPY

1. Principle of the method.

FWM experiments are generally used to measured the nonlinear response of a semiconductor in terms of third-order polarization induced by the optical fields. While the diffracted signal in FWM experiment measures only the coherent response of the semiconductor, the differential transmission or reflection signal is determined by the coherent as well as the incoherent dynamics.

The four wave mixing process can be interpreted as an interaction between four photons. A photon of frequency ω_1 combines with a photon of frequency ω_2 to produce a photon of frequency ω_3 and another of frequency ω_4. The conservation of energy and momentum are $\omega_1 + \omega_2 = \omega_3 + \omega_4$ and $\mathbf{k_1} + \mathbf{k_2} = \mathbf{k_3} + \mathbf{k_4}$ respectively. If all four waves are of the same frequency, $\omega_1 = \omega_2 = \omega_3 = \omega_4 = \omega$, the process is then called degenerate four wave mixing

[3].
In the simplest form of FWM spectroscopy, the two-beam Degenerate Four Wave Mixing (DFWM), the pump photons with wavevector $\mathbf{k_1}$ generate a coherent polarization in the sample, and the probe photons with wavevector $\mathbf{k_2}$ arrive at the sample after a time delay τ_d. If τ_d is smaller than the dephasing time of polarization, an interference grating is produced from which $\mathbf{k_2}$ can be self-diffracted along the phase matched direction $\mathbf{k_d} = \mathbf{2k_2} - \mathbf{k_1}$.
In a three-beam DFWM, the sample is excited by three beams traveling along $\mathbf{k_1}$, $\mathbf{k_2}$ and $\mathbf{k_3}$ arriving at the sample at times t_1, t_2 and t_3 respectively. τ_{21}, τ_{31} and τ_{32} $(\tau_{ij} - t_i - t_j)$ define the delays between these beams. The diffracted total energy along the phasematched direction $\mathbf{k_d} = \mathbf{k_3} + \mathbf{k_2} - \mathbf{k_1}$ is measured as a function of various time delays to obtain different information about the system being investigated. Figure 3 shows the schematic diagram for a three-beam FWM experiment.

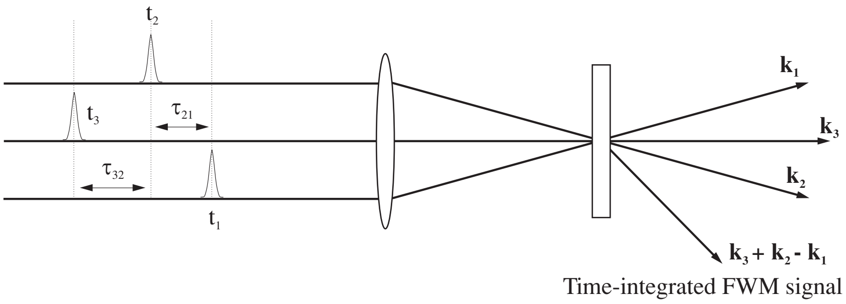

Figure 3: *Schematic diagram for a three-beam FWM experiment.*

2. FWM and third-order nonlinear effects.
The nonlinear effect is characterized by a nonlinear relation between polarization density of a dielectric medium and external electric field. The relation can be customary written in the form

$$P = \epsilon_0 \chi E + 2dE^2 + 4\chi^{(3)} E^3 + ...$$

where P and E are the polarization and electric field, χ, d and $\chi^{(3)}$ are coefficients describing the linear effect, the second- and third-order nonlinear effects, respectively [3].
We now examine the case of two-beam DFWM experiments where one measures the diffracted signal along $\mathbf{k_d}$. If the pulse along $\mathbf{k_2}$ is delayed by time delay τ_d with respect to the pulse along $\mathbf{k_1}$, then the electric field can be expressed as

$$\mathbf{E}(\mathbf{R}, t) = E_1 A(\mathbf{R}, t) exp[i(\mathbf{k_1}.\mathbf{R} - \omega t)] + E_2 A(\mathbf{R}, t - \tau_d) exp[i(\mathbf{k_2}.\mathbf{R} - \omega t)]$$

Here $E_1(E_1)$ is the amplitude of the first(second) electric field, $A(\mathbf{R}, t)$ is the electric field pulse shape at $\mathbf{R}$, ω corresponds to the peak of the mode-locked laser spectrum. Also we have ignored polarization effects on dipole matrix and the relative phase, absorbing 1/2 in the field amplitude and assumed that directions of linear polarization of the first and second fields are equal [1].
For a two-level system with the ground state with state vector $|a\rangle$ and energy E_a and an excited state with the state vector $|b\rangle$ and energy E_b, by solving the optical Bloch equations

[1, 2], we can obtain the third-order polarization

$$
\begin{aligned}
p_{221}^{(3)} = & - & 2i(e|r_{ba}|/\hbar)^3 exp(-gt)E_1E_2^2 \int_{-\infty}^{t} dt''' \int_{-\infty}^{t'''} dt'' \int_{-\infty}^{t''} dt' \\
& \times & \{exp(Gt''')exp[-g_r(t'''-t'')]A(\mathbf{R},t'''-\tau_d)\} \\
& \times & \{A(\mathbf{R},t''-\tau_d)A(\mathbf{R},t')exp[-G*(t''-t')] \\
& & +A(\mathbf{R},t'-\tau_d)A(\mathbf{R},t'')exp[-G(t''-t')]\}
\end{aligned}
$$

where $\mathbf{r}$ is the electron coordinate with respect to the location of the nucleus at $\mathbf{R}$, $G = g - i\omega$ with $g = (i\Omega + 1/T_2)$, $\Omega = (E_b - E_a)/\hbar$, $g_r = 1/T_1$, T_1 is the lifetime of the state b and $1/T_2$, the transverse relaxation rate, is the sum of the recombination rate $(1/T_1)$ and pure dephasing rate (one can think of T_2 as the lifetime of the coherent superposition state) [1]. The diffracted signal along $\mathbf{k_d}$ can be approximated by the macroscopic polarization $P_{221}^{(3)}$. In practice, the temporal evolution of the FWM (TR-FWM) signal in the $\mathbf{k_d}$ direction

$$S_{221}^{(3)}(t) = |P_{221}^{(3)}|^2$$

and the time-integrated (TI-FWM) signal as a function of time delay τ_d

$$I_{221}^{(3)}(\tau_d) = \int_{-\infty}^{\infty} S_{221}^{(3)}(t)dt$$

3. Example of Application: Quantum Beats of Excitons

Consider the case of a three-level system which has a common ground state and two excited states that are close in energy (Figure 4.a). If the system is excited resonantly with a short laser pulse such that the spectral width of the laser is greater than the energy separation between the two excited states $|1\rangle$ and $|2\rangle$, then the laser pulse excites a coherent superposition state, i.e. a linear superposition of the two wavefunctions. This is a non-stationary state whose time evolution is determined by the time-dependent Schrdinger equation. Since the energies of the eigenstates comprising the non-stationary state are different, the phase of the two eigenstates will evolve at a different rate. It can be shown that the spontaneous emission intensity from this system following the pulsed excitation oscillates with a period given by $h/\Delta E$ where ΔE is the energy separation between levels $|1\rangle$ and $|2\rangle$ and h is the Planck's constant. This is the phenomenon of quantum beats. The term "quantum beats" has also

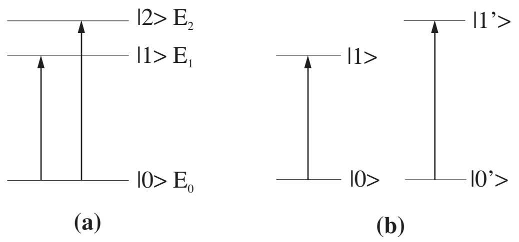

Figure 4: *(a) Schematic of a three-level system with two closely-spaced excited states and a common state; (b) two two-level systems with close transition energies but no common state [1].*

been used to describe interference phenomena which arise when the two transitions not sharing a common ground or excited states are excited by a short laser pulse whose spectrum is wider than the difference in transition energies (Figure 4.b). Although such interference has been considered as a special case of "quantum beats", there is no wavefunction superposition or quantum interference. The following part will present results obtained by several authors in studying heavy hole-light hole (HH-LH) beats in semiconductor heterostructures.

It is well known that removal of the translational symmetry in the growth direction in quantum wells lifts the degeneracy between the HH and the LH valence bands at $k = 0$, and leads to well-separated HH and LH excitons in quantum wells. This is an intrinsic three-level system with a common ground state. Figure 5 shows TI-FWM signals from a 10 period 170/150 $\AA$ $GaAs/Al_{0.36}Ga_{0.64}As$ MQW sample at 5 K and a comparison of the experiment data with a theoretical calculation. The laser is ≈ 500 fs pulse width and 4.2 meV spectral width. It can be seen clearly the strong periodic modulation of the TI-FWM signal when the laser is tuned to an energy between the HH and the LH excitons. The period of the oscillations is 960 ± 40 fs, in good agreement with the period of 980 fs calculated from the 4.2 meV HH-LH exciton splitting observed in the PLE spectrum [1].

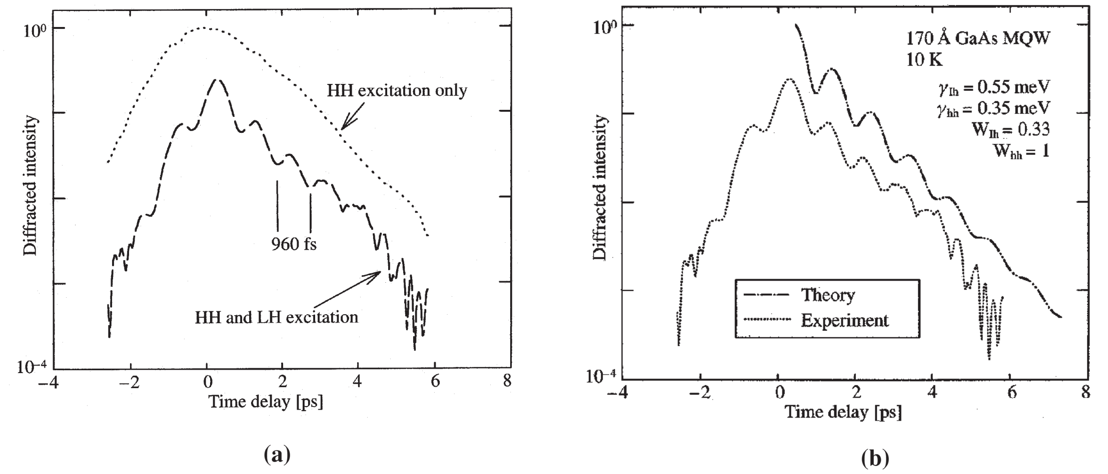

Figure 5: *(a) TI-FWM signal from a 10 period 170/150* $\AA$ $GaAs/Al_{0.36}Ga_{0.64}As$ *MQW sample at 5 K and 4.2 meV; dash-dot curve: 1.5265 eV excitation (exciting primary the HH exciton) showing smooth decay; dash curve: 1.5310 eV, exciting both the HH and the LH excitons showing the quantum beats; (b) comparison with calculations based on a simple three-level model (dash-dot line) [1].*

References

[1] J. Shah, *Ultrafast Spectroscopy of Semiconductors and Semiconductor Nanostructures* , Springer, Berlin, 1999.

[2] C. Rullire, *Femtosecond Laser Pulses: Principles and Experiments* , Springer, Berlin, 1998.

[3] B. E. A. Saleh, M. C. Teich, *Fundamentals of Photonics* , John Wiley, New York, 1991.