Optical properties of Er-doped Si-based media

Minissale, S.

Citation for published version (APA):
Experimental

1.1 Sample details and preparation

The following section will give an overview of the materials on which experiments have been carried out in my doctoral research.

Both structural description and preparation methods will be addressed.

1.1.1 Si/Si:Er multinanolayers and reference samples

A great part of the work described in this thesis has been performed in Si/Si:Er multinanolayers. They have been grown by sublimation molecular beam epitaxy (SMBE) on Si(100) $p$-type substrate ($\rho \approx 10 - 20$ $\Omega$cm) under a pressure of $2 \times 10^{-7}$ mbar [22]. These structures comprise interchanged Si and Si:Er layers stacked along the $\langle 100 \rangle$ growth direction (Fig. 1.1).

Many samples, with different thickness of both doped and undoped Si layers and different number of repetitions of the structure (Si/Si:Er) have been produced and two of them have been used in the measurements reported in this thesis. The first of the two, called NLA (S51 in Ref. [12]), is composed of 400 periods with a thickness of the doped ($d_{act}$) and undoped ($d_{Si}$) layer of, respectively, 2.7 and 1.7 nm. Following the growth procedure, the structure was annealed at 800 $^\circ$C for 30 minutes.

The second one, named NLB (or S56, Ref. [12]) comprised 19 periods with thickness of 5 and 100 nm for the doped and undoped layers, respectively. The Er concentration in the doped regions was determined by SIMS as $3.5 \times 10^{18}$ cm$^{-3}$ in both samples. The total Er areal density was $2 \times 10^{14}$ and $2 \times 10^{13}$ cm$^{-2}$ for structures NLA and NLB, respectively.

For comparison purposes, two references have been used:

1. A SiO$_2$:Er “standard” sample (labelled STD), prepared by a triple Er-implantation ($1.5 \times 10^{14}$ cm$^{-2}$ at 200 keV; $2.8 \times 10^{14}$ cm$^{-2}$ at 500 keV and $5.6 \times 10^{14}$ cm$^{-2}$ at 1000 keV) to the total dose of $9.9 \times 10^{14}$
1. Experimental

1.1 Si/Si:Er multilayer structure

Figure 1.1: The Si/Si:Er multilayer structure.

- cm$^{-2}$, and subsequent 30 min. annealed in nitrogen at 1000 °C.

2. An implanted Si:Er sample (labelled IMP, Er dose of $3 \times 10^{12}$ cm$^{-2}$).

- It has been prepared by Er implantation ($3 \times 10^{12}$ cm$^{-2}$, 300 keV) and oxygen ($3 \times 10^{13}$ cm$^{-2}$, 40 keV) co-implantation followed by 30 min 900 °C annealing with these preparation conditions being optimized for the maximum emission intensity [23].

1.1.2 Si/Si$_{1-x}$Ge$_x$:Er nanolayers

Also another type of nanolayers has been investigated. In this case the active medium, an Er-doped Si$_{1-x}$Ge$_x$ layer, is placed in between two layers of undoped Si [24]. Si/Si$_{1-x}$Ge$_x$:Er/Si (only one period of repetition) structure is grown on a Si substrate. The Er concentration, for all the samples, is of the order of $2 \times 10^{18}$ cm$^{-3}$. They have been fabricated by SBE as for the Si/Si:Er nanolayers at growth temperature of 500 °C, while Ge doping occurred through the pyrolysis reaction of GeH$_4$ gas close to the heated substrate.

- Three samples with different thickness $d_{act}$ of the active layer and different
Ge content $x_{Ge}$ have been produced, as summarized in the following table:

<table>
<thead>
<tr>
<th>Sample</th>
<th>$d_{SiGe:Er}$ (nm)</th>
<th>$x_{Ge}$ (%)</th>
<th>RES (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>168III</td>
<td>2300</td>
<td>27.1</td>
<td>3</td>
</tr>
<tr>
<td>185II</td>
<td>150</td>
<td>24</td>
<td>46</td>
</tr>
<tr>
<td>142II</td>
<td>150</td>
<td>12</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 1.1: $Si/Si_{1-x}Ge_x:Er$ nanolayers structure paremeters: thickness of the active layer $d_{SiGe:Er}$, Ge content $x_{Ge}$ and residual elastic stress (RES).

Residual elastic stress (RES) is an important parameter because it gives an indication of the amount of defects in the structures. The difference in lattice parameter between Si and Ge causes biaxial compressive strain within the $Si_{1-x}Ge_x$ layer that is linearly proportional to the Ge content. The biaxial compression of the SiGe layer on Si results in an elastic energy proportional to the square of the strain. This elastic energy increases linearly with the thickness of the layer and drives to relaxation of the strain, through
formation and gliding of misfits dislocations, and undulation and roughening of the surface. Therefore, samples with a low percentage of residual stress have been relaxed and produced a concentration of defects inversely proportional to the RES factor. In particular, XRD analysis was employed to determine the Ge content $x_{Ge}$ and the value of the RES factor.

1.1.3 **Er-doped Si-rich SiO$_2$ waveguides**

A series of differently prepared waveguides based on SiO$_2$:(Er+Si-nc) material, characterized by various levels of Si-nc and Er doping, and annealing temperature, which implies different conditions of Si-nc aggregation and therefore different size and structure of Si-nc’s have been produced. Slab waveguides were fabricated by multiple-energy co-implantation of Si and Er into a 10 $\mu$m thick thermal oxide layer grown on a (100) silicon substrate [26,27].

Sets of samples with three different silicon excess levels of 10%, 20% and 30%, and two different Er concentrations of $10^{18}$ cm$^{-3}$ and $10^{20}$ cm$^{-3}$ were prepared. Subsequently, each set of samples was annealed at 600 $^\circ$C, 900 $^\circ$C, or 1150 $^\circ$C for 1 hour in N$_2$ gas. The annealing was applied in order to form Si-nc’s and to optically activate Er$^{3+}$ dopants. From numerical calculations we have estimated that the 10% excess Si corresponds to a Si-nc concentration similar to the lower Er concentration, approximately $4 \times 10^{18}$ cm$^{-3}$, while the 30% excess Si results in Si-nc’s concentration somewhat lower than the higher Er concentration, approximately $10^{19}$ cm$^{-3}$.

Samples have been labelled as ErXSiY, where X=1,2 for lower and higher Er concentration, respectively, and Y=1,2,3 specifies the Si excess of 10%, 20%, and 30%, respectively. For reference, SiO$_2$ samples implanted only with erbium, with no Si excess, were prepared under identical conditions.

1.1.4 **Er-doped amorphous Si:H**

Chapter 6 is dedicated to the characterization of Er-doped hydrogenated amorphous Si samples. The samples present interesting and unexpected properties that will be discussed in detail.

The films of a-Si:H (600 - 700 nm thick) were obtained by dc-magnetron sputtering of metal erbium in an atmosphere of silane, argon and oxygen onto a quartz substrate. The substrate temperature during the sputtering process was maintained at 265 $^\circ$C. The growth rate was 111 Angstrom/h. After the completion of the sputtering, the films were slowly cooled in the chamber. The concentration of hydrogen and oxygen in the films was, according to the IR spectroscopy data, 8 - 10 at% and 6 - 8 at%, respectively [28]. The Er concentration was of the order of $10^{20}$ cm$^{-3}$. 
1.2 Experimental Techniques

Numerous techniques have been employed in the measurements that are the object of this thesis. All of them are optical techniques, namely, using optical excitation of the samples and detecting the consequent optical de-excitation as a function of different parameters.

1.2.1 Photoluminescence experiments

Photoluminescence is a widely employed technique that allows to characterize different materials. The idea is quite simple: a sample is excited with a laser at a certain wavelength $\lambda_{exc}$ and luminescence (from the sample) at different wavelengths $\lambda_{det}$ is detected. Luminescence can be collected, as mentioned before, as a function of different parameters, for instance, scanning continuously the detection wavelength or fixing the detected wavelength and varying the power of the excitation. Different lasers have been used to provide the optical excitation, continuous wave lasers: Millennia (active medium: Nd:YVO$_4$) at 532 nm and an argon laser tunable at different wavelengths (350, 488, 514, 521 nm). Pulsed excitation was provided by the 3$^{rd}$ harmonic of a Nd:YAG laser at 354 nm pumping an OPO (Optical Parametric Oscillator) with a BBO crystal (allowing continuous excitation from 430 - 650 nm and 780 - 2000 nm). The laser gives pulses of 5 ns duration time with a frequency of 10 Hz. OPO, because of its wide and continuous range of excitation wavelengths, permits to perform photoluminescence excitation (PLE) measurements [29]. In these kind of experiments the detection wavelength is fixed and the excitation wavelength is varied. Samples have been measured at different temperatures from 4 K up to room temperature. Two cryostats have been utilized: a flow cryostat, in which the sample is placed in a chamber with a helium gas atmosphere, and a cold finger one, that cools down the sample by means of contact heat exchange. Both cryostats can be nominally cooled down to 4 K, however the determination of the temperature in the cold finger is not really accurate since the contact between the sample and its holder is not always perfect and the heat exchange is not efficient. Two spectrometers were used: a high resolution one (Jobin-Yvon THR 1500), allowing a maximum spectral resolution of few Angstroms, and a second one (TRIAX 320) with a resolution of 2 nm. A photomultiplier tube (Hamamatsu instruments, 0.16 - 09 $\mu$m) with a temporal response of $\sim$30 $\mu$s and a GE photodiode (Edinburgh Instruments, EI-A, 0.9 - 1.7 $\mu$m) with a slower response time of 300 $\mu$s have been employed for detection respectively in the visible and in the infrared range. The PL signal was collected by lock-in technique with a good signal-to-noise ratio and eventually processed by a computer, equipped with a home made software. PL transients
were instead captured with a TDS 3032 Tektronix oscilloscope.

1.2.2 Time-correlated single photon counting

PL kinetics taken by directly feeding the signal from detector to the oscilloscope do not offer really good time resolution. In fact, the resolution is limited by the response time of the detector which is of the order of 30 $\mu$s. The time-correlated single photon counting (TC/SPC) technique allows much better time resolution (in our case, with a maximum resolution of 0.25 ns). TC/SCP consists of many components as briefly described in the paragraph (Fig. 1.3 illustrates schematically the photon counting setup).

![Figure 1.3: Schematic view of the TC/SPC experimental setup.](image)

A light source provides the excitation energy required for luminescence (In our experiment, the same pulsed excitation as in common PL measurements has been used). A “Start” signal is taken from the light source (detected by a diode or directly by the pulse generator of the laser) and sent to an instrument called Time to Amplitude Converter (TAC). After the sample is excited,
system relaxes back to the ground state emitting photons. When a photon hits
the photocathode of a Photomultiplier Tube (PMT), it generates an electrical
signal that adds up to the inherent noise of the PMT, sending the resulting
“signal plus noise” to the Stop Discriminator (SD). The latter is a crucial
component of the system, since it makes the system basically insensitive to
the noise. A discriminator differentiates between levels of electrical signal: if
the signal level is below a certain threshold is ignored, otherwise is recorded.
In fact, reality is more complicate than this. One of the strong assumption
behind the validity of TC/SPC is the fact that we detect and record single
photon events. It is common that multiple photon are detected and processed
as they were single photon events.

![Pulse sequences in the PMT.](image)

The occurrence of such event is detrimental for the reliability of the data,
since the intensity levels displayed in the decay curve are lower than the ones
actually emitted by photoluminescence. A way to eliminate such a problem
is to set an upper value to the discriminator, in order to ignore signal levels
higher than a ceratin threshold. Summarizing, events with a signal level lower
than a threshold, called LLD or Lower Level Threshold (this level is set to a
value higher than the noise one, and higher than a second upper threshold, ULD or Upper Level Threshold (with a value lower than the sum of the signal coming from two photons or from a photon plus noise), are recorded. Fig. 1.4 shows schematically how the discriminator works.

1.2.3 Free electron laser

Many of the measurements discussed have been performed with the free electron laser (FEL), at the Dutch Free Electron Laser facility for Infrared eXperiments (FELIX) in the FOM Institute for Plasma Physics “Rijnhuizen” [30]. A FEL is a light source that produces coherent electromagnetic radiation by the acceleration of electron bunches. In fact, electrons are accelerated to a relativistic speed and sent to an undulator, an array of magnets with alternating poles, forcing the electron beam to follow a sinusoidal path. The deflection of the electron beam from a straight path induces the emission of synchrotron radiation and the matching of the phase between the motion of the beam and the emitted light produces coherent radiation. The wavelength of the emitted radiation can be easily tuned by changing either the energy of the electron beam or the magnetic field strength of the undulator. The wavelength range of the Felix is 4.5 - 250 µm with a spectral width between 0.4 and 7 %. The beam consists of micropulses with a duration of 7 ps separated by intervals of zero intensity of either 1 or 40 ns. The micropulses form a macropulse with a duration up to 10 µs with frequency repetition of 5 Hz [30,31].

1.2.4 Two-color experiments

Two-color experiment consist of simultaneously excitation of the sample with two lasers at different wavelengths. In particular a primary band-to-band excitation is provided by a frequency doubled pulsed Nd:YAG laser (532 nm) with duration of 100 ps and repetition rate of 5 Hz. PL is collected by the usual configuration already described in Sec. 1.2.1.

The second beam from the free electron laser is spatially overlapped to the band-to-band excitation but fired with an adjustable delay time between the two beams. The wavelength of the second beam can be tuned at will in all its range.

In such configuration, we can monitor the influence on amplitude, spectra and kinetics of PL bands from the IR radiation as a function of its wavelength, power and temporal delay in respect to the Nd:YAG excitation.
1.2.5 Pump-Probe and Transient Grating spectroscopy

Two techniques that make use of the FEL excitation have been employed (Chapters 3 and 4) in this thesis. The first of them, pump-probe, is a well known technique to measure relaxation time constants of atomic or vibrational transitions.

A really intense beam (pump) excites the sample under investigation bringing the system from the ground state to the excited state. A second beam (a much less intense one, called probe), delayed in time respect to the first, probes the population of the ground state that is relaxing back from the excited state with its own characteristic time constant. The intensity of the transmitted probe beam is monitored as a function of the delay time between the two beams and compared to the intensity of a third beam probing the system much later with respect to both pump and probe. This third beam, called the reference, is always much more absorbed by the system since the effect of the pump has already gone away and the system is completely relaxed back to its ground state (Fig. 1.5 illustrates the process).

The recorded signal, for each delay time $t$ between pump and probe, gives a value corresponding to the following equation:

$$I_S(t) = \frac{I_{\text{probe}}(t) - I_{\text{reference}}(t)}{I_{\text{reference}}(t)}$$ \hspace{1cm} (1.1)

The schematic of the experimental pump-probe setup is shown in Fig. 1.6.

The FEL beam is divided by a beam splitter, with the two new created beams carrying 90% and 10% of the original intensity. The less intense beam is further divided in two beams of equal intensity, the probe and the reference, with a fixed mutual delay of 2 ns. The intense beam is directed to a moving stage in order to vary its delay in respect to the probe (and the reference). The stage can be moved continuously and provide a maximum time delay of 2 ns. After being reflected by different mirrors, the pump and probe beam (and reference too), are focused on the sample by a parabolic mirror. This point is really crucial. The spatial overlap of the beams has to be perfect for the experiment to work. In order to have a reliable overlap, collimation has to be tested first by focusing the beams into a pinhole (100 - 400 $\mu$m radius, in our case). Subsequently, the sample has to be placed in the same longitudinal position as the pinhole with respect to the direction of the beams. The orientation of the sample with respect to the incoming beams has to be carefully chosen. In fact, perpendicular orientation with respect to the axis connecting the two parabolic mirrors has to be avoided in order to prevent reflections of the beams back and forth between the mirrors. After the sample, the beams are collected by a second parabolic mirror, symmetrically
Figure 1.5: Pulse sequences in a pump-probe experiment: a) Pump beam brings the system to the excited state. b-e) Transmitted probe beam, coming with a time delay in respect to pump, is detected. Absorption of the probe beam depends on the relaxation of the system back to the ground state, with a characteristic time. f) 2 ns after the excitation of the pump the system has relaxed back to the ground state. A reference beam, identical to the probe beam, is detected after being partly absorbed by the system, giving a reference for the maximum absorption of the probe.

placed with respect to the first mirror (with the sample in the focal point of the two), and eventually probe and reference are detected by a GeGa detector, after being reflected by a mirror and one more parabolic mirror.

Transient grating spectroscopy is, to some extent, an analogous technique to the pump-probe, described before. Here, two pumps with the same intensities, are focused on the sample by a parabolic mirror, reaching it at the same time and with their mutual directions forming an angle $\theta$. The interference of the two pump beams creates a modulation of the electromagnetic field on the sample leading to the modulation of the concentration of excited species in the sample and, consequently, of the refractive index of the medium. In order to probe the interference pattern, a third beam is necessary. The probe beam,
reaching the sample with a variable delay time with respect to the pump beams, is diffracted by the interference pattern. The diffracted beam is then detected and recorded as a function of the delay time between the probe and the pump beams, monitoring in such away the change in population of the ground state with time. Although this techniques, as mentioned before, is quite akin to the pump probe, the lack of background makes it more sensitive. In addition, the interference pattern is created close to the surface, therefore the substrate plays a minor role in the measurement. A scheme of the setup is presented in Fig. 1.7.

The felix beam is separated in two by a first beam splitter, the transmitted one being the intense pump beam and the weak reflected one, the probe beam. The pump beam (red in the figure) is reflected by a mirror on a moving stage, as for pump-probe, in order to delay the pump and probe with respect to each other. The pump beam is then directed to a second beam splitter and divided in two new beams, called pump 1 and 2. Small adjustments have to be made in order to make the two pumps parallel to each other before being
reflected from the parabolic mirror and focused on the sample. They travel at the same height and with a distance of 5 cm to each other. The probe beam, instead, is directed to a polarizer, that make its polarization different from the pumps beam, and it travels parallel to one of the two pump but at different height with respect to the plane of the two pumps. The polarizer has a double function: the different polarization of the probe with respect to the pumps prevents the interference of the three beams with each other, provoking undesired effects, and makes easier the detection of the signal beam (with the same polarization as the probe). It is then reflected by the parabolic mirror and directed to the sample as for the other beams. On the other side of the sample, instead of the three incident beams, more beams emerge with the new ones being the diffracted signal.

The new beams coming out have directions given by the combination of the k vectors of the incoming beam, satisfying the following equation \[32,33\]:

\[
\vec{k}_D(t) = \vec{k}_1(t) + \vec{k}_2(t) \pm \vec{k}_3(t)
\]  

(1.2)
where $k_D$, $k_1$, $k_2$ and $k_3$ are the diffracted, the two pumps and the probe beams, respectively. $k_4$ in Fig. 1.8 represents one of the two diffracted beams $k_D$, given by the linear combination of the the three incoming beams in eq. 1.2.

All the beams are reflected by the second parabolic mirror with their image position specular in respect to the incident points on the first mirror. The signal goes through an analyzer (that let beams with the same polarization as the probe go through) and is eventually detected.

Fig. 1.8 displays the incident beams, the pattern created by them on the sample and the outcoming beams. $\Lambda$ represents the size of the fringe and it depends on the wavelength ($\lambda$) of the incident beams and on their angle ($\theta$) of incidence on the sample with the following relation [34,35]:

$$\Lambda = \frac{\lambda}{2 \sin(\theta/2)}.$$  \hspace{1cm} (1.3)