THE DOUBLE HETEROSTRUCTURE: CONCEPT AND ITS APPLICATIONS IN PHYSICS, ELECTRONICS AND TECHNOLOGY

Nobel Lecture, December 8, 2000

by

Zhores I. Alferov

A. F. Ioffe Physico-Technical Institute, Russian Academy of Sciences, 26 Politechnicheskaya, St. Petersburg 194021, Russian Federation.

1. INTRODUCTION

It is impossible to imagine now modern solid state physics without semiconductor heterostructures. Semiconductor heterostructures and, particularly, double heterostructures, including Quantum Wells, Wires and Dots are today the subject of research of 2/3 of the semiconductor physics community.

It can be said that if the possibility of controlling the type of conductivity of a semiconductor material by doping with various impurities and the idea of injecting nonequilibrium charge carriers were the seeds from which semiconductor electronics developed, heterostructures could make it possible to solve the considerably more general problem of controlling the fundamental parameters inside the semiconductor crystals and devices: band gaps, effective masses of the charge carriers and the mobilities, refractive indices, electrons energy spectrum etc.

Development of the physics and technology for the semiconductor heterostructures resulted in remarkable changes in our everyday life. Heterostructure electronics is widely used in many areas of human civilization. It is hardly possible to imagine our recent life without double heterostructure (DHS) laser-based telecommunication systems, without heterostructure-based light-emitting diodes, heterostructure bipolar transistors, without low-noise HEMT for high frequency applications including, for example, satellite television. The DHS laser enters now practically every house with CD-players. Heterostructure solar cells have been widely used for space and terrestrial applications.

Our interest in semiconductor heterostructures was not occasional. Systematic studies of semiconductors were started in the early 1930s at Physico-Technical Institute under the direct leadership of it's founder – Abraham Ioffe. V. P. Zhuze and B. V. Kurchatov studied the intrinsic and impurity conductivity of semiconductors in 1932 and the same year A. F. Ioffe and Ya. I. Frenkel created a theory of rectification in a metal-semiconductor contact based on the tunneling phenomenon [1]. In 1931 and 1936, Ya. I. Frenkel published his famous articles where he predicted, gave the name and developed the theory of excitons in semiconductors and E. F. Gross experimental-
ly discovered excitons in 1951 [2]. The first diffusion theory of p-n heterojunction rectification, which became the base for W. Shockley's p-n junction theory, was published by B. I. Davydov in 1939 [3]. By A. F. Ioffe's initiative in the late 1940s at the Physico-Technical Institute the research into intermetallic compounds was started. Theoretical prediction and experimental discovery of semiconductor properties of A\textsuperscript{5}B\textsuperscript{5} compounds were done independently by H. Welker and on the example of InSb by N. A. Gorunova and A. R. Regel at the Physico-Technical Institute [4]. We benefited a lot from the high theoretical, technological and experimental level in this area that had already existed at the Ioffe Institute at that time.

2. CLASSICAL HETEROSTRUCTURE

The idea of using heterojunctions in semiconductor electronics was put forward already at the very dawn of electronics. In the first patent concerned with p-n junction transistors W. Shockley [5] proposed a wide-gap emitter to obtain unidirectional injection. A. I. Gubanov at our Institute first theoretically analyzed volt-current characteristics of isotype and anisotype heterojunctions [6], but the important theoretical considerations at this early stage of heterostructure research have been done by H. Kroemer, who introduced the concept of quasi-electric and quasi-magnetic fields in a graded heterojunction and made an assumption that heterojunctions might exhibit extremely high injection efficiencies in comparison to homojunctions [7]. In the same period there were various suggestions about applying heterostructures in semiconductor solar cells.

The proposal of p-n junction semiconductor lasers [8], the experimental observation of effective radiative recombination in a GaAs p-n structure with a possible stimulated emission [9] and the creation of p-n junction lasers and LEDs [10] were the seeds from which semiconductor optoelectronics started to grow. However, lasers were not efficient because of high optical and electrical losses. The thresholds currents were very high and low temperature was necessary for lasing. The efficiency of LEDs was very low as well due to high internal losses.

The important step was made immediately after the creation of p-n junction lasers when the concept of double heterostructure laser was formulated independently by us and H. Kroemer [11]. In his article H. Kroemer proposed to use the double heterostructures for carriers confinement in the active region. He proposed that “laser action should be obtainable in many of the indirect gap semiconductors and improved in the direct gap ones, if it is possible to supply them with a pair of heterojunctions injectors”.

In our patent we also outlined the possibility to achieve high density of injected carriers and inverse population by “double” injection. We specially pointed out that homojunction lasers “do not provide CW at elevated temperatures” and as an additional advantage of DHS lasers we considered the possibility “to enlarge the emitting surface and to use new materials in various regions of the spectrum”.

Initially the theoretical progress was much faster than the experimental realization. In 1966, [12] we predicted that the density of injected carriers could by several orders of magnitude exceed the carrier density in the wide-gap emitter ("superjunction" effect). At the same year in the paper [13] submitted to a new Soviet Journal "Fizika i Tekhnika Poluprovodnikov" (Sov. Phys. Semiconductors) I summarized our understanding of the main advantages of the DHS for different devices, especially for lasers and high power rectifiers: "The recombination, light emitting, and population inversion zones coincide and are concentrated in the middle layer. Due to potential barriers at the boundaries of semiconductors having forbidden bands of different width, the through currents of electrons and holes are completely absent, even under strong forward voltages, and there is no recombination in the emitters (in contrast to p-i-n, p-n-n', n-p-p' homostructures, in which the recombination plays the dominant role) ... Because of a considerable difference between the permittivities, the light is completely concentrated in the middle layer, which acts as a high-grade waveguide, and thus there are no light losses in the passive regions (emitters)".

Here are the most important peculiarities of semiconductor heterostructures we underlined at that time: (i) superinjection of carriers, (ii) optical confinement, (iii) electron confinement.

The realization of the wide-gap window effect was very important for photodetectors, solar cells and LEDs applications. It permitted to broaden considerably and to control precisely spectral region for solar cells and photodetectors and to improve drastically the efficiency for LEDs. Main physical phenomena in double and single classical heterostructures are shown in Fig. 1. Then it was only necessary to find heterostructures where these phenomena could be realized.

![Diagrams](image-url)

Figure 1. Main physical phenomena in classical heterostructures. (a) One-side injection and superinjection; (b) Diffusion in built-in quasi-electric field; (c) Electron and optical confinement; (d) Wide-gap window effect; (e) Diagonal tunneling through a heterostructure interface.
At that time general skepticism existed with respect to the possibility of creating the "ideal" heterojunction with a defect-free interface and first of all with theoretical injection properties. Even a very pioneering study of the first lattice-matched epitaxially grown single-crystal heterojunctions Ge-GaAs by R. L. Anderson [14] did not give any proof of the injection of nonequilibrium carriers in heterostructures. Actual realization of the efficient wide-gap emitters was considered as simply close to the impossible and the double heterostructure laser patent was often referred to as a "paper patent".

Mostly due to this general skepticism there existed only a few groups trying to find out the "ideal couple", which was, naturally, a difficult problem. There should be met many conditions of compatibility between thermal, electrical, crystallochemical properties and between the crystal and the band structure of the contacting materials.

A lucky combination of a number of properties, i.e. small effective mass and wide energy gap, effective radiative recombination and a sharp optical absorption edge due to the "direct" band structure, a high mobility at the absolute minimum of the conduction band and its strong reduction of the nearest minimum at the (100) point ensured for GaAs even at that time a place of honor in semiconductor physics and electronics. Since the maximum effect is obtained by using heterojunctions between the semiconductor serving as the active region and more wide band material, the most promising systems looked at in that time were GaP-GaAs and AlAs-GaAs. To be "compatible", materials of the "couple" should have, as the first and the most important condition, close values of the lattice constants; therefore heterojunctions in the system AlAs-GaAs were preferable. However prior to starting work on preparation and study of these heterojunctions one had to overcome a certain psychological barrier. AlAs had been synthesized long ago [15], but many properties of this compound remained unstudied since AlAs was known to be chemically unstable and decompose in moist air. The possibility of preparing stable specimens adequate to applications of heterojunctions in this system seemed to be not very promising.

Initially, our attempts to create DHS were related to a lattice-mismatched GaAsP system. And we succeeded in fabricating by VPE the first DHS lasers in this system. However, due to lattice mismatch the lasing like that in heterojunction lasers occurs only at liquid nitrogen temperature [16]. From a more curious point of view, I would like to mention that it was the first practical result obtained for a lattice mismatched, even partially relaxed, system.

Our experience, which we got from studying the GaAsP system, was very important for understanding many specific heterojunction physical properties and basics of heteroepitaxy. Development of a multichamber VPE method for the GaAsP system permitted us to create in 1970 superlattice structures with a 200 Å-period and to demonstrate the splitting of the conduction band [17].

But from the general point of view at the end of 1966, we came to a conclusion that even a small lattice mismatch in heterostructures such as GaP$_{0.15}$As$_{0.85}$-GaAs did not permit to realize potential advantages of the DHS. At that time my co-worker D. N. Tret'yakov told me that small crystals of
Al$_x$Ga$_{1-x}$As solid solutions of different compositions, which had been prepared two years ago by cooling from a melt, were put in the desk drawer by Dr. A. S. Bortshevsky and nothing happened to them. It immediately became clear that Al$_x$Ga$_{1-x}$As solid solutions turned out to be chemically stable and suitable for preparation of durable heterostructures and devices. Studies of phase diagrams and the growth kinetics in this system and development of the LPE method especially for heterostructure growth soon resulted in fabricating the first lattice-matched AlGaAs heterostructures. When we published the first paper on this subject, we were lucky to be the first to find out a unique, practically an ideal lattice-matched system for GaAs, but as it frequently happened simultaneously and independently the same results were achieved by H. Rupprecht and J. Woodall at T. Watson IBM Research Center [18].

Then the progress in the semiconductor heterostructure area was very rapid. First of all we experimentally proved the unique injection properties of the wide-gap emitters and the superinjection effect [19], the stimulated emission in AlGaAs DHS [20], established the band-diagram of Al$_x$Ga$_{1-x}$As–GaAs$_x$ heterojunction, carefully studied luminescence properties, diffusion of carriers in a graded heterostructure and very interesting peculiarities of the current flow through the heterojunction – that is similar, for instance, to diagonal tunneling-recombination transitions directly between holes of the narrow-band and electrons of the wide-band heterojunction components [21].

At the same time, we created the majority of the most important devices with realization of the main advantages of the heterostructure concepts:
- Low threshold at room temperature DHS lasers [22] (Fig. 2).

![Emission Spectrum of Al$_x$Ga$_{1-x}$As DHS Laser](image)

**Figure 2.** Emission spectrum of the first low-threshold Al$_x$Ga$_{1-x}$As DHS laser operating at room temperature (300 K), $J_0 = 4500$ A/cm$^2$. The current rises: (a) from (1) 0.7 A to (2) 8.5 A and then to (3) 13.6 A; (b) from (1) 13.6 A to (2) 18 A; $s = 2.2 \times 10^{-3}$ cm$^2$. 
– High effective SHS and DHS LED [23]
– Heterostructure solar cells [24]
– Heterostructure bipolar transistors [25]
– Heterostructure p-n-p-n switching devices [26].

One of the first successful applications in industrial scale production in our
country was heterostructure solar cells in space research. We transferred our
technology to the “Quant” company and since 1974, GaAlAs solar cells have
been installed on many of our sputniks. Our space station “Mir” (Fig. 4) has
been using them for 15 years.

Most of these results were achieved afterwards in other laboratories in 1–2
years and in some cases even later. But in 1970, the international competition
became very strong. Later on one of our main competitors I. Hayashi, who was
working together with M. Panish at Bell Telephone Lab in Murray Hill, wrote
[27]: “In September 1969, Zhores Alferov of the Ioffe Institute in Leningrad
visited our laboratory. We realized he was already getting a \( J_{th}^{(300)} \) of 4.3
kA/cm² with a DH. We had not realized that the competition was so close and
redoubled our efforts ... Room temperature CW operation was reported in
May 1970 ...” . In our paper published in 1970, [28] the CW lasing was real-
ized in stripe-geometry lasers formed by photolithography and mounted on
copper plates covered by silver (Fig. 3). The lowest \( J_{th} \) density at 300 K was 940
A/cm² for broad area lasers and 2.7 kA/cm² for stripe lasers. Independently,
CW operation in DHS lasers was reported by Izu Hayaishi and Morton Panish
[29] (for broad area lasers with diamond heatsinks) in a paper submitted
only one month later than our work. Achievement of CW at room tempera-
ture produced an explosion of interest in physics and technology of semi-
conductor heterostructures. If in 1969, AlGaAs heterostructures were studied
just in a few laboratories mostly in USSR and US (A. F. Ioffe Institute, “Polyus”
and “Quant” – industrial lab’s where we transferred our technology for appli-
cations in USSR; Bell Telephone, D. Sarnoff RCA Research Center, T. Watson
IBM Res.Center in US), in the beginning of 1971, many Universities, indus-
trial labs in USA, USSR, United Kingdom, Japan and even in Brazil and
Poland started investigations of III–V Heterostructures and Heterostructure
device.

![Figure 3. Schematic view of the structure of the first injection DHS laser operating in the CW regime at room temperature.](image-url)
At this early stage of the development of the heterostructure physics and technology it became clear that we needed to look for new lattice-matched heterostructures in order to cover a broad area of the energy spectrum. The first important step was done in our laboratory in 1970: in the paper [30] we reported that various lattice-matched heterojunctions based on quaternary III–V solid solutions were possible, which permitted independent variation between lattice constant and band gap. Later on G. Antipas with co-workers [31] came to the same conclusions. As a practical example utilizing this idea we considered different InGaAsP compositions and soon this material was recognized among the most important ones, for many different practical applications: photocathodes [32] and especially lasers in infra-red region for fiber optical communications [33] and the visible [34].
In the early 1970s a “world map” of ideal lattice-matched heterostructures was shown (Fig. 5). Not before a decade later this “world map” was drastically changed (Fig. 6). Nowadays, it is necessary to add III-nitrides.

The main ideas of a semiconductor distributed-feed-back laser were formulated by us in the patent in 1971 [35]. The same year H. Kogelnilk and C. V. Shank considered the possibility of replacing the Fabry–Perot or similar types of resonator in dye-lasers with volume periodical inhomogeneities [36]. It is necessary to note, that their approach is not applicable to semiconductor lasers and all laboratories that carried out research in DFB and DBR semiconductor lasers used the ideas formulated in [35]:

1. Diffraction grating created not in volume, but on a surface waveguide layer
2. Interaction of waveguide modes with a surface diffraction grating, giving not only distributed feedback but also highly collimated light output.

Detailed theoretical analyses of the semiconductor laser with surface diffraction grating was published in 1972 [37]. In this paper the authors established the way for the single-frequency generation. First semiconductor lasers with surface diffraction grating and distributed feedback were realized practically simultaneously at the Physico-Technical Institute [38], Caltech [39] and Xerox Lab in Palo Alto [40].

In the early 1980s H. Kroemer and G. Griffiths published a paper [41] which stimulated strong interest in staggered line-up heterostructures (type II heterojunction). Spatial separation of electrons and holes at the interface results in a tunability of their optical properties [21(c), 42]. Staggered band alignment gives a possibility to realize optical emission with a photon energy much smaller than the band-gap energy of each of the semiconductors form-

![Energy gaps vs lattice constants for semiconductors of IV elements, III-V and II(IV)-VI compounds and magnetic materials in parentheses. Lines connecting the semiconductors, solid for the III-V's and dotted for the others, indicate quantum heterostructures that have been investigated.](image)
ing a heterojunction. The demonstration of an injection laser based on a
type-II GaInAsSb–GaSb heterojunction [42] opens good perspectives for the
creation of effective coherent light sources in the infrared optical range.
Radiation in such a device is due to the recombination of electrons and holes
localized in self-consistent potential wells at different sides of the hetero-
interface. Thus, type-II heterostructures open possibilities both for funda-
mental physics and for device applications, which cannot be realized with
type-I heterostructures in the III–V material system. However, practical appli-
cations of these structures are still hampered by a poor understanding of
their fundamental properties and the few actual systems, which have been
studied experimentally up to now.

To summarize this brief review of the classical heterostructures develop-
ment, it would be very convenient to classify the most important results in the
following way:

**Classical Heterostructures**

*I. Fundamental Physical Phenomena (Fig. 1)*
- One-side injection
- Superinjection
- Diffusion in Built-in Quasi-Electric Fields
- Electron confinement
- Optical confinement
- Wide-gap window effect
- Diagonal tunneling through heterostructure interface

*II. Important Applications in Electronics*
- Semiconductor lasers – Low threshold and CW at room temperature, DFB
  and DBR Lasers, Vertical Surface Emitting Lasers, IR II-type Heterostruc-
  ture Lasers
- High efficient LED
- Solar cells and photodetectors, based on the wide-gap window effect
- Semiconductor integrated optics, based on semiconductor DFB and DBR
  lasers
- Bipolar wide-gap transistors
- Transistors, thyristors, dynistors with photonic signal transmission
- High power diodes and thyristors
- Infra-red to visible converters
- Efficient cold cathodes.

*III. Important Technological Peculiarities*
- Lattice-matched structures are a necessity of principle
- Multi-component solid solutions are used for lattice-matching
- Epitaxial growth technology is needed of principle.

Concluding this concise summary of early development of bulk hetero-
structures one may say that the invention of an “ideal” heterojunction
and the introduction of the heterostructure concept into semiconductor physics and technology have led to the discovery of new physical effects, cardinal improvement of characteristics of practically all known semiconductor devices, and the invention of new ones.

3. HETEROSTRUCTURE QUANTUM WELLS AND SUPERLATTICES

Owing to electron confinement in DHS, the double-heterostructure laser became an important precursor of a quantum well structure: when a middle-layer had a thickness of some hundred ångströms, the electron levels would split due to the quantum-size effect. The development of heterostructure growth techniques gave a possibility to fabricate high quality double heterostructures with ultrathin layers. Two main methods of growth with very precise control of thickness, planarity, compositions etc. were developed in the 70s. A modern molecular beam epitaxy method became practically important for III–V heterostructure technology first of all due to the pioneering work of A. Cho [43]. Metallo-organic chemical vapor deposition originated from an early work of H. Manasevit [44] and found broad application in III–V heterostructure research after R. Dupuis and P. Dapkus reported the room temperature injection of AlGaAs DHS lasers which had been grown by the MOCVD method [45].

Clear manifestation of the quantum-size effect in optical spectra of GaAs–AlGaAs semiconductor heterostructures with ultrathin GaAs layers (quantum wells) was demonstrated by Raymond Dingle et al. in 1974 [46]. The authors observed a characteristic step-like behavior in absorption spectra and systematic shifts of the characteristic energies with a quantum well width decrease.

Studies of superlattices were started by the work of L. Esaki and R. Tsu in 1970 [47] who considered the electron transport in a superlattice, i.e. an additional periodic potential created by doping or changing the composition of semiconductor materials with the period bigger, but comparable with a lattice constant of a crystal. In this, as Leo Esaki called it “man-made crystal”, a parabolic band would break into mini-bands separated by small forbidden gaps and having Brillouin zones determined by this period. Similar ideas were described by L. V. Keldysh in 1962 [48] when considering the periodic potential produced on a semiconductor surface by an intense ultrasonic wave. At the Physico-Technical Institute R. Kazarinov and R. Suris theoretically considered the current flow in superlattice structures in the early 1970s [49]. It was shown that the current between wells is determined by tunneling through the potential barriers separating the wells and the authors predicted a very important phenomena: tunneling under electric field when the ground state of a well coincides with an excited state of the next well and stimulated emission resulting from photon-assisted tunneling between the ground state of one well and excited state of a neighboring well, which is lower by the energy due to applied electric field. At that time L. Esaki and R. Tsu independently considered resonant tunneling in superlattice structures [50].
The pioneering experimental studies of the superlattice structures were carried out by L. Esaki and R. Tsu: the superlattices were grown by VPE in the system GaP$_{x}$As$_{1-x}$–GaAs. At the same time in our laboratory we developed the first multichamber apparatus and, as it was mentioned before, prepared a superlattice structure GaP$_{0.9}$Al$_{0.1}$–GaAs with the thickness of each layer equal to 100 Å and total number of the layers being 200 [17]. Observed peculiarities of the volt-current characteristics, their temperature dependence and photoconductivity were explained by the splitting of the conduction band due to the one-dimensional periodic potential of the superlattice. These first superlattices were also the first strained-layer superlattices. E. Blakeslee and J. Matthews who were working with L. Esaki and R. Tsu at IBM succeeded in the mid-70’s in growing strained-layer superlattices with a very low concentration of defects. But many years later, after G. Osbourn’s theoretical study [51] at Sandia lab and the first successful preparation of a high quality strained-layer superlattice GaAs–In$_{0.2}$Ga$_{0.8}$As by M. Ludowise at Varian, N. Holonyak at the University of Illinois achieved the CW room-temperature laser action on those structures[52]. It became clear that in a strained-layer superlattice the lattice strain became an additional degree of freedom and by varying the layer thicknesses and compositions one can vary continuously and independently of one another the forbidden gap, lattice constant etc. of the overall superlattice.

In the early 1970s L. Esaki with co-workers passed to MBE technology in the AlGaAs system [53] and in March 1974, they submitted a paper on resonant tunneling [54]. It was the first experimental demonstration of quantum well heterostructure physics. They measured the tunneling current and conductance as a function of an applied voltage in GaAs–GaAlAs double barriers and found current maxima associated with resonant tunneling. Later in the same year L. Esaki and L. L. Chang observed resonant tunneling in a superlattice [55]. Strong interest in resonant tunneling obviously was connected with its potential applications in high-speed electronics. In the late 1980, picosecond operation has been achieved in a double resonant tunnel diode and oscillations up to 420 GHz were reported in a GaAs resonant tunnel diode at room temperature.

The restriction of the electron motion to two dimensions in field effect transistors were recognized long time ago [56] and for trapped electrons in inversion layers was first verified by the magneto-conductance experiment by A. B. Fowler et al. in 1966 [57]. Spectral effects due to spatial quantization were observed in thin bismuth films in 1968 by V. N. Lutskii and L. A. Kulik [58].

The pioneering work on modulation-doped superlattices [59] demonstrating a mobility enhancement with respect to the bulk crystal, stimulated research data on application of the high-mobility two-dimensional electron gas for microwave amplification. In France and Japan practically simultaneously new types of transistors based on a single nAlGaAs–nGaAs modulation-doped heterostructure were created that were labeled TEGFET (two dimensional electron gas FET) in France [60] and HEMT (high electron mobility transistor) in Japan [61].
The first quantum well laser operation was demonstrated by J. P. van der Ziel et al. [62] but parameters of the lasing were much worse than for average DHS lasers. In 1978, R. Dupuis and P. Dapkus in collaboration with N. Holonyak first reported about the quantum well laser with parameters comparable with conventional DHS lasers [63]. The name "quantum well" was used in that paper. Real advantages of quantum well lasers were demonstrated much later by W. T. Tsang at Bell Telephone Lab. Thanks to many improvements of MBE growth technology and introducing an optimized structure (GRIN SCH) he found threshold currents as low as 160 A/cm² [64].

We started to develop MBE and MOCVD methods of growth of III–V heterostructures only in the late 1970s. First of all, we stimulated the design and construction of the first Soviet MBE machine in our electronic industry. During a few years there were developed three generations of MBE machine and the last, which had the name "Cna" (the nice river not very far from Ryazan – the city where NITI – industrial laboratory of the Electronic Industry was located, this NITI carried out development of MBE machine) were good enough for our goals. In parallel, later on, we started to develop MBE system with NTO AN – scientific Instruments company of the Academy of Sciences in Leningrad. In the middle of the 80's we got a few systems of this version. Both types of MBE systems are still working at the Ioffe Institute and other laboratories.

The MOCVD systems we developed just in our Institute and later, in the 1980s, a Swedish company "Epiquip" specially designed with our participation a couple of systems for our Institute, which are still used in our research.

The strong interest in experimental study of low-dimensional structures and lack of equipment for MBE and MOCVD growth technology stimulated our research on the development of LPE suitable for quantum well heterostructures.

However, until the late 1970s it seemed impossible to grow III–V heterostructures with an active-region thickness of less than 500 Å by LPE because of the existence near the heterojunction of extended interface regions with varying chemical compositions.

The situation was changed due to work of N. Holonyak et al. [65] for superlattice like InGaAsP structures by using a rotating boat system. In our laboratory we developed a new LPE method with the usual translational motion in a standard horizontal system for InGaAsP heterostructures [66] and a low-temperature LPE method for AlGaAs heterostructures [67]. These methods permitted us to prepare practically any kind of excellent quality quantum well heterostructures with the thickness of the active region up to 20 Å and with the size of the interface regions comparable to one lattice constant. Of great practical importance for InGaAsP laser heterostructures was the creation of a record threshold current density for InGaAsP/InP (λ = 1.3 and 1.55 µm) and for InGaAsP/GaAs (λ = 0.65–0.9 µm) single quantum well separate confinement lasers [68]. For high power InGaAsP/GaAs (λ = 0.8 µm) lasers a total efficiency of 66 % with CW power 5 W for 100 µm width of the strips, a stripe-geometry laser was achieved [69]. In this lasers effective cooling of the semi-
conductor power device by recombination radiation was for the first time realized as it had been predicted much earlier [13]. Another important conclusion about InGaAsP heterostructure was its unusual resistance to multiplication of dislocations and defects (Fig. 7) [70]. It was this research that made a start to broad application of Al-free heterostructures.

![Figure 7. Time evolution of DHS active region under high level photoexcitation. AlGaAs/GaAs (a), InGaAsP/GaAs (b). Diameter of Kr^+ laser excitation beam ~ 40 μm. Excitation level (a) 10^4 W/cm^2, (b) 10^5 W/cm^2.](image)

A most complicated quantum well laser structure, which combined a single quantum well with short period superlattices (SPSL), for the creation of GRIN SCH (the most favourable for the lowest Jth) was demonstrated in our laboratory in 1988 [71] (Fig. 8). Using SPSL we achieved not only the desirable profile of a graded wave-guide region, thus creating a barrier for dislocation movement to the active layer, but also got a possibility to grow different

![Figure 8. MBE grown SPSL QWSCH laser structure.](image)
parts of the structure at sufficiently large differences of the temperature. In this way, we have obtained both an excellent surface morphology and a high internal quantum efficiency on a planar GaAs (100) surface. The lowest $J_{\text{th}} = 52 \text{ A/cm}^2$ and, shortly after some small optimization, $40 \text{ A/cm}^2$ was for a long time a world record for semiconductor injection lasers and a good demonstration of the application of quantum wells and superlattices in electronic devices.

The idea of stimulated emission in superlattices that had been published by R. Kazarinov and R. Suris [49] was realized nearly a quarter of century after the proposal, by Federico Capasso [72]. The proposed structure was strongly improved and a cascade laser developed by F. Capasso gave rise to a new generation of unipolar lasers operating in the middle-infrared range.

The history of the semiconductor lasers is, from a certain point of view, the history of evolution of the semiconductor laser current threshold, which is shown in Fig. 9. The most dramatic changes happened just after the introduction of the DHS concept. Impact of SPDS QW led practically to a theoretical limit of this most important parameter. What can happen as a result of application of new quantum wires and quantum dots structures will be discussed in the next part of our paper.

![Figure 9. Evolution of the threshold current of semiconductor lasers.](image)

Maybe the crown of the quantum well studies was the discovery of the Quantum Hall effect [73]. This discovery and its comprehensive studies in AlGaAs-GaAs heterostructures, which shortly led to the fractional quantum Hall effect discovery [74], had a principal effect on the whole physics of the solid state. Observation of the effect which deals only with fundamental quantities and does not rely on peculiarities of the band structure, carrier mobility and densities in a semiconductor, has shown that heterostructures can be used to model some very basic physical effects. Recently many studies in this area have been concentrated on understanding condensation of electrons and search of Wigner crystallization.
To summarize this part of our paper we present the summary in a way like that in the previous classical heterostructure part.

**Heterostructure Quantum Wells and Superlattices**

**I. Fundamental Physical Phenomena**

- 2D electron gas
- Step-like density-of-state function
- Quantum Hall effect
- Fractional Quantum Hall effect
- Excitons at room temperature
- Resonant tunneling in double-barrier structures and superlattices
- Energy spectrum in superlattices is determined by choice of potential and strain
- Stimulated emission at resonant tunneling in superlattices
- Pseudomorphic growth of strained structures.

**II. Important Consequences for Applications**

- Shorter emission wavelength, reduced threshold current, larger differential gain and reduced temperature dependence of the threshold current for semiconductor lasers
- IR quantum cascade laser
- SPSL QW laser
- Optimization of Electron and Light Confinement and Waveguiding for semiconductor lasers
- 2D electron gas transistors (HEMT)
- Resonant-tunneling diodes
- Precise resistance standards
- SEEDs and electro-optical modulators
- IR photodetectors based on quantum size level absorption.

**III. Important Technological Peculiarities**

- Lattice-match is unnecessary
- Low growth-rate technology (MBE, MOCVD) is needed of principle
- Submonolayer growth technique
- Blockading mismatch dislocations during epitaxial growth
- Sharp increase of the variety of heterostructure components.

**4. HETEROSTRUCTURE QUANTUM WIRES AND QUANTUM DOTS**

The principal advantage of application of quantum-size heterostructures for lasers originates from the noticeable increase of the density of states with reducing dimensionalities for the electron gas (Fig. 10).

During the 1980s, progress in 2D-quantum well heterostructures physics and its applications attracted many scientists to studying systems of far less dimensionality – quantum wires and quantum dots. In contrast to quantum
“wells” where carriers are localized in the direction perpendicular to the layers but move freely in the layer plane, in quantum “wires” carriers are localized in two directions and move freely along the wire axis. And being confined in all three directions quantum “dots” – “artificial atoms” with a totally discrete energy spectrum are created (Fig. 11).

Experimental work on fabrication and investigation of quantum wire and dot structures began more than 15 years ago. In 1982, Y. Arakawa and H. Sakaki [75] theoretically considered some effects in lasers based on heterostructures with size quantization in one, two, and three directions. They wrote: “Most important, the threshold current of such a laser is reported to be far less sensitive than that of a conventional laser reflecting the reduced dimensionality of the electronic state.” The authors performed experimental studies on a QW laser placed in high-magnetic fields directed perpendicular to the QW plane and demonstrated that the characteristic temperature ($T_0$) describing the exponential growth of the threshold current with temperature increases in a magnetic field from 144 to 313 °C. They pointed to a possibility to weaken the threshold current dependence on temperature for QWR lasers and full temperature stability for QD lasers (Fig. 12). By now there is a significant number of both theoretical and experimental papers in this field.

Figure 11. Schematic representation of energy diagrams in case of a single atom (left), a bulk crystal (center), and a quantum dot (right).
The first semiconductor dots based on II-VI microcrystals in glass matrix were proposed and demonstrated by A.I. Ekimov and A.A. Onushchenko [76]. However, since the semiconductor quantum dots were introduced in an insulating glass matrix and the quality of the interface between glass and semiconductor dot was not high, both fundamental studies and device applications were limited. Much more exciting possibilities appeared after three dimensional coherent quantum dots had been fabricated in a semiconductor matrix [77].

Several methods were proposed for the fabrication of these structures. Indirect methods, such as the post-growth lateral patterning of 2D quantum wells suffer often from insufficient lateral resolution and interface damage caused by the patterning procedure. A more promising way is the fabrication by direct methods, i.e., growth in V-grooves and on corrugated surfaces which may result in formation of quantum wires and dots. The groups of the Ioffe Institute and Berlin Technical University – last years we carried out this research in close cooperation – contributed significantly to the last direction.

Finally we came to the conclusion that the most exciting method of the formation of ordered arrays of quantum wires and dots is the self-organization phenomena on crystal surfaces. Strain relaxation on step or facet edges may result in formation of ordered arrays of quantum wires and dots both for lattice-matched and lattice-mismatched growth.

The first very uniform arrays of three dimensional quantum dots exhibiting also lateral ordering were realized in the system InAs–GaAs both by MBE and MOCVD growth methods [78, 79].

Elastic strain relaxation on facet edges and island interaction via the strained substrate are driving forces for self-organization of ordered arrays of uniform, coherently strained islands on crystal surfaces [80].

In lattice-matched heteroepitaxial systems the growth mode is determined
solely by the relation between the energies of two surfaces and the interface energy. If the sum of the surface energy of epitaxial layer $\gamma_2$ and energy of interface $\gamma_{12}$ is lower than the substrate surface energy, $\gamma_2 + \gamma_{12} < \gamma_1$, i.e., if the material 2 being deposited wets the substrate, then we have the Frank–van der Merve growth. Changing the $\gamma_2 + \gamma_{12}$ value may result in a transition from the Frank–van der Merve mode to on Volmer–Weber one where 3D islands are formed on a bare substrate.

In a heteroepitaxial system with lattice mismatch between the material being deposited and the substrate the growth may initially proceed in a layer-by-layer mode.

However, a thicker layer has a higher elastic energy, and the elastic energy tends to be reduced via formation of isolated islands. In these islands the elastic strains relax and, correspondingly, the elastic energy decreases. This results in a Stranski–Krastanow growth mode (Fig. 13). The characteristic size of islands is determined by the minimum in the energy of an array of 3D coherently strained islands per unit surface area as a function of the island size (Fig. 14) [80]. Interaction between islands via elastically strained substrate would results in lateral island ordering typical of the square lattice.

Experiments show in most cases rather narrow size distribution of the islands, and on top of that coherent islands of InAs form under certain conditions a quasi-periodic square lattice (Fig. 15). The shape of quantum dots can be significantly modified during regrowth or post-growth annealing, or by applying complex growth sequences. Short period alternating deposition of strained materials leads to a splitting of QDs and to formation of vertically coupled quantum dot superlattice structures (Fig. 15) [81]. Ground state QD emission, absorption and lasing energies are found to coincide [78].
Observation of ultranarrow (< 0.15 meV) luminescence lines from single quantum dots [78], which do not exhibit broadening with temperature, is the proof of the formation of an electronic quantum dot (Fig. 16).
Quantum dot lasers are expected to have superior properties with respect to conventional QW lasers. High differential gain, ultralow threshold current density and high temperature stability of the threshold current density are expected to occur simultaneously. Additionally ordered arrays of scatterers formed in an optical waveguide region may result in distributed feedback and (or) in stabilization of single-mode lasing. Intrinsically buried quantum dot structures spatially localize carriers and prevent them from recombining non-radiatively at resonator faces. Overheating of facets, being one of the most important problems for high-power and high efficiency operation of AlGaAs–GaAs and AlGaAs–InGaAs lasers, may thus be avoidable.

Since the first realization of QD lasers [82], it has become clear that the QD size uniformity was sufficient to achieve good device performance. But even at that time, it was recognized that the main obstacle for QDHS laser operation at room and elevated temperatures was connected to temperature-induced evaporation of carriers from QD's. Different methods were developed to improve the laser performance: (i) the increase of the density of QD's by stacking of QD's (Fig. 17); (ii) the insertion of QD's into a QW sheet; (iii) the use of a matrix material with a higher bandgap energy. As a result, we got many parameters of QDHS lasers better than ones for QWHS lasers based on the same materials. As an example, the world-record threshold-current density of 19 A/cm² has been recently achieved [83]. Further, the cw-output power up to 3.5–4.0 W (CW) for a 100-μm strip width, the quantum efficiency of 95 % and the wall-plug efficiency of 50 % were obtained [84].

Significant activities in theoretical understanding of QD lasers with realistic parameters have been performed. For a QD size dispersion of about 10 % and other practical structure parameters, the theory [85] predicts typical threshold-current densities of 5 A/cm² at room temperature. The value of 10 A/cm² at 77 K [86] and even 5 A/cm² at 4 K [87] have been experimentally observed.

Figure 17. Transmission electron microscopy image of the active region of high-power QDHS laser.
Figure 18. Schematic view of the QD VCSEL structure (a).
Basic advantages of quantum dots: (i) no interface recombination at oxide-defined apertures; (ii) reduced lateral spreading of carriers out of the aperture region. Single QD laser at ultralow threshold current is possible. Dependence of the threshold current on the aperture size in a QD VCSEL (b). (i) low threshold current densities (170 A/cm² at 300 K); (ii) low threshold currents at ultrasmall apertures; (iii) 1.3 μm range on GaAs substrate?

In view of the advanced device applications of QD’s, the incorporation of QD’s in vertical-cavity surface-emitting lasers (VCSEL’s) seems to be very important. QD VCSEL’s with parameters, which fit to the best values for QW devices of the similar geometry, have been demonstrated (Fig. 18) [88]. Recently, very promising results for 1.3-μm QD VCSEL’s on a GaAs substrate to use in fiber optical communications have been obtained (Fig. 19) [89].

In a free-standing 3D island formed on a lattice-mismatched substrate, the strains can relax elastically, without the formation of dislocations. Thus, sufficiently large volume of a coherent narrow-gap QD material can be realized. This makes it possible to cover a spectral range of 1.3–1.5 μm using a GaAs substrate and to develop wavelength-multiplexing systems on the basis of QD VCSEL’s in the future.
It is very important to emphasize that we always realized the DHS concept for QWR and QD structures because in both cases we have a narrow gap-band material in a wide-gap matrix.

Let us summarize this part again by the same way as it had been done for other parts.

**Heterostructure Quantum Wires and Dots**

**I. Fundamental Physical Phenomena**
- 1D electron gas (wires)
- Density-of-state function with sharp maximums (wires)
- 0D-electron gas (dots)
- Delta-function type of density-of-state function (dots)
- Increasing binding energy of excitons.

**II. Important Applications in Electronics**
- Reduced lasing threshold current and larger differential gain
- Reduced temperature dependence of threshold current (wires)
- Temperature stability of the threshold current (dots)
- Discrete amplification spectrum and a possibility of obtaining performance characteristics similar to those of solid-state or gas lasers (dots)
- Higher modulation factor in electro-optical modulators
- A possibility of creation “single-electron” devices
- A new possibility for the development of FET transistors.

**III. Important Technological Peculiarities**
- The application of self-organization effects for growth
- Epitaxial growth in V-groves (wires)
- High resolution lithography of HSQWL.
5. FUTURE TRENDS

Recently very impressive results for short wave-length light sources have been achieved on the basis of II–VI selenides and III–V nitrides. The success in this research was mostly determined by application of heterostructure concepts and methods of growth which had been developed for III–V quantum wells and superlattices. The natural and most predictable trend is the application of the heterostructure concepts as well as technological methods and peculiarities to new materials. Different III–V, II–VI and IV–VI heterostructures, developed in recent time, are good examples of this statement.

But from a general and more deep point of view, heterostructures (it concerns all of them: the classical, QWs and SLs, QWRs and QDs) are a way of the creation of new types of materials – Heterosemiconductors. By using Leo Esaki words – instead of “God made crystals” we create by ourselves – “Man made crystals”.

The classical heterostructures, quantum wells and superlattices are quite mature and we exploit many of their unique properties. Quantum wires and dots structures are still very young: exciting discoveries and new unexpected applications are awaiting us on this way. Even now we can say that ordered equilibrium arrays of quantum dots may be used in many devices: lasers, light modulators, far-infrared detectors and emitters, etc. Resonant tunneling via semiconductor atoms introduced in larger band-gap layers may lead to significant improvement in device characteristics. More generally speaking, QD structures will be developed both “in width” and “in depth”.

In width means new material systems to cover new energy spectrum. The life-time problems of the green and blue semiconductor lasers and even more general problems of the creation of defect-free structures based on wide-gap II–VI and III–V (nitrides) would be solved by using QDs structures in these systems.

As to in depth it is necessary to mention that the degree of ordering depends on very complicated growth conditions, materials constants, concrete values of the surface free energy. The way to resonant tunneling and “single” electron devices including optical ones is a deep detailed investigation and evaluation of these parameters in order to achieve the maximal possible degree of ordering. In general, it is necessary to find out more strong self-organization mechanisms for ordered arrays of QDs creation.

In the early 1980s I was invited to deliver a lecture about heterostructures and applications at the Amoco Photonic Center near Chicago.

The summary of my lecture was as follows:
1. Heterostructures – a new kind of semiconductor materials:
   – expensive, complicated chemically & technologically but most efficient
2. Modern optoelectronics is based on heterostructure applications
   – DHS laser – key device of the modern optoelectronics
   – HS PD – the most efficient & high speed photo diode
   – OEIC – only solves problem of high information density of optical communication systems
3. Future high speed microelectronics will mostly use heterostructures
4. High temperature, high speed power electronics – a new broad field of heterostructure applications
5. Heterostructures in solar energy conversion: the most expensive photo-cells and the cheapest solar electricity producer
6. In the 21st century heterostructures in electronics will reserve only 1 % for homojunctions.

And 20 years later I do not like to change any word there.

It is hardly possible to describe even the main directions of the modern physics and technology of semiconductor heterostructures. There is much more than was mentioned. Many scientists contributed to this tremendous progress which not only defines to a great extent the future prospects of condensed matter physics and semiconductor laser and communication technology but, in a sense also, the future of the human society. I would like also to emphasize the impact of scientists of previous generations who prepared our way. I am very happy that I had a chance to work in this field from the very beginning. I am even more happy that we can continue to contribute to the progress in this area now.

REFERENCES


[34] Zh. I. Alferov, I. N. Arsent’ev, D. Z. Garbuzov, S. G. Konnikov, and V. D. Rumyantsev, "Generation of coherent radiation in nGa0.5In0.5P-pGa0.55In0.45As heterolayers, nGa0.5In0.5P-nGa0.5In0.5P" Pisma Zh. Tech. Fiz., 1, pp. 305–310, 1975 [Sov. Phys.-Tech. Phys. Lett., 1, pp. 147–148, 1975];
W. R. Hitchens, N. Holonyak Jr., P. D. Wright, and J. J. Coleman, “Low-threshold LPE In$_{1-x}$Ga$_x$P$_{1-y}$As$_y$/In$_{1-x}$Ga$_x$As$_y$/In$_{1-x}$Ga$_x$P$_{1-y}$As$_y$: yellow double-heterojunction laser diodes (J < 10$^5$ A/cm$^2$, I = 5850 A, 77 K),” *Appl. Phys. Lett.*, **27**, p. 245, 1975.


[65] E. Rezek, H. Shichijo, B. A. Vojak and N. Holonyak, "Confined-carrier luminescence of a thin InxGa1-xP0.1As0.9 well (x = 0.13, z = 0.29, ~ 400 A) in an InP p-n-junction". Appl. Phys. Lett. 31, p. 534, 1977.


[73] K. v. Klitzing, G. Dorda and M. Pepper, “New method for high-accuracy determina-
tion of the fine-structure constant based on quantized Hall resistance”.


