Classification of Ge hut clusters in arrays formed by molecular beam epitaxy at low temperatures on the Si(001) surface

L V Arapkina, V A Yuryev

Contents

1. Introduction 279
   1.1 Statement of the problem; 1.2 Formation of Ge hut clusters. A brief excursion into the historical domain 279
   2. Experimental 283
   3. Classification 284
      3.1 Main species of clusters; 3.2 Derivative species of clusters
   4. Conclusion 289
      References 290

Abstract. Ge hut clusters forming quantum dot arrays on the Si(001) surface in the process of low-temperature ultrahigh-vacuum molecular beam epitaxy are morphologically investigated and classified using in situ scanning tunnelling microscopy. It is found that two main Ge hut cluster types—pyramidal and wedge-shaped—have different atomic structures, and it is concluded that shape transitions between the two are impossible. Derivative cluster species—obelisks (or truncated wedges) and accreted wedges—are revealed and investigated for the first time and shown to start dominating at high coverage. The uniformity of cluster arrays is shown to be controlled by the scatter in the lengths of wedge-like clusters. At low growth temperatures (360 °C), cluster nucleation during the growth of the array is observed for all values of Ge coverage except for a particular point at which the arrays are more uniform than at higher or lower coverages. At higher temperatures (530 °C), no cluster nucleation is observed after the initial formation of the array.

1. Introduction

1.1 Statement of the problem

The development of processes for the controllable formation of germanium quantum dot (QD) arrays on a silicon surface, as well as multilayer Ge/Si epitaxial heterostructures based on them, has been a subject of significant and constantly increasing effort for a number of years [1–4], primarily due to their potential applications in prospective devices of microelectronics and integrated microphotronics, whose formation process is compatible with monolithic silicon VLSI technology. Both the high density of the germanium nanoclusters (> 10^{11} \text{cm}^{-2}) and the high uniformity of the cluster shapes and sizes (dispersion < 10\%) in the arrays are required for the many practically important applications of such structures [2, 4–11].

The main technique of formation of germanium nanoclusters on a silicon surface is molecular beam epitaxy (MBE) [2, 3]. A high density of self-assembled clusters can be obtained in the MBE process of Ge/Si(001) structure formation when depositing germanium on a silicon substrate heated to a moderate temperature (≤ 550 °C). In this case, the lower the temperature of the silicon substrate in the process of the germanium deposition, the higher the density of the clusters at the permanent quantity of the deposited germanium [1, 12, 13]. For example, the density of germanium clusters in an array reached 6 × 10^{11} \text{cm}^{-2} at the substrate temperature T_{\text{gr}} = 360 °C during the deposition and an effective thickness h_{\text{Ge}} = 8 \text{A} of the deposited germanium layer, whereas a cluster density of only about 2 × 10^{11} \text{cm}^{-2} was obtained at T_{\text{gr}} = 530 °C and the same value of h_{\text{Ge}} [1].

There are also other ways to increase the cluster density in the arrays. Thus, the authors of Ref. [5] succeeded in reaching a cluster density of about 9 × 10^{11} \text{cm}^{-2} in an array using the pulsed irradiation of the substrate by a low-energy Ge^{+} ion beam during the MBE growth of a Ge/Si(001) heterostructure at T_{\text{gr}} = 570 °C.

Obtaining arrays of densely packed Ge QDs on an Si(001) surface is an important task, but the problem of formation of uniform arrays of Ge clusters is an even more challenging one. The process of Ge/Si(001) heterostructure formation with Ge

1 Notice also that the lowering of the array formation temperature to values of ≤ 450 °C is required for the compatibility of the Ge/Si(001) heterostructure formation process with a CMOS device fabrication cycle [1]. This is another reason to lower the temperature of all technological treatments starting from the Si surface preparation.

2 That is, the Ge coverage, or more accurately the thickness of the Ge film, measured by a graduated in advance film thickness monitor with a quartz sensor installed in the MBE chamber.
QD dense arrays and predetermined electrophysical and photoelectric parameters cannot be developed until both of these tasks are solved. The uniformity of the cluster sizes and shapes in the arrays determines not only the widths of the energy spectra of the charge-carrier bound states in the QD arrays [5] but in a number of cases the optical and electrical properties of both the arrays themselves and the device structures produced on their basis [14]. To find an approach to improving the Ge QD array uniformity on an Si(001) surface it is necessary to carry out a morphological investigation of the clusters constituting the arrays and first of all classify them.

At present, two genera of the self-assembled Ge clusters formed on an Si(001)† surface have been discriminated — huts and domes. The former are smaller and faceted by \( f_{105} \) planes, while the latter are much larger and have more sophisticated faceting [6, 13, 15–18]. Our investigations of densely packed Ge nanocluster arrays on an Si(001) surface [1] have shown that the composition of an ensemble of the hut clusters is by no means homogeneous — there are several species of hut clusters different in their geometrical shapes, as well as their behavior in the array formation process.3 Examples of the Ge cluster arrays formed on an Si(001) surface at different \( T_{gr} \) and \( h_{Ge} \) are given in Figs 1 and 2, which obviously demonstrate that the arrays grown at low temperature always consist of a set of morphologically different hut clusters. Some of them have often been discussed in the literature, since the classical letter by Mo et al. [16, 17, 20–27], and some others have not.

This paper is devoted to a study of morphological differences in Ge hut clusters formed in the process of molecular beam epitaxy at low substrate temperatures on an Si(001) surface and their classification derived from the revealed species distinctions.

In writing about the classification of the hut clusters, our goal is not to simply introduce new terminology, as it might seem, although this is also proposed.4 The aim of classification is to sort out the hut clusters in accordance with their structural peculiarities, which are much more important attributes than geometrical shapes. The differences in geometrical shapes do not necessarily imply a difference in atomic structures. The latter may be identical. Hence, if so, it may be assumed that one type of clusters originates from another, as is usually accepted ‘by default’ (postulated) for pyramids and ‘elongated’ hut clusters in the literature [24–26]. On the contrary, if the atomic structure of two species of clusters is different, the shape transitions in them seem to be very unlike

---

3 Here, we do not consider defects in the arrays. Another article will be devoted to their investigation. At present, all the available information about the morphology, structure, and density of array defects, as well as their effect on the array parameters, can be found in our reports [1] or [19].

4 Since the pioneering paper by Mo et al. [16], a good few descriptive and often confusing terms have been used in the literature to designate two known types of hut clusters. In this paper, we introduce new strictly stereometrical terminology to emphasize the structural differences between the cluster species and avoid muddle in the future. We shall name each species of the clusters in accordance with the denominations of the geometrical bodies which most accurately describe the shapes of the clusters.
because of the probable high potential barrier to be overcome to change the cluster atomic configuration and symmetry.

1.2 Formation of Ge hut clusters.

A brief excursion into the historical domain

It is well known [28] that in equilibrium conditions Ge layers grow on an Si(001) surface following the Stranaski–Krastanow mode [29, 30]. This means that initially Ge grows layer by layer, until it reaches a thickness of a few monolayers, then nucleation of the three-dimensional islands begins [2].

Because of lower Ge surface energy, growing Ge wets the Si surface (the etymology of the term ‘wetting layer’). Despite the mismatch of Si and Ge crystal lattices, which is \( \approx 4.2\% \), Ge atoms deposited on the Si(001) surface stay in correspondence with the Si atoms for several monolayers. The Ge layer surface is dimerized and \( (2 \times 1) \) reconstructed [28]. Due to dimer buckling, it comprises a mixture of the c\((4 \times 2)\) and p\((2 \times 2)\) structures [31, 32] which manifest themselves in STM images as characteristic antiphased and cophased zigzags. As Ge atoms arrive onto the surface, the compression along the dimer rows is relieved by arising dimer vacancies. Further ordering of the dimer vacancies into a nearly regular array of parallel trenches (or formation of so-called \( (2 \times n) \) reconstruction [28]), which are perpendicular to the dimer rows, goes on to reduce the surface strain energy beyond the Ge coverage of 0.8 monolayer (thickness of 1 ML \( \approx 1.4 \) A) [33].

Finally, the formation of the dimer-row vacancies and a grating of the quasiperiodic \( (M \times N) \) patched structure [32–35] exhausts the ability of the dimer vacancies to accommodate the strain increasing in the wetting layer with the growth of the quantity of deposited Ge. When the Ge coverage of the Si surface exceeds 3 ML, the three-dimensional nanoislands (or Ge nanoclusters) start to nucleate on the surface. They are free of dislocations at the Si/Ge interface (coherent with the Si substrate), are faceted, and have their base sides aligned with two orthogonal \( (100) \) axes. At moderate growth temperatures, the composition of the cluster arrays is always bimodal: a part of the clusters have the shape of regular square-based pyramids, while others have rectangles in their bases [16, 36]. Due to the shapes resembling huts, both square-based and rectangular-based clusters are usually referred to as ‘hut’ clusters. Hut clusters (coherent islands) were theoretically shown to be (under some conditions) more energetically stable than strained films or dislocated islands [20, 37]. Their appearance was also found to be kinetically favorable compared to the nucleation of dislocations [37].

The first STM observation of hut clusters was reported by Mo et al. [16] in 1990 (the term ‘hut’ was introduced by its authors). That article presented an experimental investigation of a newly discovered metastable phase of Ge clusters, which arose in the process of Stranaski–Krastanow growth to relieve the increasing stress in the wetting layer when Ge was deposited by MBE on Si(001) at moderate temperatures. As distinct from macroscopic clusters, hut clusters were found to have \([105]\) faceting on all sides. The first model of cluster \([105]\) faces, according to which these facets consist of \((001)\) terraces composed of \( \text{pairs of dimers} \) of the Ge(001)\( \text{-}(2 \times 1) \) reconstructed surface, was suggested in that communication (this model is now usually referred to as the PD model [36, 38, 39]). Huts were found to have “predominantly a prism shape (with canted ends), in some cases four-sided pyramids, with the same atomic structure on all four facets.” The observed length of the huts was very large (up to 1000 A), while their widths did not exceed 200 A (the aspect ratio often reached 10). The question was raised for the first time in that article as to what caused the elongation and specific base orientation of the huts. Unfortunately, no convincing answer to this question has been proposed thus far.

It should be noted that since issuing the article by Mo et al. [16], pyramidal and ‘elongated’ clusters have always been considered in the literature as structurally identical, differing only in their base aspect ratio [21–27]. The only argument for this assumption — the identity of faceting — does not seem to us as being very solid. Stress relaxation via the formation of \([105]\)-faceted structures, such as islands or pits [21], appears to be energetically favorable and the structures are more stable compared to those with different faceting [39]. This means that faceting itself cannot be considered as the only sign of belonging to some specific group of morphologically identical clusters. The total atomic structure of clusters is defined by both the structure of their facets and the configuration of their apexes. If the latter are different, the clusters should be regarded as members of different species. A sound counterargument to the assumption about cluster identity, which is usually disregarded, was given, for instance, in Ref. [40]: ‘elongated’ clusters completely disappeared from arrays after annealing at 550°C for 600 s, whereas pyramids and domes remained. Moreover, a commonly adopted pathway of dome cluster formation is as follows: some ‘prepyramid’ (we have never observed such formations in our low-temperature MBE experiments) \( \rightarrow \) pyramid \( \rightarrow \) dome [17, 41]. Elongated huts have never been met on this pathway. This led us to suggest that elongated huts\(^5\) differ from pyramids not only by the shape but mainly by their atomic structure and, which is

\(^5\) Hereinafter, we shall, as a rule, refer to them as wedges, wedge-like, or wedge-shaped clusters.
more important, by the genesis and role in array development.

Nevertheless, the question about the mechanism of hut elongation was asked and required answering. Before long, the simplest, and at first glance most plausible, scenario was proposed and immediately adopted by the scientific community. According to this scenario, wedges arise by elongation of a pyramid due to the growth of one of its \{105\} facets \[21–27\]. This hypothesis would explain everything unless the necessity to explain why the symmetry of the pyramid is violated. Due to its shape, the square-based regular pyramid seems to be stable enough, at least unless some exterior anisotropic agent affects it removing the degeneracy of its facets. Otherwise it is unclear why one facet gathers arriving Ge atoms to the detriment of the remaining three (or at least two) equivalent planes.

The next often observed miracle is the formation of two closely neighboring clusters separated by only a few nanometers, one of which is a pyramid, while the other is a wedge. This means that the acting agent violating the system symmetry is localized within a few nanometers around the growing wedge and does not affect the adjacent growing pyramid. At high enough coverages, such clusters often start to coalesce—this, in turn, means that the agent that previously violated the symmetry of one cluster in the pair has reached the growing pyramid but does not affect it now!

It may be, however, that the symmetry is violated only once, at the beginning of pyramid elongation, then the agent stops acting and the cluster grows on triangular facets faster than on trapezoid ones. It should now be concluded that, as distinct from the case of annealing \[40\], pyramids are much less stable than wedges in the process of Ge deposition, perhaps due to supersaturation of Ge atoms on the surface. This conclusion seems to be in agreement with experimental data, although the nature of the fluctuating agent remains obscure.

Goldfarb et al. \[26\] proposed that an asymmetry of the stress field due to the presence of \{105\}-faceted pits may result in cluster anisotropic lateral growth because of so-called equilibrium-driven elongation. They revealed that in a thick hydrogen-rich Ge wetting layer (between 7.7 and 8.3 ML), which is formed in the process of gas-source-molecular-beam-epitaxy (GS-MBE) \[6\] growth of Ge on Si(001) from GeH\(_4\), the strain at 690°C is relieved by formation of \{105\}-faceted pits rather than islands \[21\], which is in agreement with the earlier conclusion reached by Tersoff and LeGoues \[37\], according to which pits always have a lower energy than islands of the same shape and equal size. (Goldfarb et al. \[21\] also observed that at 620°C—and hence in a thinner wetting layer—only hut clusters arose.) Further, with the increase in Ge coverage of the surface, when the capability of pits to release the stress is exhausted, hut clusters nucleate in the vicinity of pits \[21, 22, 25\]. Acknowledging the model by Jesson et al. \[27\], which explains the instability of the hut cluster shapes (read ‘elongation’) by nucleation and growth on the facets, as more common, the authors of Ref. \[21\] illustrate the elongation process by an example of cluster coalescence. We agree with Goldfarb and coworkers that such an event sometimes happens and will present below picturesque evidence of it. However, this explanation of the cluster elongation phenomenon, certainly, can by no means be accepted as universal.\(^6\)

As mentioned above, investigations by Goldfarb and co-authors \[26\] eventually gave a weighty argument in support of so-called equilibrium-driven elongation or, in other words, elongation governed by energy minimization. They considered the evolution model developed by Tersoff and Tromp \[20\] for conditions when an isolated faceted strained Stranski–Krastanow island formed on a wetting layer grows more slowly in height than in the lateral direction. It may be shown in this case by minimizing the total energy per unit volume that there is a critical size \((x_0)\) of a pyramidal cluster, and the cluster grows isotropically up to some point \((e_{0})\), then anisotropic elongation in one direction starts. The authors of Ref. \[26\] indicated the discrepancy in estimates made on the basis of the model built by Tersoff and Tromp \[20\] with experimental evidences. The model predicts \(e_{0} \approx 100 \text{ nm} \ (x_0 \approx 37 \text{ nm})\), whereas in the opinion of the authors of Ref. \[26\] anisotropic elongation of huts starts at a nucleus size of about 5—8 nm \[21\], if at all.

Another model of equilibrium-driven growth was proposed by Li, Liu, and Lagally for a two-dimensional rectangular island \[42, 26\]. The model is as follows: let the dimensions of the island sides be \(s\) and \(l\). Until its size is less than some critical value, the total energy is a minimum for \(s = l\). For a greater island, its square shape becomes unstable because of the strain: the island begins elongation in one of two degenerate orthogonal directions until the total energy reaches a minimum at some value of arctan \((s/l) = \pi/4 \pm \Delta\). To explain the anisotropic elongation, the authors of Ref. \[42\] had to introduce anisotropy of edge energies, so as to violate the symmetry of a square. According to this model, the island grows along the direction of the lower edge free energy, in which both strain and edge energies are at a minimum. The advantage of this construction is that for strong anisotropy the elongation occurs at any size of island. Goldfarb and co-authors extend this model to the case of a 3D faceted island with slowly growing height \[26\]. According to them, when the island elongates its perimeter grows faster than the area of the strained base, resulting in more effective relaxation. Nevertheless, the origin of anisotropy remains an issue in this model.

As mentioned above, however, an explanation was proposed in the same letter \[26\]. It was observed that huts interact with parent pits as well as with different adjacent ones, their lengthwise growth often starting at one pit and ending at the other. Sometimes the elongation is finished when a hut grows along the pit boundary and reaches its corner. The effect of the pit interaction with hut was analyzed by the finite element technique and found to cause hut elongation. The pit dimensions used in calculations were 10 × 10 nm, and the wetting layer was as thick as 2 nm. The inference was made in the article that the following hut evolution scheme takes place: as soon as a stable critical nucleus is formed \[25\], it grows in the energetically favorable direction along its mutual boundary with the parent pit until it reaches the pit corner or attains an equilibrium \(s/l\) ratio. This finishes the first phase of cluster elongation. The second

\(^6\) GS-MBE much more resembles chemical vapor deposition (CVD) than the solid-source ultrahigh vacuum (UHV) MBE in which an atomic beam of Ge supplies the growing layer with Ge atoms rather than a flux of GeH\(_4\) or GeH\(_6\).

\(^7\) Note, by the way, that Goldfarb and co-authors were the first to report the STM observation of incomplete trapezoid facets of the wedges in Ref. \[22\]. Incomplete triangular facets have not been observed up to now, however.
phase implies that the cluster continues the equilibrium growth in the perpendicular direction (from the pit) until either an equilibrium $\frac{s}{h}$ ratio is attained or impingement on a pit occurs.

These observations and reasoning, which may be correct for the particular case of GS-MBE, were extended to all deposition methods. We have solid objections. First of all, only relatively rarefied arrays on very thick hydrogen-rich wetting layers were grown by GS-MBE and investigated in Ref. [26]. Dense arrays obtained by UHV-MBE grow in an absolutely different manner (Figs 1 and 2). Second, as is seen from Figs 1 and 2, no pits are available in the wetting layer grown by UHV-MBE at moderate temperatures. Third, according to our observations, which are as a rule carried out with atomic resolution, neither pits nor steps are required for cluster nucleation and growth by UHV-MBE (Figs 1 and 2; see also Ref. [43]). And finally, the above-described tricky evolutions of huts would leave their imprints on the cluster shape and structure, which would be seen with a good microscope. Unfortunately, this is not the case and no marks of the above evolutions are seen even at atomic resolution.

A model competing with those of equilibrium-driven elongation is usually referred to as kinetically driven elongation [23, 24, 44]. In this model, one of the pyramid facets begins to grow randomly due to a fluctuation. Coverage by a germanium monolayer decreases the likelihood of the next attachment of a Ge adatom to newly appeared trapezoid facets, whereas the probability of attachment to triangular ones remains unaltered. Due to an increase in the areas of the trapezoid facets and, consequently, in the height of the barrier for adatom attachment, further elongation in the randomly chosen (100) direction goes on with a decreasing probability of Ge atom attachment to the trapezoid facets and hence with a decreasing rate of in-height growth in comparison with the rate of the longitudinal one. Unfortunately, this model appeared to disagree with experiments. Observations by Goldfarb et al. [26] indicate that in the case of interaction of pits and huts, the two (100) directions of elongation are not equally probable as follows from the model of kinetically driven elongation. Our elaborated observations, which are presented below in Section 3.1, do not support this model, either.

Summarizing this brief historical review, we would like to emphasize the following. To date, the main milestones on the pathway of the Stranski–Krastanow growth of Ge film on Si(001) at moderate temperatures ($\leq 550^\circ$C) from the pure silicon surface to the appearance of hut clusters are recognized as consecutive steps of stress relief. Schematically, they are the following: $(2 \times 1) \rightarrow (2 \times n) \rightarrow (M \times N) \rightarrow$ huts. However, despite the efforts made, there is no clarity on the issues of hut nucleation and further transformation. The process of hut nucleation is appealing for detailed experimental investigation in different growth conditions by instruments assuring atomic resolution and using different deposition methods.

The formation processes and mechanisms of the longitudinal growth of huts have not been understood thus far. The elongation of pyramidal clusters has never been unambiguously observed in experiments. Likewise, the better of the two theoretical models explaining the elongation of wedges has not been chosen yet.

Issues of the evolution of cluster arrays during Ge deposition have been by-passed by researchers, too. There has been no systematic investigation of the stages of this very important and complicated process presented in the literature. The final phase of the array growth cycle—growth at high coverages and transition to the 2D mode—has never been the focus of investigations. So, we can conclude now that despite a widely adopted standpoint, investigations on the discussed problem are still far from completion.

The above analysis of the available literature poses a number of new questions, the most obvious of which are as follows. First, it is unclear whether the structures of pyramids and wedges are identical and whether they belong to the same morphologically uniform class. Second, does the structures of pyramids and wedges coincide at the moment of nucleation or have they got different structures already at the stage of emergence, and, consequently, different nuclei? Third, are their roles the same in the array formation and evolution? Fourth, what are the driving forces of their evolution during array growth? Why are the results of evolution different for pyramids and wedges? And last, is it possible to control cluster evolution in such a manner as to obtain uniform and defectless dense arrays suitable for industrial applications?

In this paper, we shall study some of the above issues, focusing mainly on the morphology of hut clusters and partly on the growth cycle of densely packed arrays.

2. Experimental

Experiments were made using an integrated ultrahigh vacuum system based on the Riber EVA 32 molecular beam epitaxy chamber coupled through a transfer line with a scanning tunneling microscope (STM) GPI-300 [45]. This setup enables the STM study of samples with atomic resolution at any stage of Si surface cleaning and MBE growth. The samples can be consecutively moved into the STM chamber for analysis and back into the MBE chamber for further processing, never leaving the ambient UHV and their surfaces stay atomically clean throughout the overall experiment.

The procedure for preparing the samples was as follows: initial substrates were $8 \times 8$-mm squares cut from B-doped CZ Si(100) wafers (p-type, $\rho = 12 \Omega$ cm). After washing and chemical treatment following a standard procedure described elsewhere (see, e.g., Refs [1, 46]), the silicon substrates mounted on a molybdenum STM holder and clamped with tantalum fasteners were loaded into an airlock and transferred to the preliminary annealing chamber, where they were outgassed at a temperature of around 565$^\circ$C and a pressure of about $5 \times 10^{-9}$ Torr for about 24 h. After that, the substrates were moved for final treatment into the MBE chamber evacuated down to about $10^{-11}$ Torr. There were two stages of annealing in the process of substrate heating in the MBE chamber—at approximately 600$^\circ$C for about 5 min, and at approximately 800$^\circ$C for about 3 min (Fig. 3). The final annealing at a temperature greater than 900$^\circ$C was carried out for nearly 2.5 min, with the maximum temperature being about 925$^\circ$C ($\approx 1.5$ min). Then, the temperature was rapidly lowered to about 750$^\circ$C. The rate of further cooling was around 0.4$^\circ$C s$^{-1}$. The pressure in the MBE chamber rose to nearly $2 \times 10^{-9}$ Torr during the process.

The surfaces of the silicon substrates were completely cleaned from the oxide film as a result of this treatment. The high-order $(8 \times n)$ surface reconstruction described in
Ref. [47] was always revealed by the STM on the deoxidized substrates, whereas the reflected high-energy electron diffraction (RHEED) patterns obtained from the same cleaned surfaces always corresponded to either (2 × 1) or (4 × 4) surface reconstruction [48]. This observation is in good agreement with the model put forward by us in Ref. [47], as well as with the generally accepted opinion based on RHEED measurements that the (2 × 1) reconstruction is formed on the Si(001) surface due to deoxidization in an MBE chamber.

Germanium was deposited directly onto the cleaned silicon surface from the source by electron beam evaporation. The rate of Ge deposition was about 0.15 Å s⁻¹, and \( h_{Ge} \) was varied from 6 to 14 Å for different samples. The deposition rate and the effective Ge film thickness \( h_{Ge} \) were measured by an XTC film thickness monitor graduated in advance, with the quartz sensor installed in the MBE chamber. The substrate temperature \( T_{Gr} \) was 360 °C or 530 °C during the process. The pressure in the MBE chamber did not exceed \( 10^{-9} \) Torr during Ge deposition. The rate of sample cooling to room temperature equaled approximately 0.4 °C s⁻¹ after the deposition.

The samples were heated by Ta radiators from the rear side in both preliminary annealing and MBE chambers. The temperature was monitored with chromel-alumel and tungsten–rhenium thermocouples in the preliminary annealing and MBE chambers, respectively. Thermocouples were mounted in a vacuum near the rear side of the samples and in situ graduated beforehand against the IMPAC IS 12-Si pyrometer, which measured the sample temperature through chamber windows with an accuracy of ±(0.3 % \( T \) °C + 1 °C).

The atmospheric composition in the MBE chamber was monitored using the SRS RGA-200 residual gas analyzer with a quadrupole mass spectrometer before and during the process.

After Ge deposition and cooling, the prepared samples were moved for analysis into the STM chamber, in which the pressure did not exceed \( 10^{-10} \) Torr. The STM tip was ex situ made of tungsten wire and cleaned by ion bombardment [49] in a special UHV chamber connected to the STM chamber. Images were obtained in the constant tunneling current mode at room temperature. The STM tip was zero-biased, while the sample was positively or negatively biased for mapping empty or filled surface states.

WXM software [50] was utilized for processing the STM images.

### 3. Classification

#### 3.1 Main species of the clusters

**3.1.1 Pyramidal and wedge-shaped clusters.** As mentioned in Section 1, an array of self-assembled germanium hut clusters formed on an Si(001) surface consists of a set of morphologically different clusters. All the clusters have the edges of bases oriented along the (100) directions in common. Yet in spite of the apparent variety of the cluster forms (Figs 1 and 2), an analysis of the STM images gives evidence that only two main species of the hut clusters exist—those having square bases and the shape of regular pyramids, and others with rectangular bases which have the shape of wedges. We have already cited the paper by Mo et al. [16], in which both species of hut clusters were described for the first time, as well as a number of publications which investigated the details of their formation [20–26]. Unfortunately, their structure has not yet been clearly visualized and identified.

Let us dwell on the descriptions of each species of the clusters in more detail.

Typical high-resolution STM images of the pyramidal clusters in arrays with different growth parameters are displayed in Fig. 4. A regular shape of the clusters is clearly seen in the figures (a) and (b) presenting the arrays obtained at \( T_{Gr} = 530 ^\circ C, h_{Ge} = 11 \) Å, and \( T_{Gr} = 360 ^\circ C, h_{Ge} = 10 \) Å; a fine structure of the faces is resolved, as well as the (\( M \times N \)) patched structure of the wetting layer [32–35]. Lines of the solidifying steps are revealed for the first time on the incomplete cluster faces in the image (c) of the array obtained at \( T_{Gr} = 360 ^\circ C, h_{Ge} = 14 \) Å (one of them is marked by an arrow in the image). Figure 4d shows a vertical view of the small pyramid grown at \( T_{Gr} = 360 ^\circ C, h_{Ge} = 6 \) Å and having only a 5-monolayer height over the wetting layer. The fine structure of the pyramid vertex and edges, as well as the stepped structure of its \{105\} facets are resolved in detail in image (d). And lastly, the same structure as that seen in Fig. 4d on the pyramid vertex is clearly resolved in image (e) of the pyramid nucleus (1 monolayer high over the wetting layer) situated on a block of the Ge (\( M \times N \)) surface.⁹

Figure 5 demonstrates STM images of wedge-like clusters. Being hut clusters, they are bounded by \{105\} planes, i.e., the proportion of their heights to the base widths is 1:10. A distinctive feature of this species of clusters is that their base lengths are not connected with cluster heights and are rather random. To some extent, the base lengths of the wedge-like clusters depend on their nearest neighbors. Nevertheless, it is impossible to confidently point out the factors which affect the lengths of Ge wedges, based upon the available data. Reasoning from the results of the STM image analysis, it may only be asserted that their base length-to-width ratio is distributed randomly and rather uniformly in the interval from a little greater than 1 to more than 10.

As long as the factors determining the base lengths of the wedges remain unclear, it is unknown if there is the possibility of controlling the array growth parameters in such a way as to minimize the dispersion of the base lengths of the wedge-like clusters. Moreover, the factors governing the nucleation of either pyramidal or wedge-like clusters with very different lengths on the wetting layer are also vague. Speculations

---

*Note also that a similar configuration of the pyramid vertex can be discerned by an attentive observer in Fig. 4b, which presents an image of a ‘ripe’ pyramid in a well-developed array.*
about the role of the strain in the wetting layer explain nothing. It remains inexplicable why clusters with different shapes (or even symmetry) may arise in a very close vicinity to one another. Perhaps the nucleation process and consequently the strain field (its distribution and local symmetry) are controlled by the underlying Si surface at the Si/Ge interface — its reconstruction before deposition and defect (in particular, vacancy) distribution before and in the course of Ge deposition — especially in the case of low-temperature processes. This is one of the reasons to consider the surface pregrowth treatment as a key to the controllable Ge/Si heterostructure formation process. That is why in situ studies of the Si surface on the atomic scale immediately before Ge deposition, as well as investigations of its influence on the deposited Ge layer, should be considered as a task of high importance.

Figure 5c demonstrates the fine structure of ridges of two close wedge-like clusters \( T_{\text{p}} = 360^\circ\text{C}, h_{\text{Ge}} = 8\ \text{Å} \). It is interesting that one can discern the same configuration of the ridge in the first published image of the hut cluster (see Fig. 2 in Ref. [16]).

STM images of the fine structure of the vertices and the ridges, similar to those shown in Figs 4d, e and 5c, helped us propose structural diagrams of both species of the clusters \([43, 51]\). It is clearly seen from the images that the fine structures of apexes of the clusters are different. The features in the uppermost parallel rows on the ridges of the wedge-like clusters are shifted with respect to one another. (They are marked by the rows of shifted arrows in the STM image.) We interpret these features as Ge dimer pairs in accordance with the simple structural model of the hut cluster facets (the PD model) proposed by Mo et al. [16]. Similar features in the images of vertexes of the pyramids are gathered in the straight rows. The difference between the symmetry of the pyramid vertex and the symmetry of the elementary unit of the wedge ridge is distinctly evident when compared in Figs 4e and 5c.
As we have already mentioned in Section 1.1, the difference in the atomic structure prohibits shape transitions between the pyramidal and wedge-shaped hut clusters, which were intensively discussed in the literature [24–26]. In addition, particular nuclei should be sought for pyramidal and wedge-shaped clusters. The question also arises as to why two structurally different species of hut clusters arise on the wetting layer.

Figure 6 depicts cross section profiles of the adjacent wedge-like and pyramidal clusters presented in Fig. 4a. Both clusters are seen to have an equal ratio of the base width to the cluster height close to 10. The base sides of the pyramidal cluster and the base length of the wedge-like one are nearly equal, whereas the base width of the wedge-like cluster is about 1.6 times less than the sides of the pyramid base. It is a common rule which is not affected by the length of a particular wedge-like cluster: the pyramidal clusters are usually higher than the wedge-like ones if \( h_{Ge} \) is high enough (see also, e.g., images (a) to (c) in Fig. 4). This inference may be explained if the height limitation is assumed to occur for the wedge-like clusters at \( h_{Ge} \) greater than some value, and not to occur for the pyramids. We observed such a limitation for wedge-shaped clusters (see Section 3.2) and did not succeed in fixing the height limitation on the pyramidal ones.10

3.1.2 Density and fraction of each species of clusters in the arrays. It was revealed that wedge-like and pyramidal clusters differ not only in their atomic structure and geometrical shapes. Wedge-like clusters dominate in arrays formed at low temperatures, and their fraction grows with an increase in \( h_{Ge} \) (Fig. 7).

Figure 7a plots the dependence of the cluster density on \( h_{Ge} \) for different clusters in arrays grown at 360 °C and 530 °C. It is seen that for \( T_{Ge} = 360 \) °C the density of wedges rises, starting from \( D_w \approx 1.8 \times 10^{11} \) cm\(^{-2}\) at the beginning of the Ge three-dimensional growth (the estimate was obtained by data extrapolation to \( h_{Ge} = 5 \) Å) and reaches a maximum of \( \approx 5 \times 10^{11} \) cm\(^{-2}\) at \( h_{Ge} = 8 \) Å; the total density of clusters at this point, \( D_E \approx 6 \times 10^{11} \) cm\(^{-2}\), is also a maximum. Then, both \( D_w \) and \( D_E \) slowly decrease until the Ge two-dimensional growth starts at \( h_{Ge} = 14 \) Å and \( D_E \approx D_w \approx 2 \times 10^{11} \) cm\(^{-2}\) (the contribution of \( D_p \) of pyramids to \( D_E \) becomes negligible — around \( 3 \times 10^{10} \) cm\(^{-2}\) — at this value of \( h_{Ge} \)). The pyramid density exponentially drops at \( T_{Ge} = 360 \) °C as the value of \( h_{Ge} \) grows (\( D_p \approx 5 \times 10^{11} \exp [-2.0 \times 10^5 h_{Ge}] \), with \( h_{Ge} \) measured in centimeters). The maximum value of \( D_p \approx 1.8 \times 10^{11} \) cm\(^{-2}\) obtained from extrapolation to \( h_{Ge} = 5 \) Å coincides with the estimated initial value of \( D_{Ge} \).

For \( T_{Ge} = 530 \) °C, the total density of clusters exhibits the same trend as \( D_w \) for \( T_{Ge} = 360 \) °C: \( D_E \approx 7 \times 10^{11} \exp (-4.3 \times 10^5 h_{Ge} \text{[nm]}) \). The maximum (initial) value of \( D_E \) is estimated as \( 8 \times 10^{11} \) cm\(^{-2}\) by extrapolation to \( h_{Ge} = 5 \) Å.

The graphs of cluster fractions in the arrays versus \( h_{Ge} \), are presented in Fig. 7b. For \( T_{Ge} = 360 \) °C, portions of pyramids and wedges, being initially very close (~ 50% at \( h_{Ge} < 5 \) Å), rapidly become different as \( h_{Ge} \) rises. The content of pyramids monotonically falls. The fraction of wedge-like clusters amounts to approximately 57% at the early stage of array growth (\( h_{Ge} = 6 \) Å) and becomes 82% at \( h_{Ge} = 8 \) Å. Upon further growth of the array, the content of the wedges reaches saturation at the level of approximately 88% at \( h_{Ge} = 10 \) Å.

At moderate values of \( h_{Ge} \), the proportion of pyramids to wedges at \( T_{Ge} = 530 \) °C was found to be nearly the same as for \( T_{Ge} = 360 \) °C. The content of the pyramidal clusters in the array is about 20% at \( h_{Ge} = 8 \) and 10 Å.

The inference may be made from this observation that, contrary to what is intuitively expected from the symmetry considerations, the wedge-like shape of the clusters is energetically more advantageous than the pyramidal one, and the more advantageous, the more Ge atoms (and the more the number of atomic layers) constitute the cluster. The probability of nucleation appears to be close to 1/2 for both wedge-like and pyramidal clusters at the initial stage of array formation and low growth temperatures. Then, as the array grows, the formation of pyramids becomes hardly probable and most of them, which have already been formed, vanish, whereas the nucleation and further growth of wedges are being continued (see Fig. 1). Ge pyramids on the Si(001) surface appear to be less stable species than the wedges and, in accordance with the ‘bourgeois principle’ (‘the survival of the fittest’), they lose their substance in favor of the wedge-like clusters.

At higher temperatures, no nucleation of new clusters was observed in the process of array growth (see Fig. 2). The bourgeois principle decreases the cluster density in the arrays and increases their sizes despite the species they belong to.

It should be noted that the above analysis demonstrates that pyramidal and wedge-like clusters are really different objects which belong to different cluster species rather than varieties of the same structurally uniform species, as usually postulated in the literature [16, 23–26].

Notice also that at \( T_{Ge} = 360 \) °C and the Ge atom flux \( dh_{Ge}/dt = 0.15 \) Å s\(^{-1}\), the point \( h_{Ge} = 10 \) Å is distinguished. Not only does the fraction of pyramids saturate at this point but also the array as a whole has the most uniform sizes of the clusters composing it (see Fig. 1). We reached this conclusion not only on the basis of analysis of the STM images of the Ge/Si(001) arrays but also from data of Raman light scattering by Ge/Si heterostructures with different low-temperature arrays of Ge quantum dots [52, 53]. We refer to such arrays as optimal.

---

10 Perhaps there is no height limitation for pyramids and consequently they give rise to large clusters affecting the properties of Ge/Si(001) heterostructures and classified by us as one of the types of defects of arrays (see report [1] or the article on defects of arrays [19]).
A qualitative model accounting for the presence of a particular point at low-temperature array growth is simple; at low enough temperatures of array growth, the new Ge cluster nucleation competes with the growth process for earlier formed clusters. The height of the clusters (at least, the dominating wedge-like ones) is observed to be limited by some value governed by $T_{gr}$. At small $h_{Ge}$, Ge clusters are small enough and the distances between them are large enough compared to the Ge atom (or dimer) diffusion (migration) length on the surface for nucleation of new clusters on the Ge wetting layer in the space between the clusters (Fig. 1a, b). At $h_{Ge} \approx 10$ Å and the above-given $d_{h_{Ge}}/dt$ value, the equilibrium of parameters (cluster sizes and distances between them, diffusion length at a given temperature, Ge deposition rate, etc.) sets in, the rate of new cluster nucleation decreases, and the abundant Ge atoms are mainly used to grow the existing clusters (Fig. 1c). After the clusters reach their maximum height, and in spite of this, Ge atoms continue to form their facets. As soon as most of the clusters reach the height limit, nucleation of new clusters becomes energetically advantageous again and the nucleation rate rises. A second phase of clusters appears on the wetting layer and fills its entire free surface as $h_{Ge}$ is increased (Fig. 1d). A further increase in $h_{Ge}$ results in a two-dimensional growth mode. It is clear now why the array is the most homogeneous (optimal) at $T_{gr} \approx 360$ °C and $h_{Ge} \approx 10$ Å, whereas the dispersion of cluster sizes increases at higher and lower values of $h_{Ge}$ because of the small clusters included in the array. It is also clear that the optimal array will appear at a different value of $h_{Ge}$, when $T_{gr}$ or $d_{h_{Ge}}/dt$ takes another value.

In addition, some threshold value of $T_{gr}$ must exist, beyond which the cluster growth process always dominates and the nucleation of clusters, having happened once, will never be repeated. An example of such arrays formed at $T_{gr} \approx 530$ °C exceeding the threshold value is given in Fig. 2.

### 3.2 Derivative species of clusters

#### 3.2.1 Obelisks (truncated wedges with two ridges)

Except for the above-described main species of Ge hut clusters, different clusters are also formed on the Si(001) surface, which cannot be classified as independent species because they originate from wedge-like clusters but have specific shapes, particular formation mechanisms, and maybe peculiar properties, and hence should be defined as separate but derivative species. Figure 8 shows clusters related to one of the species of derivative clusters — truncated wedge-like clusters with two ridges or obelisk-shaped clusters ($T_{gr} = 360$ °C, $h_{Ge} = 14$ Å).
Cross section profiles of the Ge obelisk shown in Fig. 8a and taken along the short and long base sides are presented in Fig. 9. Although these clusters are huts and have a slope of the facets equal to $11:3$, the ratio of the cluster height to its base width is $0.06$.

Figure 10 presents STM images of Ge wedges with two ridges, explaining the mechanism of their formation. The double ridges are seen to arise as a result of building the trapezoid facets in the cluster growth process. Initially, the clusters have one ridge and are ordinary wedge-like clusters. Wedges with double ridges arise because a limit value of the cluster height exists at any temperature $T_{gr}$. If the maximum height is reached, further growth of the cluster always goes on by building its trapezoid facets and increasing its width.

Cross section profiles of the Ge obelisk shown in Fig. 8a and taken along the short and long base sides are presented in Fig. 9. Although these clusters are huts and have a slope of the facets equal to $\sim 11.3^\circ$, the ratio of the cluster height to its base width is $\sim 0.06$.

This species of cluster dominates in arrays at high values of $h_{Ge}$, which depend on the value of $T_{gr}$ (see Figs 1 and 2).

A unique illustration of the process of trapezoid facet growth is drawn in Fig. 10c. Several (from four to six) incomplete (001) terraces are seen near the bottoms of the clusters (the dimer pairs are distinctly resolved, with the arrows showing these new growing facets in the STM image). These incomplete facets are seen to repeat the shapes of the former faces on which they grow and which are also incomplete. The highest terraces start to grow before the lower ones terminate the completion process, while new facets nucleate before the old ones finish their growth. It is also observed that the new facets nucleate far from the corners adjoining the base sides. As a result, a complex structure of trapezoid facets is formed, which can be seen in the truncated wedges represented, for example, in Fig. 4c. We propose that the reader compare the above description with the theoretical expectations of Jesson et al. [23], which draw a different picture of facet growth. First of all, they consider a triangular face as a preferential site of new facet nucleation, explaining in such a way the elongation of clusters. Then, according to their model, the facets nucleate in the corners rather than somewhere else, and so forth.

We would like to remark that we have not succeeded in observing the growth of the triangular faces at temperatures as low as $360^\circ$C. Nevertheless, we did observe this process at $T_{gr} \approx 530^\circ$C. Figures 10d–f demonstrate this phenomenon. The peculiarities of the process are as follows: new facet formation takes place when the cluster has already reached its height limit and its additional trapezoid facets are well developed. The growing triangular facets are clearly observed on only one side of the clusters. The triangular facets can nucleate both far from the bottom corner (Fig. 10d) and close to the corner (Figs 10e, f). The growing (incomplete) facets replicate the shape of the initial facet, even if the latter is complex (comprised of intersecting triangles due to developed additional trapezoid facets: see Figs 10e, f in which the
the above-discussed significant elongation of wedge-shaped clusters. In Fig. 10c, the nearly square-based clusters are seen to form because of the widening of wedges with two ridges. They resemble truncated pyramids but actually preserve the structure of the wedge. Nearly square-based clusters are also seen in the upper-left corner of Fig. 10e and the upper-right corner of Fig. 10f. These clusters were formed from formerly wedge-like ones by the successive addition of new incomplete facets. Their facets have complex structure and their shapes are far from the shape of an ideal regular pyramid. Certainly, their structure remains that of a wedge. The genuine pyramid cluster revealed in the upper-right corner of Fig. 10d grows nearly uniformly on all four its triangular facets (compare also with Fig. 4c). We would also like to indicate the formation of serial incomplete facets resolved on the sides of both pyramidal and wedge-shaped clusters presented in Fig. 10d (even the dimer pairs are seen on the ‘parallel steps’). This process does transform the shape of the clusters and may create nearly square-based clusters from the wedges, as well as so-called ‘rectangular-based clusters’ from the pyramids (the latter process may be imagined, e.g., if a pyramid is closely surrounded by its neighbors from all sides except for one and has some room for elongation only in one direction). Of course, such transformed clusters are always ‘truncated’ and have a complex ‘stepped’ structure of successive incomplete facets and an atomic structure of apexes characteristic of their precursors, as seen in the STM images presented.

### 3.2.2 Accreted wedges.

In Fig. 11, the STM images of wedge-like Ge clusters accreted together are shown: accreted clusters formed clusters with two ridges ($T_{gr} = 360 \degree C, h_{Ge} = 8 \text{ Å}$); accreted clusters which gave rise to a Γ-like one ($T_{gr} = 360 \degree C, h_{Ge} = 10 \text{ Å}$), and clusters accreted approximately at half width, which formed an extended one with a zigzag on the ridge and kinks on lateral faces ($T_{gr} = 360 \degree C, h_{Ge} = 10 \text{ Å}$).

Like the obelisks, the accreted clusters cannot be classified as independent species because they also originate from wedge-like clusters. Nevertheless, they also have specific shapes and, probably, peculiar properties and, consequently, like obelisks should be separated in a special but derivative species. It was found from the analysis of STM images obtained at different stages of the array formation that the nuclei of such clusters are situated at a distance of $\leq 32 \text{ Å}$ from one another at the initial stage of array formation in the applied array growth conditions.

From the physical viewpoint, these clusters are obviously single and integral objects, the properties of which may be different from the properties of the usual wedge-like clusters. Their influence on the properties of arrays as a whole awaits further investigation.

The coalescence of clusters is also seen in Fig. 4c. The truncated wedges are merged with or even completely absorbed by growing pyramids (the pyramids are always greater than the wedges). Such a process usually proceeds at high values of $h_{Ge}$ just before the onset of a two-dimensional growth.

### 4. Conclusion

Summarizing the above, we would like to emphasize the central ideas of the paper.

Morphological investigations and classification of Ge hut clusters forming arrays of quantum dots on an Si(001) surface at low temperatures in the process of ultrahigh vacuum molecular beam epitaxy have been carried out using *in situ* scanning tunneling microscopy. The study reported in this paper was made in view of the necessity to controllably produce highly uniform and very dense arrays of Ge quantum dots at low temperatures in a process compatible with the CMOS one. Although this task is still far from being solved, an important step has been made in understanding the object properties to be controlled.

The Ge/Si(001) system appeared to be much more sophisticated than had seemed to most researchers, and the knowledge about it, which is present in the literature, now appears to be very deficient and sometimes incorrect. This seems to be the main cause of failure in the last two decades to develop electronic or photonic devices based on ensembles of Ge quantum dots on an Si(001) surface.

Analysis of high-quality STM images which can be obtained only using an integrated high-resolution UHV STM-MBE instrument allowed us to introduce a new
classification of germanium hut clusters formed on the Si(001) surface. The hut clusters were found to be subdivided into four species, two of which are basic and structurally different — wedge-like and pyramidal clusters — and the rest of which are derivative — the obelisk-shaped and accreted wedge-shaped clusters. The conclusion was made that shape transitions between pyramids and wedges are prohibited. The growing trapezoid and triangular cluster facets have been visualized, and the peculiarities of facet completion have been described. It was shown that the growth of incomplete facets results in a complex structure of the growing hut clusters.

It was also shown that the uniformity of arrays is governed by the lengths of the wedge-like clusters. This parameter is hardly controllable, as distinct from cluster width which is linked to the cluster height and hence is much more predictable. The cluster lengths are now absolutely unpredictable. Moreover, the origin determining their values is unknown at present. This difficulty requires extensive investigations and intensive efforts to be overcome, but it is worthwhile doing because both stochastic (disordered) and predictable patterns of cluster growth are observed upon completion of the initial stage of array formation.

The growing trapezoid and triangular cluster facets have been visualized, and the peculiarities of facet completion have been described. It was shown that the growth of incomplete facets results in a complex structure of the growing hut clusters.

The authors appreciate the Science and Innovations Agency of the Russian Federation Ministry of Education and Science for funding this research under the State Contract No. 02.513.11.3130.

References

1. Yuryev V A et al. “Razrabotka fiziko-tekhnologicheskikh osnov upravljaemogo formirovaniya massivov platonopakovanymnykh nanoklasterov Ge na povr’chnosti kremnya Si(001) metodom sverkhvysokov’ukhovnikov molekul’yarno-luchevoi epitaksi”, shift 2007-3-13-25-01-303 (“Development of physical and technological basis of the controllable formation of densely packed Ge nanocluster arrays on the silicon (001) surface by means of ultrahigh vacuum molecular beam epitaxy”), Report on Research Project 2007-3-13-25-01-303 of the Science and Innovations Agency of the Russian Federation (Moscow: A M Prokhorov General Physics Institute of the Russian Academy of Sciences, 2007), VNITC No. 0220.0 802501

2. Pchelyakov O P et al. Fiz. Tekh. Poluprovodn. 34 1281 (2000) [Semicond. 34 1229 (2000)]

3. Brunner K Rep. Prog. Phys. 65 27 (2002)

4. Wang K L, Tong S, Kim H J Mater. Sci. Semicond. Process. 8 389 (2005)

5. Smagina J V et al. Zh. Eksp. Teor. Fiz. 133 593 (2008) [JETP 106 517 (2008)]