STM image interpretation: how theory influences experimental results

Karel Majer

Abstract

Scanning tunneling microscopy (STM) is generally valued for its ability to provide an atomically resolved direct-space image of a surface, moreover on a very simple basis – an exponential decay of the tunneling current with distance. However, it is seldom considered that the process of STM imaging is far more complicated and includes stages which we cannot control. An interpretation of the image is nearly impossible without theory.

In this essay I focus on the role of theoretical tools in the interpretation of my experiments with Al and Sn on Si(100). They are DFT calculations, which were used to determine atomic structures of such systems and simulate surface migration of adatoms, and theoretical assumptions which are used in our STM experiment. I will show that experimental results are to a certain level determined by those assumptions, because we do not treat STM image as it appears to us, but we choose what the relevant information is.

1 Introduction

In his short story *Tlön, Uqbar, Orbis Tertius* [1] Jorge Luis Borges describes a fictitious, artificial world which is created by a community of great minds – complete geography, history, culture, linguistics, philosophy and science of Tlön is invented and recorded in a vast encyclopedia. What at first seemed as a forgotten dream of a non-existent middle-eastern country then breaks in to our world. People are charmed by Tlön. They are interested more in its history than in their own, sciences start to describe that imaginary world, traditional languages disappear and Earth is finally converted into Tlön. As Borges says, men are attracted to a world that is governed by laws created by men instead of the divine – because we can ultimately find sense in human laws, but never fully understand divine ones. I tell this story, because the process of shaping reality in accordance to one’s expectation and desire is something common in physics, yet we seldom realize that.

Perhaps the best example would be the almost mystical hunt for elementary particles – I personally believe that quarks and Higgs boson were rather *created* by the search for them than discovered. However, I have no ambition of going deep into
this topic. Let me focus on my own experiment with heteroepitaxy of aluminium on Si(100) instead. We will look at interpretation of STM images and try to find out, whether we are using theory only as a useful tool, or whether it carries presumptions which influence our results and may be dangerous if unrealized.

2 Theory of group III. and IV. metals on Si(100) heteroepitaxy

Although my thesis is somewhat obscurely named "Recording surface processes on silicon by means of STM", it is a part of an ongoing research of heterostructures of group III. and IV. metals on silicon surfaces Si(100) and in the future probably Si(111) as well. Exploring means of increase in scanning rate of STM and recording processes beyond the limits of traditional STM experiment are important parts of my work, however, they are not of interest here.

What is so good about group III. and IV. metals heteroepitaxy on Si(100)? The chains. Self-organized one-dimensional dimer chains spanning up to several nanometers [2, 3, 4]. They are quite simple model system for studying low-dimensional objects and their electronic properties and therefore we really like them.

Growth of those chains is allowed by anisotropy of Si(100) surface. The initially symmetric surface has unsaturated dangling bonds and reconstruction occurs after annealing. Surface atoms are bound into dimers with $\sigma$ and $\pi$ bonds. Dimers are ordered in rows along $[110]$ or $[1\bar{1}0]$ directions and they are perpendicular at edges of atomic terraces [5]. The Si(100)$2 \times 1$ is depicted on figure 1.

![Figure 1: Si(100)$2 \times 1$. Source [6].](image)

Atoms in dimers are not symmetric, but one (up-atom) is higher than the other (down-atom). Up and down atoms are ordered along the dimer row and in two adjacent dimers the up-atoms are on different sides. However, at room temperature
the dimers do rapid flip-flop motion – thermally activated flipping of dimers which is correlated along the row. The flip-flop motion causes the symmetric appearance of the dimers and effective $2 \times 1$ reconstruction. Below 200 K the system undergoes a phase transition and flip-flop motion stops. Asymmetric dimers are visible in STM and form typical zig-zag patterns. They can form two reconstructions: $p(2 \times 2)$ and $c(4 \times 2)$, the latter is generally considered the basic state [7].

Of many defects which can be found on Si(100) surface I will mention only the type C defect as it can influence metal chains growth. Type C defect is a dissociated water molecule adsorbed on two adjacent dimers in a row [8]. It may behave as a preferred nucleation site for some group III. metals (In, Ga) and influence island density and size [9, 10].

Adsorption sites of various metals on Si(100) were determined mainly using density functional theory (DFT). There has been some dissent over what the stable sites are and which one is the least energetic, however, the so-called $M$ position, where an adatom is bound by two silicon atoms on-top of a second-layer atom, is considered stable in most works [11, 12]. It is also the position in which metal atoms forming chains are bound. Symmetric points where Al adsorption sites may lie calculated by Albao et al. are depicted on figure 2.

![Figure 2: Possible adsorption sites of Al on Si(100). Green: symmetric points where adsorption sites may lie, blue and red: Si surface atoms (blue up-atoms, red down-atoms), brown: Si atoms in subsurface layers. Red line marks the primitive cell of $c(4 \times 2)$ reconstruction. Background: potential energy surface for Al adatom. Source [12].](image)

Surface migration of adatoms plays a key role in epitaxial growth. The anisotropy of Si(100) causes the migration to be anisotropic as well, with different rates along the surface dimer rows and across the rows. It appears that migration along the rows has lower activation energy, therefore it is favored and at sufficiently low tem-
Peratures migration may become one-dimensional [12].

When two adatoms meet on the surface, they can form a dimer. In this case they are bound in M sites. Each adatom is bound to two adjacent Si atoms and its partner in the dimer. In case of group III. metals all adatom bonds are saturated, group IV. metal adatoms are left with one dangling bond [13, 14]. Stability of dimers and chains generally depends on the metal.

Growth of metal chains has been described as surface polymerization reaction [15]. Metal dimer is bound by electrons previously forming a π bond in Si surface dimer. Two Si atoms next to the dimer (or at the end of a chain) are left with unsaturated bonds, thus a reactive adsorption site is found next to the dimer. There is a higher chance that adatom will rest in that site long enough to capture a migrating adatom and form a stable metal dimer. The reactive site moves to the new end of the chain and the chain grows in one direction. Adsorption sites along the chain are forbidden to adatoms, because silicon bonds are already used to bind chain atoms. Surface polymerization reaction and structure of dimer chain is depicted on figure 3. Chains grow at low coverages up to coverage of 0.5 monolayer, when all surface bonds are saturated.

![Figure 3: Surface polymerization reaction. Hatched Si atoms, black stable metal dimer and metal adatom in favored adsorption site, grey adatom completing the second dimer. Source [15].](image)

Metals we are interested in (i.e. Al, Ga, In, Sn) differ in structure and stability of dimer chains. Especially Al differs greatly from the others. An indicator of chains stability can be chain length histogram. By scaling the histogram according to a certain relationship [16] we obtain the scaling function. Scaling theory tells us that the shape of the function is determined only by kinetics of growth and that the same growth mechanism leads to the same scaling function. Several growth modes and their corresponding scaling functions were obtained using kinetic Monte Carlo (KMC) simulations. Two of them are of interest here:

- Monomodal scaling function (function with one maximum) was obtained considering partially irreversible growth – existence of a critical island size [16]. Each island bigger than this size is stable and adatoms cannot detach from it. We have observed this scaling function in Al chains length distribution.
Monotonically decreasing scaling function may be a result of three models of growth. First, a reversible growth almost in thermodynamic equilibrium leads to an exponentially decreasing function [17]. Second, an irreversible growth with strong surface anisotropy and presence of reactive surface defects which behave as favored nucleation sites. In extreme case with high defect concentration the nucleation is only heterogeneous and growth is governed by defect concentration [18]. Last, an irreversible growth with strong surface anisotropy and presence of areas near islands which are forbidden to migrating adatoms [19]. This scaling function was observed in In, Ga and Sn epitaxy on Si(100) [19, 20]. Probably more than one of the mentioned mechanisms are present in their growth.

There are other notable differences between Al and the other metals. One of them is formation of kinks in chains – the chain does not go straight in the kink, but is shifted along Si dimer row. Other metals form straight chains, however it has been observed that density of kinks increases at lower temperatures and we believe that at sufficiently low temperature all chains contain kinks.

The other important difference (and more interesting for this essay) lies in STM imaging of metal dimers. I have not covered appearance of the chains in STM at all yet, but let me say, that although In, Ga and Sn dimers appear more or less the same (there are some exceptions such as tilted Sn dimers [14]), Al, on the other hand, forms dimers with three distinct levels of brightness and their origin is a mystery to us.

I believe that all important topics concerning growth of group III. and IV. metal structures on Si(100) have been briefly covered. So, let us look at role of DFT in determination of adsorption sites and then STM images of those chains and their interpretation.

3 Sad DFT results

There has been a number of articles trying to find out Al adsorption sites on Si(100) from first principles using different computational models in framework of density functional theory [11, 12, 21, 22]. They have only one thing in common – each of them completely differs from every other.

Well, that was a bit exaggerated, but not very much. Actually there is consensus that the M position (depicted on figure 2) should be stable and that there probably is another stable position called pedestal (denoted p on figure 2). That’s about it. Other positions are usually mentioned only in some of those articles, whereas others find them metastable in the best case or do not talk about them at all. I have not figured out whether the associated adsorption energies are of any relevance, because if we compare different articles, they look like taken from a lottery ticket. The same applies for computed bond lengths and angles, if authors had ambition of computing detailed atomic structure.
I conclude from this experience, that although DFT computation is surely a powerful tool, one has to be very careful about trusting it. This means comparing different sources and particularly not taking any number too seriously.

4 Interpretation of STM images

The very first STM images of Al chains on Si(100) were obtained by Nogami et al. in early 90’s [2]. Even though chains lying across Si dimer rows were clearly distinguishable (figure 4), not much could be deduced from the images alone. In brief, conclusions looked liked this: if we denote $a$ the surface period of Si(100) ($a = 3.84\,\text{Å}$) then bright maxima in the chains are spaced $2a$ apart. As the coverage increases, chains become longer and closer to each other, but there is a minimal spacing between two chains which again equals $2a$. Note that these conclusions could be only made because the structure and STM appearance of bare Si(100)$2 \times 1$ surface was known previously.

![Figure 4: STM images of Al chains on Si(100) at different coverages.](image)

Based on these findings two model of chain structure were proposed which differed in orientation of Al dimers related to Si surface dimers. As there was no way to tell the right one from STM images, DFT-LDA was used [13, 15] and surprisingly indicated that the unfavored "parallel" model in which Al dimers are oriented as depicted on figure 3 is the right one.

Now we have seen that drawing physically relevant conclusions from STM images alone is quite a problem – and those were good looking images! Imagine an ordinary STM image full of noise, glitches, blurred objects and distortion by thermal drift and piezoceramic creep.

Facing such an image one has to realize he really doesn’t know what he is looking at. Of course, he knows at least something, but not from the image itself and he
necessarily idealizes the image. In my experiment with Al on Si(100) I knew what is on the image, because I knew that I put silicon sample inside, deposited Al from an evaporator and I knew what it should look like. If I were to study structural properties of the chains from my STM images, I would definitely not tell that Al chains grow perpendicular to Si dimer rows, because angles at the images were severely distorted. However, I already knew the DFT based structural models. This knowledge we are using to tell what is on the image consists of presumptions and should not be depended upon blindly – some unusual phenomena could then be overlooked.

My aim was to count density of chains and their lengths and create length histograms which are used for fitting model of growth by KMC simulations. To do this properly, every single dimer would have to be identified and counted, which would require using only images with superb resolution, lowering total number of chains counted and increasing statistical errors. Luckily we can count number of dimers in chains even with worse resolution (even when a chain appears as a thick blurred white line), because its number equals the number of Si dimer rows under the chain in the empty states image and that can be counted far more easily. But what about single dimers?

You would say, if it’s a bright point sitting just on one dimer row, count it as a single dimer. That would be nice, but a bright point sitting on one dimer row could be literally anything. C defects appear more or less the same as single dimers, if the resolution is not high enough, but at least they can be distinguished from them in filled states image, so a check is necessary. However, from experience I know that a number of similar bright objects remain and – based on a guess more than on anything else – some of them look strange. And here theory becomes near useless, because all computations work with pure and well-defined structures. I know how Al on Si(100) looks like, how C defect looks like, but for good reasons there are no simulated STM images, nor even LDOS simulations of random dirt which may exist in vacuum chamber adsorbed on Si(100) – and yet I would be very glad for them.

I mentioned the problem of three levels of brightness of Al dimers. This may be an example of fault coming from misuse of theory, which tells us that all Al dimers are the same and their structure is as described above. However, various levels of brightness in STM mean either various height or various electronic structure, none of which was proposed by theory. Intermixing of Al and Si atoms should not be favorable at temperatures below 600 K, but we clearly see that there is something rotten in some Al dimers, yet we don’t know what. So we usually derive from theory that all the dimers are the same and if we have to count them, we don’t make differences. But in the case we are wrong our counts would need adjustments and this would be an example of experimental results getting flawed by misused theoretical assumptions.


5 Conclusion

The message of the last section may be straightforward – if you don’t check you theoretical presumptions, you may get bad results. I believe that it’s more complicated. We are always using theoretical presumptions in STM experiment – what we see, what it should look like, what is physics and what is dirt and distortion. Danger comes when we are not aware of this. Our results may get affected in an unexpected way by some unknown phenomena or just by wrong interpretation of images. Therefore it is necessary to have on mind that we really don’t know what we are looking at in STM and all our knowledge comes \textit{a priori} and needs to be reexamined and all possible factors affecting the image have to be taken into account – stability of the tip, scanning speed, possible glitches and noise. Only then theory may help us to tell what is on images and derive useful data instead of building an appealing artificial world in which we would discover only what has been known to us.

References


[6] WWW Picture gallery based on in the Surface Structure Database (SSD, NIST Standard Reference Database 42) by P. R. Watson, M. A. Van Hove, K. Hermann. The pictures have been prepared from SSD output and postprocessed with BALSAC by K. Hermann.


