Doina Bejan

Experimental methods of surface physics

(five weeks course for the third year –undergraduate students)

Spring, 2017

1. Introduction

The study of solid surface phenomena is of great importance in physics since a solid sample is always in contact with other media via its surface. The existence of such an interface modifies, at least locally, the properties of the sample. The interaction with the outside world occurs through the surface. Therefore, surface physics finds applications in many technologies, for example, in heterogeneous catalysis, microelectronics, electrochemistry, corrosion, photography, biology.

The advent of quantum mechanics in 1920s produced a turning point in the development of surface physics. In 1923, Clinton Davisson and Charles Kunsman performed the first low energy electron diffraction experiment which proved the wave nature of quantum mechanical particles. Very rapidly, quantum mechanics was applied to investigate the electron structure of solids and the role played by the boundary conditions in the presence of a surface was raised. This led to new concepts such as surface states, surface double layers and provided a means to calculate, on a microscopic basis other quantities of physical interest. Most of the physical phenomena (such as Auger effect, diffraction of particles, field emission and field ionization) on which the modern experimental techniques of surface observation are based, were also discovered. However progress in understanding the surface physics had been hampered by severe problems of experimental reproducibility, which were due to the difficulty in obtaining structurally and chemically well characterized surfaces.

This problem was solved in the 1960s due to the appearance of ultra high vacuum technology which has led to the development of many experimental techniques as well as of chemical analysis on an atomic scale. At the same time, high speed digital computers became available allowing theoretical works to reach a degree of sophistication going far beyond the simple models developed in the previous period.

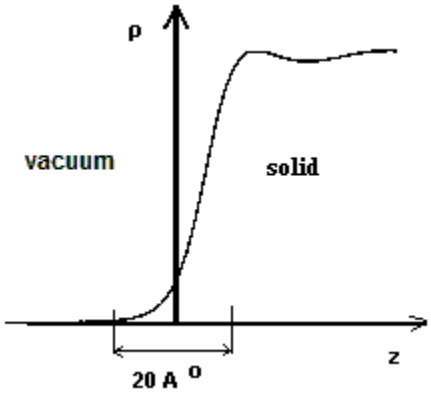

Fig. 1.1 Electronic charge density at surface

The region we are referring to as the surface is extending about 20Å around the last atomic plane, including the first three or four atomic layers. Beyond this region the electronic density almost vanishes on the vacuum side and has attained its bulk behavior on the solid side. We will deal only with crystal surfaces restricted to pure metal and semiconductor surfaces, clean or possibly in the presence of adsorbates. Some experimental techniques of surface characterization will be also presented: **FIM** (Field Ion Microscopy); **STM** (Scanning Tunneling Microscopy); LEED(Low energy electron diffraction) .

2. Atomic structure of surface

2.1 Crystal structure

The structure of all crystals can be described in terms of a lattice, with a group of atoms attached to every lattice point. The group of atoms is called the basis. When repeated in space it forms the crystal structure.

The lattice is defined by three fundamentals translation vectors $\vec{a},\vec{b},\vec{c}$ such that the atomic arrangement looks the same in every respect when viewed from the point $\vec{r}$ as when viewed from the point

$$\vec{r}' = \vec{r} + n_1\vec{a} + n_2\vec{b} + n_3\vec{c} \qquad (2.1)$$

where n_1, n_2, n_3 are arbitrary integers. The set of points $\vec{r}'$ defined by 2.1 for all n_1, n_2, n_3 defines a lattice. The lattice and the translation vectors $\vec{a},\vec{b},\vec{c}$ are said to be primitive if any two points $\vec{r}',\vec{r}$, from which the atomic arrangement looks the same, always satisfies 2.1 with a suitable choice of the integers n_1, n_2, n_3.

We often use the primitive translation vectors to define the crystal axes. They form three adjacent edges of a parallelepiped. More than one lattice is always possible for a given structure, and more than one set of axes is possible for a given lattice. The parallelepiped defined by the primitive vectors $\vec{a},\vec{b},\vec{c}$ is called the primitive cell. A primitive cell is a minimum-volume cell. The volume of the primitive cell is $V = \left|\vec{a}\times\vec{b}\cdot\vec{c}\right|$.

A lattice translation operation is defined as the displacement of a crystal by a crystal translation vector

$$\vec{T} = n_1\vec{a} + n_2\vec{b} + n_3\vec{c}\,. \qquad (2.2)$$

Any two lattice points are connected by a vector of this form. The primitive cell will fill all space by the repetition of suitable crystal translation operations.

A lattice is a regular periodic array of points in space. It is a mathematical abstraction. The crystal structure is formed when a basis of atoms is attached identically to every lattice point. The number of

atoms in the basis may be one, or N greater than one. The position of the center of an atom j of the basis, relative to the associated lattice point is

$$\vec{R}_j = x_j\vec{a} + y_j\vec{b} + z_j\vec{c} \ . \qquad (2.3)$$

We may arrange that $0 \le x_j, y_j, z_j \le 1$. The basis associated with a point of the primitive cell is called a primitive basis.

Crystal lattices can be mapped into themselves by the lattice translations $\vec{T}$ and by various other symmetry operations:

- The rotation about an axis that passes through a lattice point by an angle $\theta = 2\pi/n$ (n=1,2,3,4,6), n=5,7 are not possible at least for the perfectly periodic structures. The rotation axis are noted with *n* or C_n;
- Mirror reflections (*m* -from mirror plane or σ_v and σ_h) about a plane through a lattice point;
- The inversion operation, composed of a rotation of π, followed by reflection in a plane normal to the rotation axis; the total effect is to replace r by -r.

Above I used two kinds of notations: *n, m* or C_n, σ_v and σ_h. The latest notation is the **Schoenflies** (or **Schönflies**) **notation**, named after the German mathematician Arthur Moritz Schoenflies. It is one of two conventions commonly used to describe point groups. This notation is used in spectroscopy. The other convention is the Hermann–Mauguin notation, also known as the International notation. A point group in the Schoenflies convention is completely adequate to describe the symmetry of a molecule; this is sufficient for spectroscopy. The Hermann–Mauguin notation is used in crystallography.

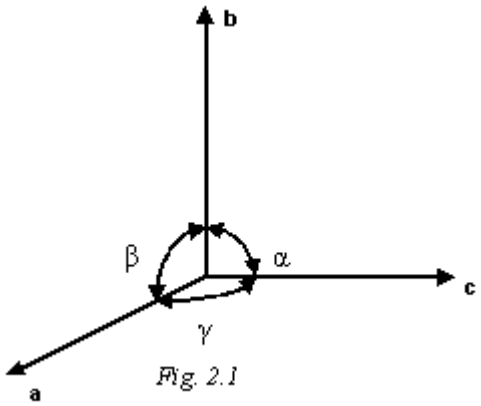

Fig. 2.1

There are 14 different lattice types called Bravais lattices (Table 2.1). These are grouped for convenience into systems classified according to seven types of cells, which are triclinic, monoclinic, orthorhombic, tetragonal, cubic, trigonal and hexagonal. The angles α, β, γ from Table 2.1 are represented in figure 2.1.

Most of the metals crystallize in the cubic system where there are three lattices: the simple cubic (SC) lattice, the body-centered cubic (bcc) lattice and the face-centered cubic (fcc) lattice.

The orientation of a crystal plane is determined by three points in the plane, provided they are not collinear. If each point lay on a different crystal axis, the plane could be specified by giving the

coordinates of the points in terms of the lattice constants a, b, c. For example, if the atoms that determinate the plane have coordinates (6,0,0), (0,2,0), (0,0,3), relative to the axes vectors, the plane may be specified by 6, 2, 3.

Table 2.1

System	Number of lattices	Symbol	Restrictions
Triclinic	1	P (primitive)	$a \neq b \neq c$ $\alpha \neq \beta \neq \gamma$
Monoclinic	2	P, C (centered basis)	$a \neq b \neq c$ $\alpha = \beta = 90° \neq \gamma$
Orthorhombic	4	P, C, I (body-centered, innenzentrierte), F (face-centered)	$a \neq b \neq c$ $\alpha = \beta = \gamma = 90°$
Tetragonal	2	P, I	$a = b \neq c$ $\alpha = \beta = \gamma = 90°$
Cubic	3	P (sc) I (bcc) F (fcc)	$a = b = c$ $\alpha = \beta = \gamma$
Trigonal	1	R (rhombohedra)	$a = b = c$ $\alpha = \beta = \gamma < 120°, \neq 90°$
Hexagonal	1	P	$a = b \neq c$ $\alpha = \beta = 90°$ $\gamma = 120°$

However, it turns out to be more useful for structure analysis to specify the orientation of a plane by the Miller indices (after William Hallowes Miller (1801-1880), british mineralogist and crystallographer) determined by the following rules:

- find the intercepts on the axes in terms of the lattice constants a, b, c , for example: p, q, r ;

- take the reciprocals of these numbers and then reduce to three integers having the same ratio, usually the smallest three integers h :k :l=1/p :1/q :1/r. The result enclosed in parentheses, (hkl), is called the index of the plane. For example, for the plane whose intercepts are 6, 2, 3, the smallest integers having the same ratio are (132).

The indices (hkl) may denote a single plane or a set of parallel planes. If a plane cuts on an axis on the negative side of the origin, the corresponding index is negative, indicated by a minus sign above the index $(h\bar{k}l)$. The cube faces of a cubic crystal are (100), (010), (001), $(\bar{1}00)$, $(0\bar{1}0)$ and $(00\bar{1})$.

The indices [hkl], of a direction in a crystal are the set of the smallest integers h, k, l that have the ratio of the components of a vector in the desired direction, referred to the axes. For example, the x axis of a cubic crystal is [100] and $-z$ is $[00\bar{1}]$. In cubic crystals the direction [hkl] is perpendicular to the plane (hkl), but this is not generally true in other crystal systems.

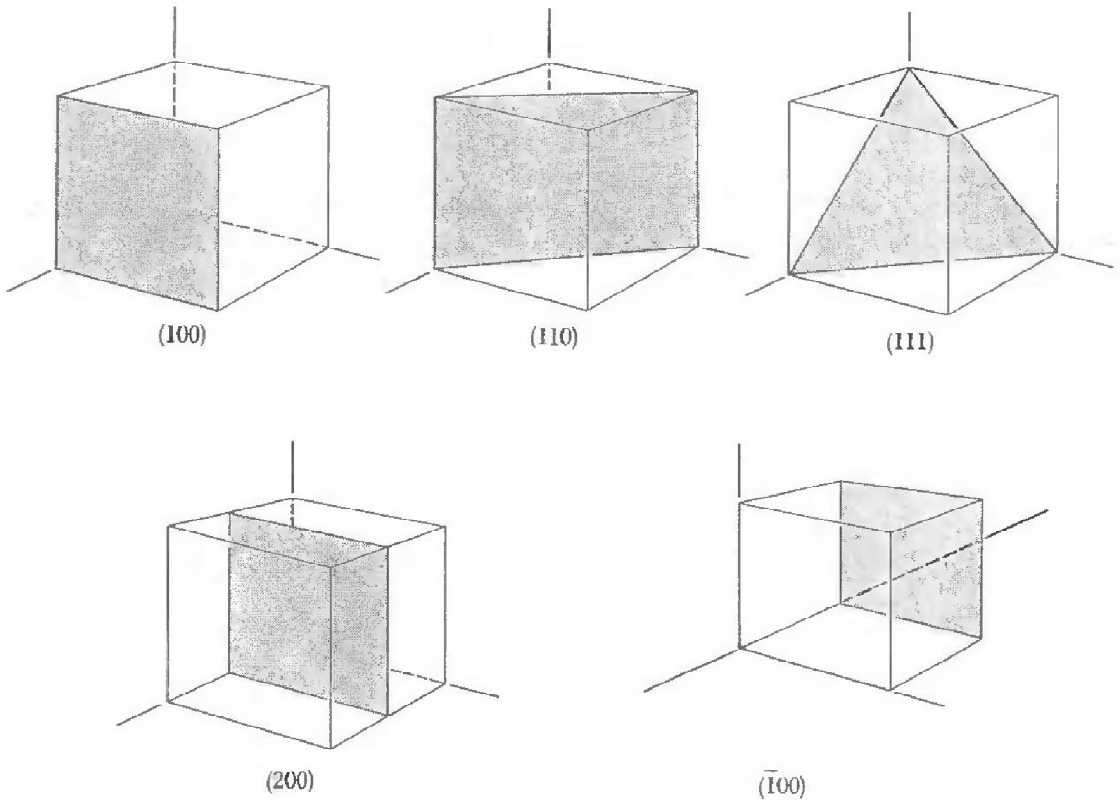

Fig. 2.2 The Miller indices of important planes in a cubic crystal. The plane (200) is parallel to (100) and $(00\bar{1})$.

2.2 Two-dimensional lattices

What happens when we cut a crystal following an atomic plane? The first thing one might guess is that cleavage of a crystal does not perturb the remaining material at all. That is, perhaps the arrangement of atoms is precisely the same as a planar termination of the bulk. However, this so called ideal surface appears to be the exception rather than the rule.

The complete translational invariance of a bulk crystal is destroyed by cleavage. At best, one retains periodicity in only two dimensions.

The surfaces are noted after the atomic plane of cleavage, for example Cu(100), Al(110) etc. The surface structure may be described by a primitive cell, the smallest cell that can be repeated for obtaining the whole surface. There are many ways of choosing the axes of the primitive cell as one can see in fig. 2.3. The vectors $\vec{a}$ and $\vec{b}$ that define the primitive cell are called primitive vectors.

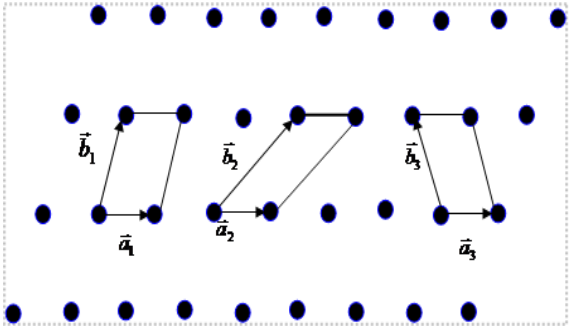

Fig. 2.3 Different ways of choosing the axes of the primitive cell. All cells are primitive and have the same area.

One of the symmetry transformations allowing us to generate two-dimensional periodic structure is translation:

$$\vec{T} = n_1\vec{a} + n_2\vec{b} \qquad (2.4)$$

The other symmetry transformations are:

- Rotational symmetry with an angle $\theta=2\pi/n$ (n=1,2,3,4,6);
- Mirror reflections m about a plane through a lattice point and perpendicular to the surface.

These different point operations yield five two-dimensional lattices, the five Bravais lattices drawn in fig. 2.4. If we now locate a set of atoms or molecules, called basis, at each lattice point, the symmetry may remain the same or be reduced. The total symmetry of the system is described by combining the Bravais lattices with the crystallographic point group of the basis.

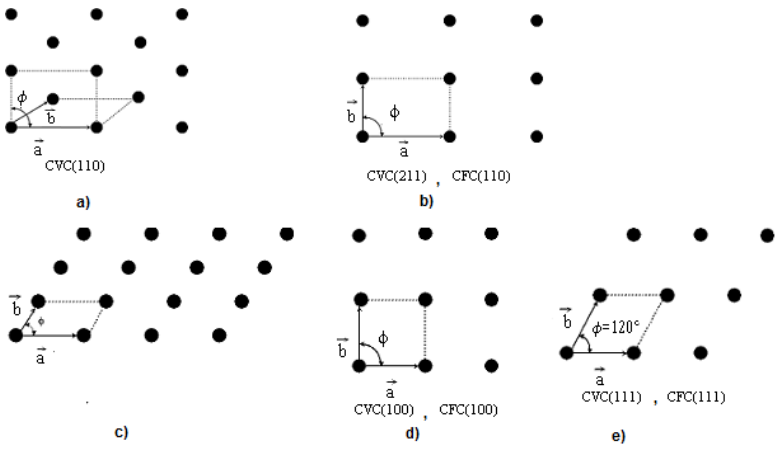

Fig. 2.4 The five Bravais lattices for surface: a) centered rectangular; b) rectangular; c) oblique; d) square; e) hexagonal.

The two-dimensional lattice being given, we can define Miller indices, h, k, to label crystal rows. The indices, [h, k], label the set of atomic rows parallel to the line which intersects the unit cell at *a/h* and *b/k*.

Table 2.2

Lattice	**Conventional cell**	**Properties**	**Symmetry of the point group (Hermann-Mauguin and Schoenflies notations)**
Oblique	Parallelogram	$a \neq b$, $\varphi \neq 90°$, arbitrar	2 (C_2)
Square	Square	$a = b$, $\varphi = 90$	4mm (C_{4v})
Hexagonal	Rhomb	$a = b$, $\varphi = 60°$	6mm (C_{6v})
Rectangular (primitive)	Rectangle	$a \neq b$, $\varphi = 90°$	2mm (C_{2v})
Centered rectangular	Rectangle	$a \neq b$, $\varphi = 90°$	2mm (C_{2v})

2.3 Relaxation, reconstruction, surstructures

At surface, the forces acting on the atoms are modified and, consequently the equilibrium structure of the semi-infinite crystal is not always one-half of the corresponding infinite crystal. Several atomic rearrangements are possible in the vicinity of the surface.

In metals, the structure and the interatomic distances in the first planes remain the same as in the bulk but there is a modification of the first interlayer spacings. This phenomenon is called normal relaxation. In most cases, the experiments show that the greater the number of broken bonds at the surface, the more the first interlayer spacing is reduced. For instance, the first interlayer spacing of Mo(100) is contracted by 12%. In addition to this normal relaxation, we can sometimes observe a uniform displacement of the first planes parallel to the surface. This is called a parallel or tangential relaxation and has been observed, for example on W(110) in the presence of hydrogen. When going deeper in the crystal these displacements are damped and often oscillatory.

The atomic structure of the first plane can be modified. This phenomenon is known as surface reconstruction which occurs, for instance, on W(100), Au(110) or Si(111).

Pure or composed semiconductors have covalent (tetrahedral) bonding, each atom being bound up with his four neighbors by a hybrid orbital sp^3. At surface, the bonds are cut and the „free" sp^3 orbitals become dangling orbitals in vacuum. In order to get rid of the nonbonding orbitals, the surface

reconstruction occurs. For example, Si(111) cut at the room temperature get a (2x1) reconstruction forming dimmers (see figure 2.5). It is a metastable state that, with the increase of the temperature, suffers an irreversible phase transition to the stable (7x7) surstructure.

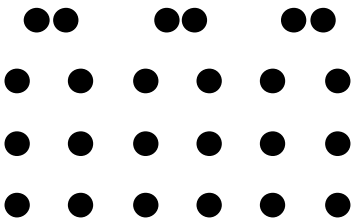

Fig. 2.5 Dimers formation at the Si(111) surface. Transversal view.

The surface may exhibits steps with various heights separating domains made of atomic planes in which point defects such as adatoms, advacancies, ledge and kinks can be detected.

The surstructures created by reconstruction of clean surfaces or by the presence of an ordered adsorbed layer need new notations to specify it. They give the dimensions, the nature and the orientation of the unit cell of the first plane relative to the underlying ones. When the primitive vectors of the surface unit cell are $\vec{a}' = n\vec{a}$ and $\vec{b}' = m\vec{b}$ the structure is labeled by $p(n \times m)$; *p* meaning primitive is sometimes omitted. This simple notation is convenient when the surface unit cell is in registry with the substrate unit cell. However it is not adequate when the surface unit cell is rotated with respect to the substrate unit cell or when the ratio of the dimensions of these cells is not an integer. For simple rotations, $(x \times y)R\theta$ means a surface structure which is obtained from the surface unit cell by a rotation of angle θ. In addition, we frequently observe surface structures which can be described as square or rectangular lattices having an atom in the center. Such a structure is denoted by $c(n \times m)$, *c* meaning centered.

2.4 Reciprocal lattice

The concept of a reciprocal lattice is important since similarly to the three dimensional case, it is directly observed in diffraction experiments. Moreover, it plays a central role in the propagation of waves in crystals.

Let us recall that in three dimensions the vectors of the reciprocal lattice are given by

$$\vec{G} = l_1\vec{A} + l_2\vec{B} + l_3\vec{C} \tag{2.5}$$

With l_1, l_2, l_3 integer numbers while the points of the direct lattice are given by $\vec{\rho} = n_1\vec{a} + n_2\vec{b} + n_3\vec{c}$. One can see that $\vec{G}\vec{\rho} = 2\pi m$ if the Laue relations for maximum diffraction intensity are satisfied:

$$\begin{aligned} &\vec{A}\vec{a} = 2\pi, \quad \vec{A}\vec{b} = 0, \quad \vec{A}\vec{c} = 0, \\ &\vec{B}\vec{a} = 0, \quad \vec{B}\vec{b} = 2\pi, \quad \vec{B}\vec{c} = 0, , \\ &\vec{C}\vec{a} = 0, \quad \vec{C}\vec{b} = 0, \quad \vec{C}\vec{c} = 2\pi \end{aligned} \tag{2.6}$$

And consequently

$$\vec{A} = 2\pi \frac{\vec{b} \times \vec{c}}{\vec{a} \cdot \vec{b} x \vec{c}}, \quad \vec{B} = 2\pi \frac{\vec{c} \times \vec{a}}{\vec{a} \cdot \vec{b} x \vec{c}}, \quad \vec{C} = 2\pi \frac{\vec{a} \times \vec{b}}{\vec{a} \cdot \vec{b} x \vec{c}} \tag{2.7}$$

represents the primitive vectors of the reciprocal lattice. They are orthogonal only if the vectors or the direct lattice are orthogonal. The vectors of the reciprocal lattice measure L^{-1}.

In two-dimensions these relationships become:

$$\vec{T} = n_1 \vec{a} + n_2 \vec{b} \tag{2.8}$$

$$\vec{G} = l_1 \vec{A} + l_2 \vec{B} \tag{2.9}$$

where $\vec{A}$ and $\vec{B}$ are related to $\vec{a}$, $\vec{b}$,by

$$\begin{aligned} &\vec{A}\vec{a} = 2\pi, \quad \vec{A}\vec{b} = 0, \\ &\vec{B}\vec{a} = 0, \quad \vec{B}\vec{b} = 2\pi. \end{aligned}$$

(2.10)

Setting $\vec{a} = (a_1, a_2, 0)$, $\vec{b} = (b_1, b_2, 0)$ and $\vec{c} = \hat{z} = (0,0,1)$ one obtains from (2.7)

$$\begin{aligned} \vec{A} &= (A_x, A_y) = \frac{2\pi}{a_x b_y - a_y b_x} (b_y, -b_x) \\ \vec{B} &= (B_x, B_y) = \frac{2\pi}{a_x b_y - a_y b_x} (-a_y, a_x) \end{aligned} \quad . \tag{2.11}$$

The parallelogram built on $\vec{A}$ and $\vec{B}$ defines the unit cell of the two-dimensional reciprocal lattice. The reciprocal cells corresponding to the five Bravais direct lattices are represented in figure 2.5.

The structure of the direct lattice at nanometric resolution is given by only three techniques:

- FIM (Field Ion Microscopy)

- STM (Scanning Tunneling Microscopy)
- AFM (Atomic Force Microscopy).

Many other techniques probe directly the reciprocal lattice and many others offer informations about the atomic structure like SEXAFS (Surface Extended X-ray Adsorbtion Fine Structure), PhD (Photoelectron Diffraction), Photon Stimulated Desorption, X-ray standing waves etc.

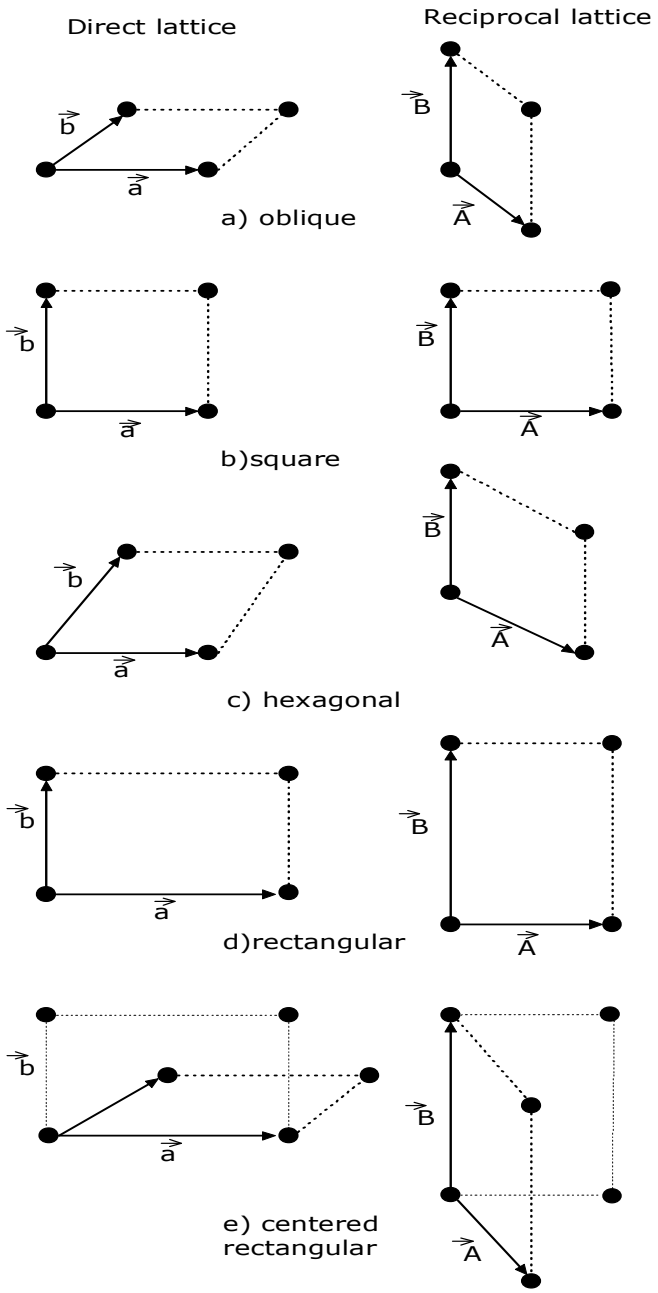

Fig. 2.6

3. Principle of Field Ion Microscope (FIM)

The **Field ion microscope (FIM)** was invented by Erwin E. Mueller in 1951 at the Pennsylvania State University. By FIM, man observed atoms for the first time. It was developed from its forerunner, the Field emission microscope, invented earlier in 1936, also by Mueller.

3.1 FEM characteristics

A Field Emission Microscope consists of a metallic sample in the form of a sharp tip (the cathode) and a conducting fluorescent screen (the anode) at a distance R from it, all enclosed in ultrahigh vacuum. The tip radius r used is typically of the order of 1000Å to 1 μm. It is composed of a metal with a high melting point, such as tungsten. The microscope screen shows a projection image of the distribution of current-density J across the emitter apex, with magnification approximately (R/r), typically 10^5 to 10^6. The sample is held at a large negative potential (1-10 kV) relative to the fluorescent screen. This produces an electric field near the tip apex of order 0.3-0.5 V/A°, which is high enough for field emission of electrons to take place. The field emitted electrons travel along the field lines and produce bright and dark patches on the fluorescent screen giving a one-to-one correspondence with the crystal planes of the hemispherical emitter.

The emission current varies strongly with the local work function. When the emitter surface is clean, this FEM image is characteristic of: (a) the material from which the emitter is made: (b) the orientation of the material relative to the needle axis; and (c) to some extent, the shape of the emitter end form. In the FEM image, dark areas correspond to regions where the local work function φ is relatively high and/or the local barrier field E_{el} is relatively low, so the current-density J is relatively low; the light areas correspond to regions where φ is relatively low and/or E_{el} is relatively high, so the current-density J is relatively high.

The spatial resolution of this technique is of the order of 2 nm and is inextricably limited by the momentum of the emitted electrons parallel to the tip surface, and, to a lesser extent by their de Broglie wavelength, neither of which could be controlled.

When in the spring days of quantum mechanics Gamow explained the radioactive alpha decay as a tunneling effect, field electron emission from metals was soon recognized by Fowler and Nordheim as another example of barrier penetration. In the same time, Oppenheimer suggested that hydrogen in an electric filed has a finite probability to be ionized by tunneling. So, the field electron emission is a paradigm example of what physicists mean by tunneling. Unfortunately, it is also a paradigm example of the mathematical difficulties that can arise. Simple solvable models of the tunneling barrier lead to equations (including the original 1928 Fowler–Nordheim-type equation- the Schrodinger equation for an exact triangular barrier (see eq. 3.1)) that get predictions of emission current density too low by a factor of 100 or more.

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2}\psi(z)-eE_{el}z\psi(z)=E\psi(z) \tag{3.1}$$

If one inserts a more realistic barrier model into the simplest form of the Schrödinger equation, then it is mathematically impossible to solve this equation exactly in terms of the usual functions of mathematical physics, or in any simple way. To get even an approximate solution, it is necessary to use special approximate methods known in physics as "semi-classical" methods, or numerical methods.

The FEM was an early observational tool of surface science. For example, in the 1960s, FEM results contributed significantly to discussions on heterogeneous catalysis. FEM has also been used for studies of surface-atom diffusion. However, FEM has now been almost completely superseded by newer surface-science techniques.

3.2 FIM characteristics

In FIM the essential features are the same, but this time the specimen tip is of a smaller radius (usually 100-300 Å, up to 2000 Å) and is kept at a high, positive potential (about 3-30 kV) to produce a field of the order 5 V/A°. The sharp needle is mounted on an electrically insulated stage that is cooled to cryogenic temperatures (20 to 100K) in an ultrahigh vacuum chamber (Fig.3.1). The field ion image of the specimen is formed on a microchannel plate and phosphor screen assembly that is positioned approximately 50 mm in front of the tip. To produce a field ion image, controlled amounts of image gas are admitted into the vacuum system. The type of image gas used depends on the material under investigation; common images gases are neon, helium, hydrogen and argon. The pressure of the gas is of few 10^{-3} Torr, low enough to let the ions travel to the screen, without disturbing collisions.

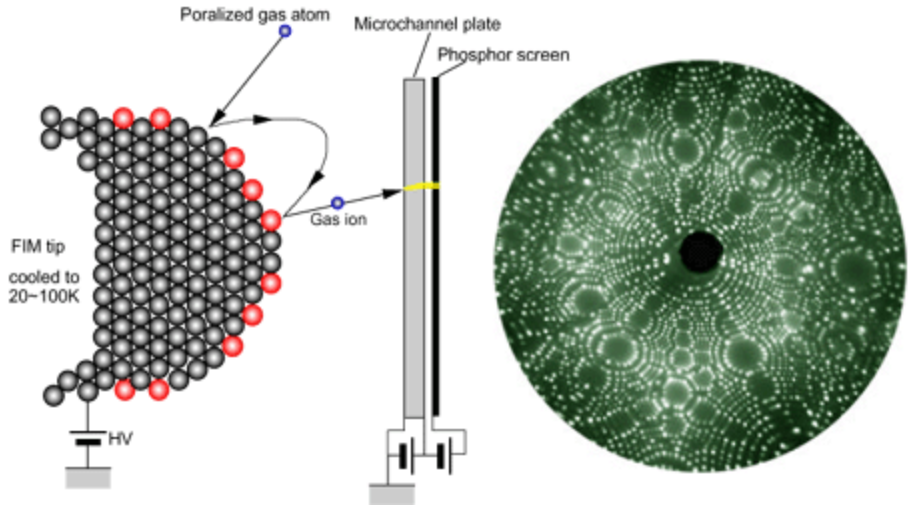

Fig. 3.1 Principle of field ion microscope (FIM).

The field ion image is produced by the projection of image gas atoms that are ionized by the high positive voltage on the tip, onto the fluorescent screen. A schematic diagram of the multiple steps imaging process is shown in Fig. 3.1. The image gas atoms in the vicinity of the tip are polarized because of the high field and then attracted to the apex region of the specimen. After a series of collisions with the tip during which the image gas atoms lose part of their kinetic energy, these image gas atoms become thermally accommodated to the cryogenic temperature of the specimen. If the field is sufficiently high, these image gas atoms are field ionized by a quantum-mechanical tunneling process. The ions produced

are then radially repelled from the surface of the tip towards the microchannel plate and screen assembly. A microchannel plate image intensifier positioned immediately in front of the phosphor screen produces between 10^3 and 10^4 electrons for each input ion. These electrons are accelerated towards the phosphor screen where they produce a visible image.

The angular width of the beam, mostly determined by the random velocity component of the ions, but also changing inversely with the tip radius, can be as narrow as 0.1°. Thus the total ion image typically encompassing about two-third of a hemisphere, can be quite sharp and finely detailed if a relatively large tip radius is used (in practice up to 2000 A°). The resolution improves with a decreasing tip radius but the images of small tips with radii down below 100Å appear quite blurred due to their unnecessary over magnification. The magnification is up to several millions tip diameter and the resolution is often between 2-3 A°. The strength of the ion beam coming from a single surface atom is fairly independent of the tip radius, and under practical conditions is some 10^5 ioni/s, generating an ion current of 10^{-14} A.

The tips are normally prepared from small cylindrical samples, mostly in the form of fine wires, made from the material to be studied. Chemical or electrochemical etching and polishing forms a conical needle ending in an extremely sharp point, which is in its dimensions well below the range of an optical microscope. For FEM the tip is annealed at a temperature above one half of the melting point, where surface migration of most metals becomes fast enough to rearrange the surface by approaching a shape of minimum free surface energy. The resulting tip shows flat low index crystal plane connected by smoothly curved intermediate regions. It is regular enough for the limited resolution of FEM but not for FIM. In FIM the tip is submitted to such a high field that metal atoms evaporate from the surface even at cryogenic temperatures. The evaporation of the most protuberant places occurs first, because of the exceedingly large local field. This process is continued as long as asperities remain on the surface, resulting in a field evaporated end form which is atomically smooth and crystallograhically as perfect as the bulk material of the specimen. However, one finds a variation of the local tip radius as large as 3 or 4. The reason of this variation was recognized to be the dependence on the crystallographic orientation of the binding energy of the atom.

The effective field strength at an evaporation site can be written as:

$$E = \beta_1 \beta_2 \frac{V}{k r_t} \tag{3.2}$$

where V is the applied voltage (above 500 V/Å), r_t is the average tip radius and k is a semi-constant taking a value around 5 for the most commonly used tip shapes. β_1 is a field enhancement factor, formally taking care of the effect of smaller or larger than average local tip radius, and β_2 is a short range enhancement factor used to account for geometric factors such as the degree of protrusion of a lattice step.

3.3 Theory of field ionization microscopy

When a very sharp metallic needle is subjected to a high voltage of a few kilovolts, an intense electric field is generated at the surface by the positive charges present at the surface. Indeed, the application of

the high voltage induces the free electrons to be, on average, displaced inwards by a small amount to screen the electric field, leaving partly charged atoms at the very surface. For a non-flat surface, protruding atoms are subject to a greater charge density, as shown in Fig. 3.2. Since the electric field at the surface is directly proportional to the charge density, it will be higher around these local protrusions. In the case of an atomically smooth curved surface, these protrusions correspond to the edges of atomic terraces. By imaging the distribution of the field intensity at the surface, the field ion microscope provides an atomically resolved image of the surface itself.

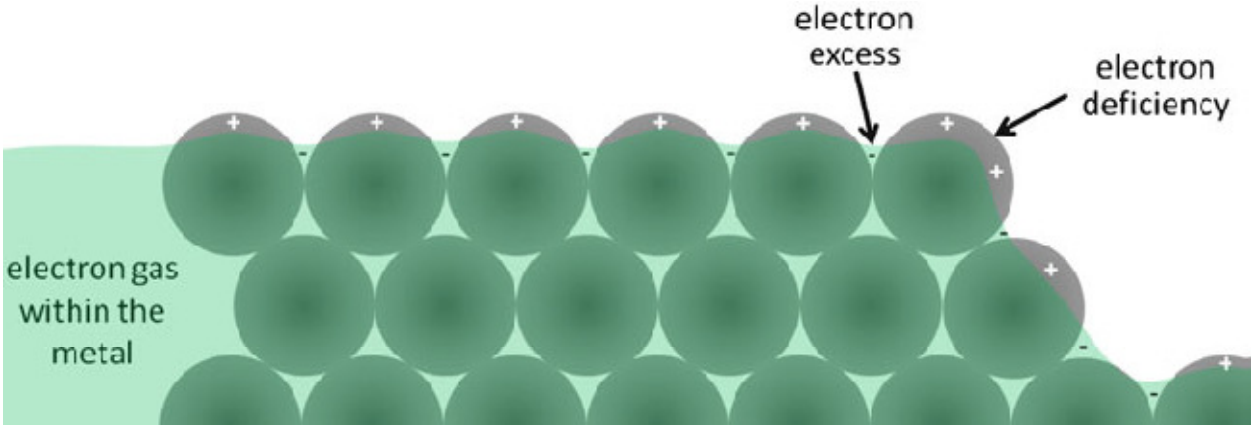

Fig. 3.2 Schematic view of the surface of a positively charged metal

Field ionization is the field-induced removal of an electron from an atom. Figure 3.3 schematically presents the potential energy level of a gas atom in the vicinity of a metal surface in the absence or presence of an electric field. The electric field polarizes the gas atom, deforming the potential curve. As the atom approaches the metal surface the potential barrier narrows. A supplementary reduction of the barrier height is caused by the image force or exchange and correlations interactions.

When subjected to a very strong electric field, an electron from the outer shell of the gas atom can tunnel through the energy barrier towards an empty energy level at the metal surface. The width of the barrier is proportional to the electric field, and thus, the ionization probability is critically dependent on the amplitude of the electric field. Field ionization will occur as close as possible to the surface, where the electric field is most intense. However, the energy of the electron from the gas atom must coincide with, or be higher than, the lowest available conduction level in the metal, which is close to the Fermi level. If this condition is not fulfilled, there are no vacant energy levels in the metal available for the tunneling electron, because all states below the Fermi level are already occupied. As a result, this process can only take place when the gas atom is beyond a critical distance away from the surface.

To a first approximation, the critical distance can be written as:

$$I - \phi = V(z) = eE_{el} z - \frac{e^2}{16\pi\varepsilon_0 z} \tag{3.3}$$

$$eE_{el} z_c = I - \phi + \frac{e^2}{16\pi\varepsilon_0 z_c} \quad . \tag{3.4}$$

where I is the energy of first ionization, Φ is the work function of the surface, and E_{el} is the electric field and the last term is the image potential. Usually, the image potential is small and eq. 3.4 can be reduced to

$$z_c = \frac{I - \phi}{eE_{el}} \quad . \tag{3.5}$$

In the case of a helium atom (I= 24.59 eV) ionized at 5 V/A° in the vicinity of a pure tungsten surface, where the work function is typically ~4.5 eV, this distance corresponds to approximately ~4Å. Hence, ionization takes place at this distance from the surface mostly within a thin zone of thickness smaller than $0.1z_c$.

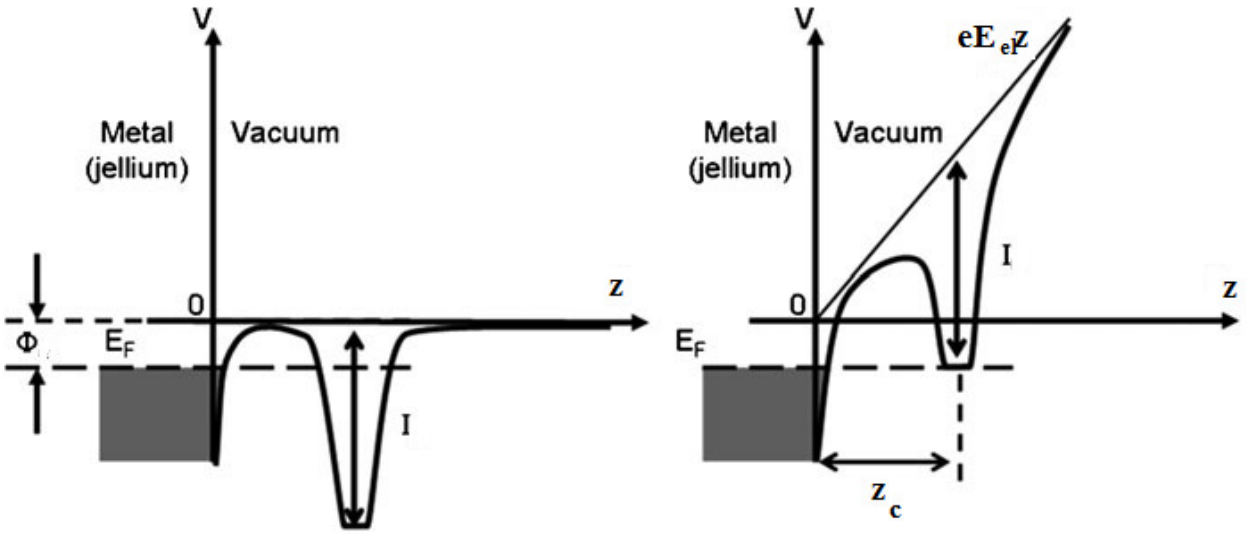

Fig. 3.3 Potential energy diagram as a function of the distance to the surface (z) of an electron from a gas atom in the vicinity of a tip (left) in the absence of an electric field, and (right) subject to an applied electric field, E_{el}. The energy of the first ionization is I, z_c is the critical distance of ionization, E_F is the Fermi energy and Φ is the work function of the surface.

The gas atoms strike the tip and bounce back and forth on its surface, losing some of their kinetic energy with each interaction. This energy is transferred to the lattice in a process that may be viewed as a thermal accommodation of the gas atoms prior to their ionization. In the best-case scenario, the energy of the gas atoms will be diminished to a level as low as the thermal energy of the tip's surface. As the kinetic energy of the gas atoms progressively decreases, there is a corresponding decrease in their velocity, which in turn increases the time spent by the atoms in the ionization zone, around the critical distance above the surface of the tip. Hence, the imaging-gas atoms will execute series of hops on and around the tip surface, with each hop diminishing their energy more than the previous, until ionization finally occurs. These new positively charged gas ions now experience the electric field force from the highly positive potential of the tip. As a result, they are repelled from the tip surface on a trajectory that is remarkably close to the normal to the tangent of the specimen surface. The gas ions accelerate away from the tip through the microscope chamber and eventually strike the screen, providing a vastly magnified projection of the specimen surface, a field ion micrograph.

3.4 The FIM image

The brightness of a micrograph is not uniform. The variation of the local radius of curvature of the tip also alters the local magnification and thereby the local brightness. In tungsten images, the brightest area is the {111} region and the darkest area is the {110} region. The radius of curvature near the {110} poles is several times larger than that near the {111} poles, the latter representing the more protruding part of the emitter. The field ionization probability is thus higher in the {111} regions. Also, the field ionization rate is affected by the work function, surface states and electron orbital directions.

Indexing of a field ion micrograph can be achieved most easily by comparing the micrograph with the expected symmetry and surface structure of the lattice. Originally a ball model was built to approach a hemisphere as closely as compatible with the steps of a bcc lattice. When the atoms along the protruding edges were painted with phosphorescent paint the model viewed in dark resembles the field ion images very well. This resemblance suggests that the indexing of a micrograph can be done by comparing it with a geometrical projection of the crystal lattice.

A symmetrical crystallographic map is obtained only if the pole of projection lays on the axis AC normal to the projection plane. In general, the point P may lie at any distance from the projection plane. Assuming that P lies at a distance nR from the projection plane, where n is a real number and R is the radius of the sphere, various types of projection can be defined:

- Orthographic projection of projection when the pole P lies at infinity (n = ∞);
- Stereographic projection when the pole P coincides with C (n = 2);
- Gnomonic projection when the pole P coincides with the center O of the sphere (n = 1).

The distance x_n from A to the image point B is:

$$x_n = \frac{n \sin \gamma}{n - 1 + \cos \gamma} R . \qquad (3.10)$$

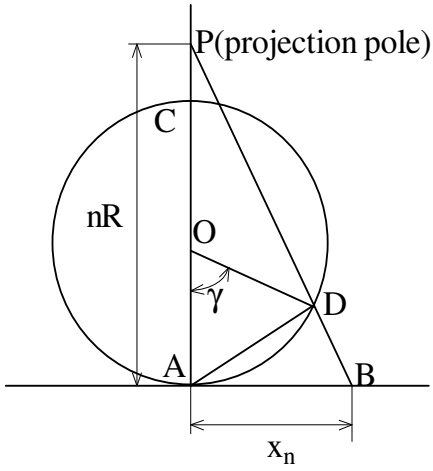

Fig.3.4 The projection of a sphere on a plane

Eq. (3.10) can be easily demonstrated geometrically for n = 2 (stereographic projection). Then, eq. (3.10) becomes

$$x_n = \frac{n \sin \gamma}{1 + \cos \gamma} R = 2R \cdot tg \frac{\gamma}{2} . \qquad (3.11)$$

In this case (P=C), one observes that $\Delta CAD \approx \Delta CAB$ (rectangular triangles having a common angle $A\hat{C}B$). The resemblance ratios are:

$$\frac{x_n}{AD} = \frac{AC}{CD} \Leftrightarrow \quad \frac{x_n}{2R\sin\frac{\gamma}{2}} = \frac{2R}{2R\sin\left(90 - \frac{\gamma}{2}\right)} \Rightarrow x_n = 2R\frac{\sin\frac{\gamma}{2}}{\cos\frac{\gamma}{2}} . \tag{3.12}$$

The projected distance between two crystallographic orientations can thus be calculated once the angle γ is known. In the cubic lattice the angle γ between the normals to the two planes with Miller indices $(h_1k_1l_1)$ and $(h_2k_2l_2)$ is also the angle between the directions $[h_1k_1l_1]$ and $[h_2k_2l_2]$ and is given by:

$$\cos\gamma = \frac{h_1h_2 + k_1k_2 + l_1l_2}{\sqrt{\left(h_1^2 + k_1^2 + l_1^2\right)\left(h_2^2 + k_2^2 + l_2^2\right)}} . \tag{3.13}$$

To have an idea of a stereographic projection let us take a cube oriented (001), meaning that the Oz axis is normal to the projection plane. The Oz axis is projected in the center of the projection circle like in figure 3.5. The Ox axis is projected on the vertical diameter, thus the vertical diameter is limited by 100 and $\bar{1}00$ points. The Oy axis is projected on the horizontal diameter, thus the vertical diameter is limited by 010 şi $0\bar{1}0$. The diameter limited by the 110 and $\bar{1}\,\bar{1}0$ points is mid-way between the vertical and the horizontal diameter. Any other direction can be projected using the angle γ and the Wulff net (named after the Ukrainian mineralogist George Yuri Viktorovich Wulff), represented in fig. 3.6.

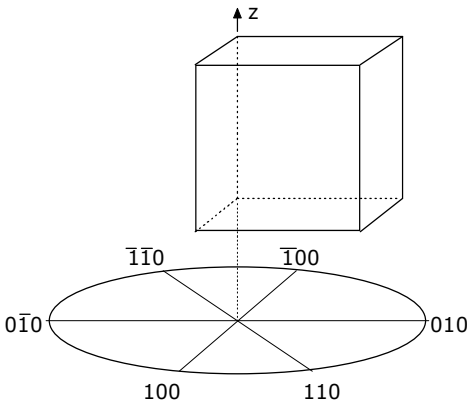

Fig. 3.5 The stereographic projection of a cube

On the Wulff net, the images of the parallels and meridians intersect at right angles. This orthogonality property is a consequence of the angle-preserving property of the stereoscopic projection.

Before the advent of the field ion microscope, the radius of a field electron emitter was either determined from the tip profile seen with an electron microscope or by the measurement of the field electron emission current. However the electron microscope, having an uncertain magnification, is not very suitable for direct determination of the emitter radius. The most accurate and yet the most simple method of determining the field ion

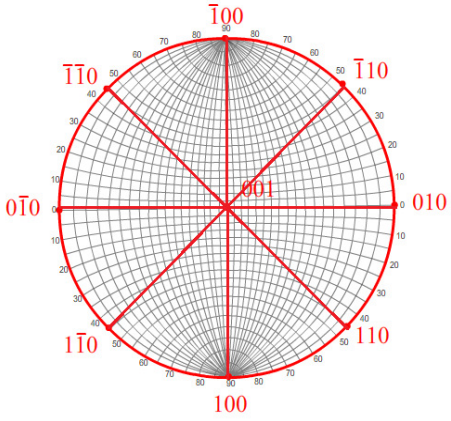

Fig. 3.6 The Wulff net and some points for the stereographic projection of a fcc lattice, (001) oriented

emitter radius is to count the number of net rings between two pole of known angular separation in the FIM micrograph.

As one can see in fig. 3.7, if between two directions $[h_1k_1l_1]$ and $[h_2k_2l_2]$ there are p atomic planes, represented like steps of height s, then the FIM image presents p net rings between the poles $h_1k_1l_1$ and $h_2k_2l_2$. From fig. 3.7 one sees that:

$$R = \frac{ps}{1-\cos\gamma}. \qquad (3.14)$$

For cubic crystals, the step heights s of $\{hkl\}$ poles, are given by

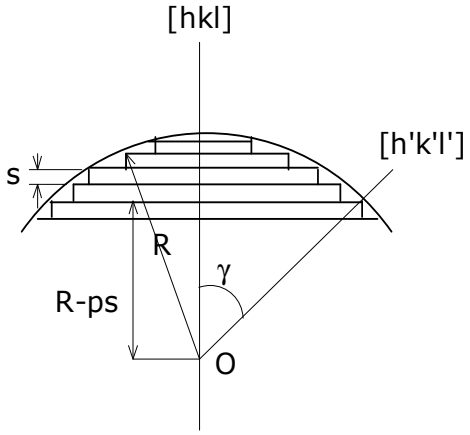

Fig. 3.7 Determination of the tip radius

$$s = \frac{a}{\delta\sqrt{h^2+k^2+l^2}}, \qquad (3.15)$$

where a is the lattice constant and δ is defined as in Table 3.1.

Table 3.1

	Simple cubic lattice	**bcc**	**fcc**
δ = 1	for all h,k,l	h+k+l is even	h,k,l all odd
δ = 2		h+k+l is odd	h,k,l of mixed parity

4. The scanning tunneling microscope (STM)

4.1 Introduction

The first scanning tunneling microscope was built by Gerd Binnig and Heinrich Rohrer from IBM Zurich Laboratory in 1982. The debut was made by resolving the structure of Si(111)-7x7 in real space (a restructured semiconductor surface providing strong evidence for surface adatoms and triangular stacking fault regions within the surface unit cell as structural features of the reconstruction). However the structure of the reconstruction, implying the last three atomic planes, was ultimately determined by a combination of experimental techniques including LEED, reflection high energy electron diffraction (RHEED), impact-collision ion scattering spectroscopy (ICISS), transmission electron microscopy (TEM)). The STM has proven itself to be as important as other techniques, and in some cases more important, but has rarely led to a structural determination by itself.

The STM can resolve local electronic structure at an atomic scale on every kind of conducting solid surface, thus also allowing its local atomic structure to be revealed. An extension of scanning tunneling microscope is the atomic force microscope that can image the local atomic structure even on insulating surfaces. The ability of STM and AFM to image in various ambiances with virtually no damage or interference to the sample made it possible to observe processes continuously. For example the entire process of a living cell infected by viruses was investigated in situ using AFM.

The field of scanning tunneling microscopy has enjoyed a rapid and sustained growth, phenomenal for a new branch of science. STMs are being developed commercially at an astonishing speed.

In STM a sharp metallic tip is stabilized at a distance d (of few Å) from a conducting sample. In this way there is a superposition between the electronic functions of the surface and of the tip. If a voltage is applied between these two electrodes, it gives rise to a tunneling current.

The tunneling microscope images the electronic charge density at the Fermi level outside of the surface rather than the true positions of the atomic nuclei in the plane of the surface. While the electronic charge density above the surface at the Fermi level is directly related to the positions of the atoms in the surface, the theory describing this relationship is a complex problem, dependent upon models of charge density, which is at the limits of present day computational capabilities especially for the semiconductor surface reconstructions with large surface unit cells. For this reason STM has not usually been utilized to determine the coordinates of the surface atoms to arrive at a structural model, but rather to eliminate some of the existent structural models, due to the inconsistencies between the model and the STM data.

The most significant impact the STM has made in the area of surface science has been a shift in the point of view of surface scientist away from the band structure like perspective (that emphasizes the collective properties of repetitive lattice structures), towards a more atomistic point of view where the influence of atomic scale features can be studied. Although this point of view was first brought into the realm of surface science with the introduction of FIM, the limitations imposed by the use of sharp FIM (the tip) samples in high electrostatic fields did not allow its use for fundamental studies of semiconductor surfaces.

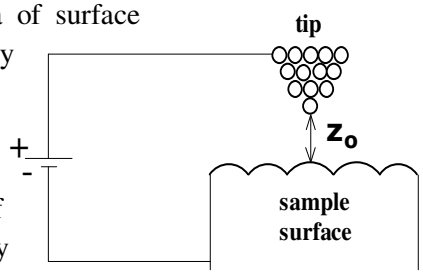

Fig. 4.1 Principle of STM

The assembly STM tip-surface is represented in figure 4.1. STM can work properly only in UHV (10^{-9} bar to 10^{-11} mbar). The microscopic tip-surface distance is the distance from the plane of the topmost nuclei of the sample to the nucleus of the apex atom of the tip, denoted z_0. Let us consider the following four regimes of interactions:

1. When the tip-sample distance is large, for example z_0>100Å, the mutual interaction is negligible. By applying a large electrical field between them, field emission may occur. Those phenomena can be described as the interaction of one electrode with an electric field without considering any interactions from the other electrode.

2. At intermediate distances, for example 10<z_0<100 Å, a long range interaction between the tip and the sample takes place. The wavefunctions of both the tip and the sample are distorted, and a Van der Waals force arises. The Van der Waals force between two atoms follows a power law:

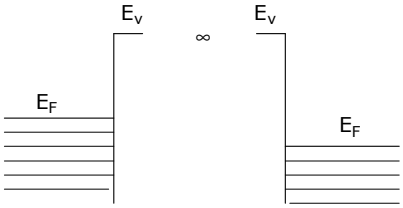

Fig.4.2 At large distances the vacuum level is the same for the tip and the sample

a) $U = -\frac{const}{r^6}$ (const>0) for atoms being in an S state, for which the dipolar and quadrupolar perturbation is different from zero only in the second order of perturbation;

b) $U = \frac{const}{r^5}$ ($const \gtrless 0$) for atoms being in the same state, but different from S, for which the quadrupol perturbation is nonzero in the first order of perturbations;

c) $U = \frac{const}{r^3}$ ($const \gtrless 0$) for atoms being in different states, different from S, for which the dipolar interactions is nonzero in the first order of perturbation;

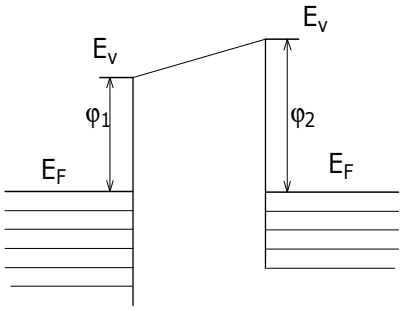

Fig. 4.3 At short distances the Fermi level is the same

3. At short distances, $3 < z_0 < 10$ A°, the electrons may transfer from one side to another. The exchange of electrons gives rise to an attractive interaction, the delocalization energy of Heisenberg and Pauling, which is the origin of the chemical bond. If a bias is applied between them, a tunneling current occurs. The delocalization energy follows an exponential law, and can be much larger than the van de Waals interaction.

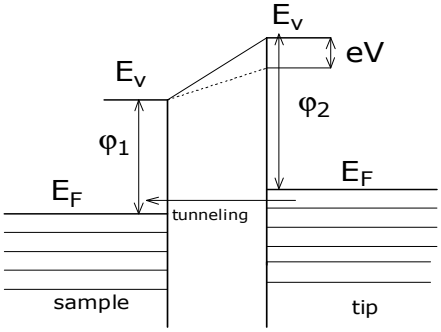

Fig. 4.4 If a voltage V>0 is applied the electrons can tunnel from the tip to the sample

4. At extremely short distances, for example, $z_0 < 3$ Å, the repulsive force becomes dominant. It has very steep distance dependence. The tip-sample distance is virtually determined by the short ranged repulsive force. By pushing the tip farther toward the sample surface, the tip and the sample deform accordingly.

4.2 The quantum theory of tunneling

An ubiquitos experimental observation is that the apparent barrier height is almost constant up to a mechanical contact, in spite of the effect of the image force. The apparent barrier, defined in terms of the rate of change (at constant bias) of the logarithm of current with tip-sample separation, is found to be substantially lower than the sample work function in the range of tip-sample separations commonly used experimentally, and becomes very small in the region just before contact between tip and sample.

This is why the tunneling current in STM may be approximated by the tunneling current through a square barrier. The Schrödinger equation for the square, unidimensional barrier of height U_b and width a is:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(z)}{dz^2}+U(z)\psi(z)=E\psi(z) \tag{4.1}$$

If $E > U_b$, than the barrier is clasically allowed. Considering an incident electron moving to the left, the wave function can be written:

$$\begin{cases} \psi(z)=\exp(-ik_1 z)+A\exp(ik_1 z) & z>a \\ \psi(z)=B_1\exp(-ik_2 z)+B_2\exp(ik_2 z) & 0<z\le a \\ \psi(z)=C\exp(-ik_1 z) & z\le 0 \end{cases} \tag{4.2}$$

where $k_1=\frac{\sqrt{2mE}}{\hbar}$ and $k_2=\frac{\sqrt{2m(E-U_b)}}{\hbar}$. The first function is a mixture of incoming and reflected waves before the barrier, the second function represents a mixture of transmitted and reflected waves in the barrier region and the third function represents the wave transmitted by the barrier. There two different conventions in defining the transmission coefficient, according to the normalization of the initial state. One is to normalize it with respect to the incoming traveling wave thus obtaining the transmission coefficient $T=|C|^2$.

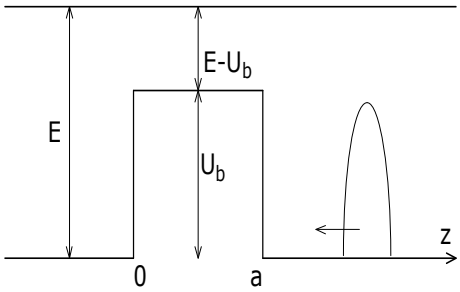

Fig. 4.5 The square potential barrier and the incident wave packet

Another is to normalize it with respect to the quasistanding wave on the left hand side, thus obtaining

$$T=\frac{|C|^2}{1+|A|^2}. \tag{4.3}$$

We will choose this latter convention because it connects more conveniently with the perturbation theory. For determination of A, B_1, B_2, C one imposes the continuity conditions in z = 0 and z = a on the wave function and its derivative. Therefore, the continuity in z=0 leads to

$$C=B_1+B_2,\ -k_1C=-k_2B_1+k_2B_2, \tag{4.4}$$

and the continuity requirement in z = a leads to

$$\begin{aligned}&\exp(-ik_1a)+A\exp(ik_1a)=B_1\exp(-ik_2a)+B_2\exp(ik_2a)\\&-k_1\exp(-ik_1a)+k_1A\exp(ik_1a)=-k_2\left(B_1\exp(-ik_2a)-B_2\exp(ik_2a)\right)\end{aligned}. \tag{4.5}$$

From (3.4) şi (3.5) the expressions of A, B_1, B_2, C may be calculated and read as:

$$\begin{aligned}A&=\exp(-2ik_1a)\frac{\left(k_2^2-k_1^2\right)\left(\exp(2ik_2a)-1\right)}{(k_2+k_1)^2-\exp(2ik_2a)(k_2-k_1)^2}\\B_1&=\exp(i(k_2-k_1)a)\frac{2k_1(k_2+k_1)}{(k_2+k_1)^2-\exp(2ik_2a)(k_2-k_1)^2}\\B_2&=\exp(i(k_2-k_1)a)\frac{2k_1(k_2-k_1)}{(k_2+k_1)^2-\exp(2ik_2a)(k_2-k_1)^2}\\C&=\exp(i(k_2-k_1)a)\frac{4k_1k_2}{(k_2+k_1)^2-\exp(2ik_2a)(k_2-k_1)^2}\end{aligned}. \tag{4.6}$$

Introducing eq. (4.6) în (4.3), the transmission coefficient of the classically allowed barrier is

$$T=\left[1+\frac{1}{2}\left(\frac{k_2}{k_1}-\frac{k_1}{k_2}\right)^2\sin^2(k_2a)\right]^{-1}. \tag{4.7}$$

If the barrier is classically forbidden then the transmission coefficient is obtained from eq. (4.7) by replacing k_2 with ik_2 (where $k_2=\frac{\sqrt{2m(U_b-E)}}{\hbar}$) thus:

$$\left(\frac{k_2}{k_1}-\frac{k_1}{k_2}\right)^2\rightarrow\left(\frac{ik_2}{k_1}-\frac{k_1}{ik_2}\right)^2=i^2\left(\frac{k_2}{k_1}+\frac{k_1}{k_2}\right)^2=-\left(\frac{k_2}{k_1}+\frac{k_1}{k_2}\right)^2$$

$$\sin^2(k_2a)\rightarrow\sin^2(ik_2a)=\left(\frac{\exp(-k_2a)-\exp(k_2a)}{2i}\right)^2=-\left(\frac{\exp(-k_2a)-\exp(k_2a)}{2}\right)^2=-sh^2(k_2a),$$

finally leading to

$$T=\left[1+\frac{1}{2}\left(\frac{k_2}{k_1}+\frac{k_1}{k_2}\right)^2sh^2(k_2a)\right]^{-1}. \tag{4.8}$$

For thick barriers: $\exp(k_2a)>>\exp(-k_2a)$ and $sh(k_2a)\sim\exp(k_2a)/2$, and eq. (4.8) becomes

$$T\approx 8\left(\frac{k_2}{k_1}+\frac{k_1}{k_2}\right)^{-2}e^{-2k_2a}. \tag{4.9}$$

We recovered here the usual formula for tunneling probability which exhibits an exponentially decaying behavior. The tunneling current is proportional to the transmission coefficient so it is dependent on the width and height of the barrier and will decay exponentially with the tip-surface distance

$$I \approx f(U_b, a)e^{-2k_2 a} \tag{4.10}$$

If U_b - E = 5 eV, a = 5Å then $2k_2a$ = 11,4 and exp(k_2a) = 89321.7, but if a = 4Å , then $2k_2a$ = 9.1 and exp(k_2a) = 8955.2, so the tunneling current increases with one order of magnitude when a decreases with 1Å. Consequently the vertical resolution of STM is excellent, lower than 0.01Å. However the lateral resolution which is the most important in STM is of the order of few Å and depends on the tip end form. The tip may be ideally made to consist of one atom like in fig. 4.1.

4.3 Uncertainty principle considerations

From a semi-classical point of view, there are two qualitatively different cases. Then the barrier is classically forbidden the process is called tunneling and when the barrier is classically allowed the process is called channeling or ballistic transport. We will show that for square potential barriers of atomic scale, the distinction between classically forbidden or allowed cases disappears. There is one unified phenomenon, the quantum transmission. It can be shown from the following two points of view:

a. The first view is based on the uncertainty relation between time and energy. When the electron is in the region of the barrier, its velocity, as determined by the kinetic energy, is:

$$v = \sqrt{\frac{2(E - U_b)}{m}} \tag{4.11}$$

The time to pass the barrier is

$$\Delta t = \frac{a}{v} = a\sqrt{\frac{m}{2(E - U_b)}} \tag{4.12}$$

The uncertainty principle for an electron in the barrier region is then

$$\Delta E > \frac{\hbar}{\Delta t} = \frac{\hbar}{a}\sqrt{\frac{2(E - U_b)}{m}} \; . \tag{4.13}$$

At E- U_b=1.5 eV, and a=3 Å, the energy uncertainty is ΔE~1.59Å. In other words, the energy uncertainty is larger than the absolute value of the kinetic energy. Therefore for atom scale barriers, the distinction between tunneling and ballistic transport disappears.

b. The second point of view is based on the uncertainty relation between the coordinate and momentum. In the region of a classically allowed barrier, the de Broglie wavelength of an electron is

$$\lambda = \frac{h}{\sqrt{2m(E - U_b)}} . \tag{4.14}$$

For example, an electron in the barrier region with kinetic energy 1.5 eV has a de Broglie wavelength of 10 Å. Therefore the position uncertainty is much greater than the barrier thickness.

4.4 The STM image

During image acquisition, the tip scans across the sample using *xy* piezoelectric elements (the tip can be moved in all dimensions with a precision better than 0.01Å), and a feedback loop adjusts the tip height *z* in order to maintain a constant current. Then the tip height signal is displayed resulting in an STM image, which contains both topographic and electronic information. The corrugation amplitude measured in STM is a quantity which is defined as the difference between the largest and smallest tip-sample distances in a constant-current experiment. Due to the exponential dependence of the tunneling current on the width of the barrier, such an experimental setup allows a high resolution vertical to the surface. Combined with the high accuracy of the positioning of the tip parallel to the surface, images with corrugation amplitude smaller than 0.01Å can be obtained. Steps and islands are mapped easily and if the experimental setup is stable enough one can achieve atomic resolution. In images with atomic resolution single atoms in the surface topmost layer are observed.

To understand what this tip height variation corresponds to, it is necessary to go beyond a one-dimensional analysis. The tunneling current is given to the first order in this formalism by:

$$I = \frac{2\pi e}{\hbar}\sum_{\mu\upsilon} f\left(E_\mu\right)\left(1-f\left(E_\upsilon + eV\right)\right)\left|M_{\mu\upsilon}\right|^2 \delta\left(E_\mu - E_\upsilon\right) \tag{4.15}$$

where

$$f(E) = \frac{1}{1+\exp\dfrac{E-E_F}{kT}} \tag{4.16}$$

is the Fermi-Dirac function showing the occupation probability of level E, E_F is the Fermi energy, Vis the applied voltage and $M_{\mu\nu}$ is the tunneling matrix element between the unperturbed state ψ_μ of the probe (tip) and ψ_υ of the surface, E_ν and E_μ are the energy of the states ψ_μ, ψ_υ in the absence of tunneling. The quantity $1-f\left(E_\upsilon + eV\right)$ is the probability that the energy level E_ν+eV be unoccupied. The expression (4.15) resembles the transition probability in the first order of perturbations but it is different because ψ_μ and ψ_υ are un-orthogonal states of different Hamiltonians and the transition takes place from level E_μ to E_ν+eV (not from E_μ to E_ν). Because STM is usually performed at room temperature or even lower and small voltages (≈ 10 meV for metal to metal tunneling), eq. (4.15) can be rearranged (after many calculations, see for example M. C. Desjonquères, D. Spanjard, *Concepts in surface physics*, Springer-Verlag, Heidelberg, 1993, appendix G) as

$$I = \frac{2\pi}{\hbar}e^2 V\sum_{\mu\upsilon}\left|M_{\mu\upsilon}\right|^2 \delta\left(E_\mu - E_F\right)\delta\left(E_\upsilon - E_F\right), \tag{4.17}$$

A simple interpretation of eq. (4.17) can be gained by considering the most simple tip: a singular point probe located at $\vec{r}_o$. In this limit case the wave function of the tip is $\psi_\mu(\vec{r}) = \delta(\vec{r} - \vec{r}_o)$ and $M_{\mu\nu}$ will be proportional to $\psi_\upsilon(\vec{r}_o)$. In this limit (4.17) becomes:

$$I \approx \sum_\upsilon |\psi_\upsilon(r_0)|^2 \delta(E_\upsilon - E_F) = \rho(r_0, E_F) \tag{4.18}$$

where $\rho(r_0, E_F)$ is the surface local density of states at E_F at the position of the point probe. Thus the microscope image is a contour map of the charge density of the surface at the Fermi level. This remains approximately true in a more complete treatment. In interpreting the STM images, an important question is how the local density of states at E_F relates to the positions of atoms on the sample surface. In order to see at what does the constant-current image correspond to, one has to calculate $\rho(r_0, E_F)$ on the basis of a structural model and to compare the image with the obtained charge density.

4.5 STM instrumentation

The common elements of STM are the coarse positioner, the piezoelectric scanner, vibration isolation, the feedback system and current amplifier.

1. The coarse positioner moves the tip versus the sample in steps exceeding the range of the piezoelectric scanner, typically a fraction of a micrometer. The initial success of STM was partly due to the piezoelectric stepper, nicknamed the louse.

2. The piezoelectric effect was discovered by Pierre and Jacques Curie in 1880. By applying a vertical tension to a quartz plate, an electrical charge was generated. Few months later Lipmann predicted the existence of the inverse piezoelectric effect: by applying a voltage to a quartz plate, a deformation should be observed. The effect was confirmed by the Curie brothers. They obtained a parallel piezoelectric coefficient:

$$d_{zz} = \frac{\delta z}{z} \cdot \frac{1}{E_z} \tag{4.19}$$

of 0.021 Å/V for quartz which matches accurately the results of modern measurements. Å/V is the natural unit in STM for the piezoelectric coefficient.

In order to obtain a greater piezoelectric coefficient in STM one uses various kinds of lead zirconate titanate ceramics (PZT). The PZT ceramics are made by firing a mixture of $PbZrO_3$ and $PbTiO_3$ together with a small amount of additives at about 1350°C under strictly controlled conditions. The result is a solid solution. Microscopically it consists of small ferroelectric crystals in random orientations. Macroscopically it is isotropic due to the random arrangement of the electrical dipoles and does not have a permanent dipole moment. To produce a permanent electric polarization a high electric field is applied for a sufficiently long period of time, for example 60kV/cm for 1h. A strong piezoelectricity is generated, that is temperature dependent. As a convention the direction of the poling field is considered the z

direction. In contrast to piezoelectric single crystals, such as quartz, the piezoelectricity of PZT ceramics decays with time due to relaxation. The piezoelectric coefficient of PZT is (1-6) Å/V.

In the early years of STM instrumentation, the tripod piezoelectric scanners were the predominant choice. The displacements along the x, y, z directions are actuated by three independent PZT transducers. If larger displacements are required, an arrangement of bimorph is used. In a bimorph two thin plates of piezoelectric material are glued together with electrodes on both sides and an embedded center electrode. By applying a voltage one plate expands and the other one contracts. The composite flexes. However, it is very difficult to construct a three dimensional scanner based on bimorph. Their main advantage seems to be their large excursion per applied voltage. The predominant STM scanner is the tube scanner. A tube made of PZT, metalized on the outer and inner surfaces, is poled in the radial direction. The outside metal coating is sectioned into four quadrants. The inner metal coating is connected to the z voltage and two neighboring quadrants are connected to the varying x, y voltages, respectively. The remaining two quadrants are connected to a certain dc voltage to improve linearity. The tip is attached to the center of one of the dc quadrants.

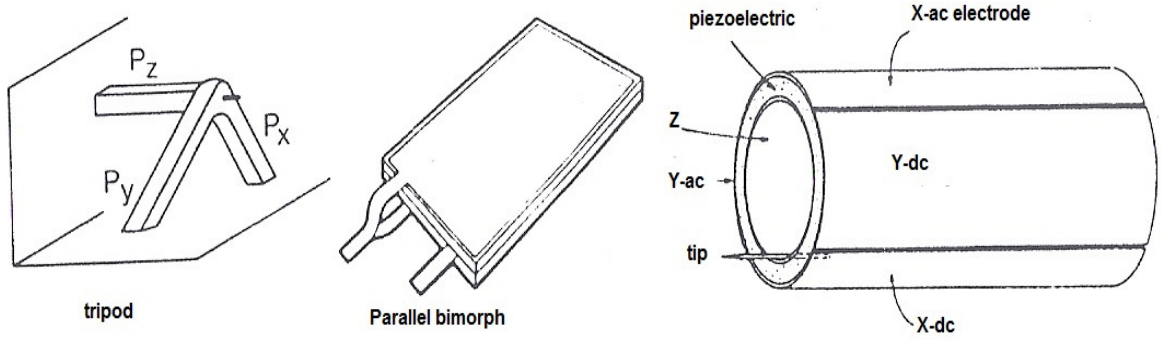

Fig. 4.6 The principal piezoelectric scanners: tripod, bimorph and tube.

3. Effective vibration isolation is one of the critical elements in achieving atomic resolution by STM. The typical corrugation amplitude for STM images is about 0.1 Å. Therefore the disturbance from external vibrations must be reduced to less than 0.01 Å=1pm. The basic idea is to make the internal resonance frequencies of the STM very high, and to mount it on a support with a very low resonance frequency. The support will follow only the low-frequency building vibrations and suppress most of the high-frequency components. The STM in turn will not be disturbed by the remaining low-frequency vibrations, because they do not introduce any internal motions into the STM (it just moves as a whole).

Before a vibration isolation device is designed, the vibration characteristic of the laboratory has to be measured. Typical frequencies are: i) 1 Hz building vibrations due to people walking around etc.; ii) 10…200 Hz building vibrations due to ventilation, appliances etc.; iii) 1…10 kHz lowest internal resonance of typical STMs.

The necessary quality of vibration isolation depends, of course, on the amount and spectral distribution of building vibrations present and on the lowest resonance frequency of the STM.

Popular vibration isolation systems used for STMs are:

- Air support tables: Heavy plate resting on inflated supports, possibly with active regulation to maintain constant height. They give a large, steady work surface, but expensive and bulky.
- "Pendulum": Massive blocks suspended by springs, rubber bands, elastic tubing. It's easy to achieve very low resonance frequencies by combining a large mass with a soft suspension. May need some damping to avoid oscillations at resonance (for example, internal friction of the suspension or eddy-current damping)
- Plate and rubber stacks: Have been successfully used with the modern, compact STM designs, which will tolerate some low-frequency vibration. It is composed of a stack of steel plates (typically postcard-sized and about 1 cm thick), with some rubbery material in between. Since the elastic material is quite stiff, the resonance frequencies of the individual stages are quite high. The isolation is effective only for vibrations with relatively high frequencies (>50Hz). However, the stack of multiple damping stages (each with different resonance frequency, since the supported mass decreases from bottom to top) is quite efficient.

4. The tunneling current in STM is very small, typically from 0.01 to 50nA. The current amplifier is thus essential element of an STM, which amplifies the tiny current and convert it into a voltage. The output of the logarithmic amplifier is compared with a reference voltage. The error signal is then sent to the feedback circuit which sends a voltage into the z piezo. If the tunneling current is larger than the preset current, then the voltage applied to the z piezo tends to withdraw the tip from the sample surface and vice versa. Therefore an equilibrium z position is established through the feedback loop. As the tip scans the contour height z also changes in time.

5. S-XRD (Surface X ray diffraction)

X-ray scattering is one of the oldest techniques used for the determination of the atomic structure of the solids. However, it is only recently applied to surface studies, because the scattered intensity from the surface is several orders of magnitude less than that from a bulk crystal and can be measured only when intense X-ray sources such as synchrotron radiation are available. A **synchrotron** is a particular type of cyclic particle accelerator, descended from the **cyclotron**.

5.1 The cyclotron

A **cyclotron** is a type of particle accelerator invented by Ernest O. Lawrence in 1932 in which charged particles accelerate outwards from the centre along a spiral path. The particles are held to a spiral trajectory by a static magnetic field and accelerated by a rapidly varying (radio frequency) electric field. Lawrence was awarded the 1939 Nobel Prize in physics for this invention. Cyclotrons were the most powerful particle accelerator technology until the 1950s when they were superseded by the synchrotron, but are still used to produce particle beams in physics and nuclear medicine. The largest single-magnet cyclotron was the 4.67 m synchrocyclotron built between 1940 and 1946 by Lawrence at the University of California at Berkeley, which could accelerate protons to 730 MeV.

A cyclotron accelerates a charged particle beam using a high frequency alternating voltage which is applied between two hollow "D"-shaped sheet metal electrodes called "Dees" inside a vacuum chamber. The Dees are placed face to face with a narrow gap between them, creating a cylindrical space within them for the particles to move. The particles are injected into the center of this space. The Dees are located between the poles of a large electromagnet which applies a static magnetic field B perpendicular to the electrode plane. The magnetic field causes the particles path to bend in a circle due to the Lorentz force perpendicular to their direction of motion.

If the particles speed were constant, they would travel in a circular path within the Dees under the influence of the magnetic field. However a radio frequency (RF) alternating voltage of several thousand volts is applied between the Dees. The frequency is set so that the particles make one circuit during a single cycle of the voltage. To achieve this, the frequency must match the particle's cyclotron resonance frequency

$$\frac{Mv^2}{r} = qvB \rightarrow r = \frac{Mv}{qB} \rightarrow \omega = \frac{v}{r} = \frac{qB}{m}, \tag{5.1}$$

where B is the magnetic field strength, q is the electric charge of the particle, and m is the relativistic mass of the charged particle. Each time after the particles pass to the other Dee electrode the polarity of the RF voltage reverses. Therefore, each time the particles cross the gap from one Dee electrode to the other, the electric field is in the correct direction to accelerate them. The particles increasing speed due to these pushes causes them to move in a larger radius circle with each rotation, so the particles move in a spiral path outward from the center to the rim of the Dees.

When they reach the rim the particles exit the Dees through a small gap between them, and hit a target located at the exit point at the rim of the chamber, or leave the cyclotron through an evacuating beam tube to hit a remote target. Various materials may be used for the target, and the nuclear reactions due to the collisions will create secondary particles which may be guided outside of the cyclotron and into instruments for analysis.

The cyclotron was the first "cyclical" accelerator. In the cyclotron, the particles encounter the accelerating voltage many times during their spiral path, and so are accelerated many times, so the output energy it is not limited to the accelerating voltage. A usual energy for a small cyclotron is 20MeV. If R is

the external ray of the Dees then v_{max} of the ion having a trajectory with this ray is $v_{max} = BR\frac{q}{m}$ and the corresponding kinetic energy is $E_c = \frac{mv_{max}^2}{2} = \frac{mq^2B^2R^2}{2}$.

Therefore, the limit to the cyclotron output energy for a given type of particle is the strength of the magnetic field **B**, which is limited to about 2 T for ferromagnetic electromagnets, and the radius of the Dees R, which is determined by the diameter of the magnet pole pieces. Very large magnets were constructed for cyclotrons, culminating in Lawrence's 1946 synchrocyclotron, which had pole pieces of 4.67 m diameter.

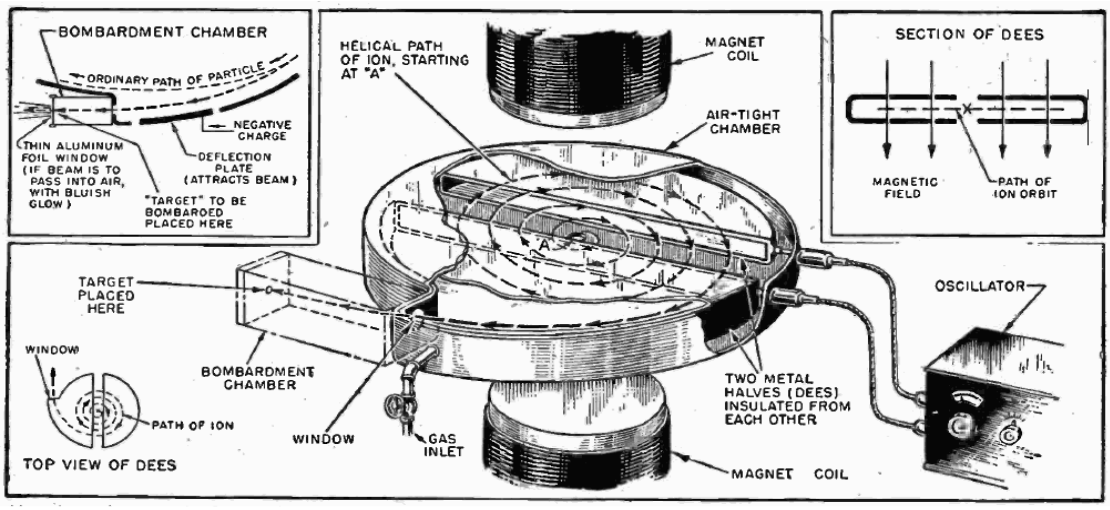

Fig. 5.1 Image of a cyclotron

In the nonrelativistic approximation, the frequency does not depend upon the radius of the particle orbit, since the particle mass is constant. As the beam spirals out, its frequency does not decrease, and it must continue to accelerate, as it is travelling a greater distance in the same time period. In contrast to this approximation, as particles approach the speed of light, their relativistic mass increases, requiring either modifications to the frequency, leading to the *synchrocyclotron*, or modifications to the magnetic field during the acceleration, which leads to the *isochronous cyclotron*. The relativistic cyclotron angular frequency can be rewritten as $\omega = B\frac{q}{m}\sqrt{1-\frac{v^2}{c^2}}$ where $m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$ and m_0 is the particle rest mass.

5.2 The synchrotron

The **synchrotron** is a particular type of cyclic particle accelerator, in which the guiding magnetic field (bending the particles into a closed path) is time-dependent, being *synchronized* to a particle beam of increasing kinetic energy. The synchrotron is one of the first accelerator concepts that enable the construction of large-scale facilities, since bending, beam focusing and acceleration can be separated into different components. The most powerful modern particle accelerators use versions of the synchrotron design. The largest synchrotron-type accelerator is the 27 kilometers circumference Large Hadron Collider (LHC) near Geneva, Switzerland, built in 2008 by CERN.

In a synchrotron, the magnetic field strength is varied in time, rather than in space. For particles that are not close to the speed of light, the frequency of the applied electromagnetic field may also change to follow their non-constant circulation time. By increasing these parameters accordingly as the particles gain energy, their circulation path can be held constant as they are accelerated. This allows the vacuum

chamber for the particles to be a large thin torus, rather than a disk as in previous, compact accelerator designs. Also, the thin profile of the vacuum chamber allowed for a more efficient use of magnetic fields than in a cyclotron, enabling the cost-effective construction of larger synchrotrons. Unlike in a cyclotron, synchrotrons are unable to accelerate particles from zero kinetic energy; one of the obvious reasons for this is that its closed particle path would be cut by a device that emits particles. Thus, schemes were developed to inject pre-accelerated particle beams into a synchrotron. The pre-acceleration can be realized by a chain of other accelerator structures like a linac, a microtron or another synchrotron; all of these in turn need to be fed by a particle source comprising a simple high voltage power supply.
Properties of the synchrotron radiation:

1. Broad Spectrum (which covers from microwaves to hard X-rays): the users can select the wavelength required for their experiment.
2. High Flux: high intensity photon beam allows rapid experiments or use of weakly scattering crystals.
3. High Brilliance: highly collimated photon beam generated by a small divergence and small size source (spatial coherence)
4. High Stability: submicron source stability
5. Polarization: both linear and circular

Super massive black holes and pulsar have been suggested for producing synchrotron radiation.

5.3 S-XRD

To optimize the surface signal, grazing incidence is used in S-XRD since it gives an enormous reduction of the background by reduction of penetration into the bulk or limited access to the detector of the beams diffracted by bulk electrons. Moreover, one can take advantage of the maximum reflectivity of the surface when the incidence angle tends to the critical angle of total reflection.

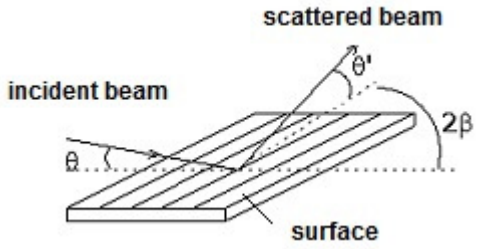

Fig. 5.2 X ray diffraction at grazing angle

The Fresnel relations for the reflectance are:

$$R_p = \left| \frac{n\cos\theta_i - \cos\theta_t}{n\cos\theta_i + \cos\theta_t} \right|^2 \qquad (5.2)$$

for p polarized waves and

$$R_s = \left| \frac{\cos\theta_i - n\cos\theta_t}{\cos\theta_i + n\cos\theta_t} \right|^2 \qquad (5.3)$$

for s polarized waves. When $\theta_t \to \pi/2$ then $R_p=R_s=1$.

Compton showed that the refraction index for X-ray in metals and semiconductors is $n = 1-\delta$ where $\delta \approx 10^{-6}$. The critical glancing angle can be deduced as:

$$\cos\theta_c = n\sin\frac{\pi}{2} = 1-\delta \Rightarrow 1-\frac{\theta_c^2}{2} = 1-\delta \Rightarrow \theta_c \cong \sqrt{2\delta}\ , \qquad (5.4)$$

which means 0.1° up to 0.6° for most materials. Some experimentalists prefer to work at *glancing* angles sufficiently larger than θ_c so the measurements cannot be affected by the rapid variations of the reflectivity around θ_c. Generally they are working at 1°.

X-ray are weakly scattered by the matter so the single scattering approximation is well justified. For this reason in the deduction of the intensity of the scattered ray we will consider a finite single monolayer of atoms with two-dimensional periodicity defined by the translations vectors $\vec{a}$ şi $\vec{b}$.

The incident X-ray can be treated as a plane wave $E_0 e^{i\vec{k}\vec{\rho}}$. The wave diffracted in the point of coordinate $\vec{\rho}$ contributes in the point P, placed at distance R, great compared to the crystal dimensions, with an electric field having the intensity

$$E_{imp} = CE_0 \frac{e^{ikR}}{R} e^{-i\vec{\rho}\Delta\vec{k}} \quad . \tag{5.5}$$

Demonstration:

$$E_{imp} = CE(\rho)\frac{e^{ikr}}{r} = CE_0 e^{i\vec{k}\vec{\rho}} \frac{e^{i(kr-\omega t)}}{r} \tag{5.6}$$

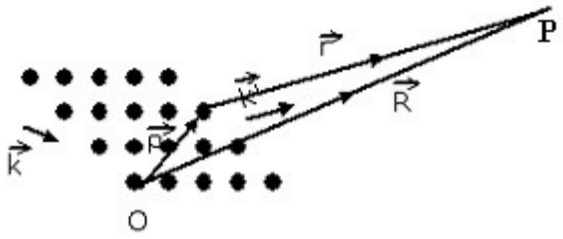

Fig. 5.3 Diffraction of a plane wave

where E_{imp} is the modulus of the of the electric field intensity vector of the scattered wave written as the product of the incident plane wave in $\vec{\rho}$ and the elementary spherical wave generated in that point following the Huygens principle. The amplitude measured in P depends on the phase factor as can be observed in Fig. 5.3. Here C is a proportionality constant that includes the characteristics of the scattering center. The factor *1/r* ensures flow conservation.

$$e^{i(\vec{k}\vec{\rho}+kr)} = e^{i(\vec{k}\vec{\rho}+k(R-\rho\cos(\rho,R)))} = e^{ikR} e^{i(\vec{k}\vec{\rho}-k\rho\cos(\rho,R))} = e^{ikR} e^{i(\vec{k}\vec{\rho}-\vec{k}'\vec{\rho})} = e^{ikR-i\vec{\rho}\Delta\vec{k}} \tag{5.7}$$

because $\vec{r} = \vec{R} - \vec{\rho}$ so that $r = R\sqrt{1 - 2\rho/R\cos(\rho,R) + \rho^2} \cong R(1 - \rho/R\cos(\rho,R))$. Also $\vec{k}' \| \vec{r}$ and $k' = k = 2\pi/\lambda$ so that $k'\rho\cos(\rho,R) = \vec{k}'\vec{\rho}$. At large distances *r* can be replaced by R.

The intensity of the electric field vector of the scattered electromagnetic wave can be computed by adding the waves scattered in each point by the electron distribution in that point:

$$E_{tot} = \int E_{imp} n(\vec{\rho}) dV = CE_0 \frac{e^{ikR}}{R} \int n(\vec{\rho}) e^{-i\vec{\rho}\Delta\vec{k}} dV \tag{5.8}$$

where $n(\vec{\rho})$ is the electron density in the point of coordinate $\vec{\rho}$ due to a basis of s atoms associated to that point. Generally, $n(\vec{\rho})$ can be written as:

$$n(\vec{\rho}) = \sum_{n_1,n_2} \sum_{j=1}^{s} C_j \left(\vec{\rho} - \vec{\rho}_j - \vec{\rho}_{n_1 n_2}\right) \tag{5.9}$$

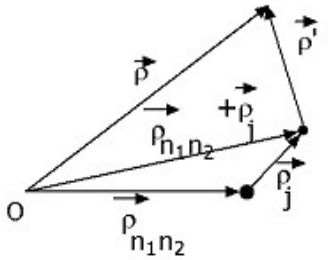

Fig. 5.4

where $\vec{\rho}_{n_1 n_2}$ is the coordinate of the lattice point indexed with $n_1 n_2$, and $\vec{\rho}_j$ is the coordinate of the atom j, from the basis of s atoms. It is obvious that in $n(\vec{\rho})$ contributes preponderantly the atoms closest to $\vec{\rho}$. The expression $\int n(\vec{\rho}) e^{-i\vec{\rho}\Delta\vec{k}} dV$ is the form factor of the atomic basis

associated to each lattice point. From Fig. 5.4 one observes that $\vec{\rho} = \vec{\rho}' + \vec{\rho}_j + \vec{\rho}_{n_1 n_2}$, so that (5.9) can be rewritten as:

$$\begin{aligned}\int n(\vec{\rho}) e^{-i\vec{\rho}\Delta\vec{k}} dV &= \int \sum_{n_1,n_2} \sum_{j=1}^{s} C_j\left(\vec{\rho} - \vec{\rho}_j - \vec{\rho}_{n_1 n_2}\right) e^{-i\vec{\rho}\Delta\vec{k}} dV = \\ &= \sum_{n_1,n_2} \sum_{j=1}^{s} e^{-i\vec{\rho}_j \Delta\vec{k}} e^{-i\vec{\rho}_{n_1 n_2}\Delta\vec{k}} \int C_j(\vec{\rho}') e^{-i\vec{\rho}'\Delta\vec{k}} dV = \\ &= \sum_{j,n_1,n_2} f_j\left(\Delta\vec{k}\right) e^{-i\vec{\rho}_j\Delta\vec{k}} e^{-i\vec{\rho}_{n_1 n_2}\Delta\vec{k}} = F\left(\Delta\vec{k}\right) \sum_{n_1,n_2} e^{-i\vec{\rho}_{n_1 n_2}\Delta\vec{k}}\end{aligned} \tag{5.10}$$

where

$$f_j\left(\Delta\vec{k}\right) = \int C_j(\vec{\rho}') e^{-i\vec{\rho}'\Delta\vec{k}} dV \tag{5.11}$$

is the atomic form factor for each atom j of the basis and

$$F\left(\Delta\vec{k}\right) = \sum_{j,n_1,n_2} f_j\left(\Delta\vec{k}\right) e^{-i\vec{\rho}_j\Delta\vec{k}} \tag{5.12}$$

is the structure factor of the basis that does not depend on the choice of the primitive cell. Also the values for atomic form factor may be found in International tables for X-ray crystallography, vol.3. It was found that the overall distribution of electrons in a crystal as seen in X-ray diffraction is fairly close to that of the free atoms. Therefore, the form factors are not very sensitive to the small redistribution of the electrons in solid.

Replacing (5.8) in (5.6) one obtains

$$E_{tot} = CE_0 \frac{e^{ikR}}{R} F\left(\Delta\vec{k}\right) \sum_{n_1,n_2} e^{-i\vec{\rho}_{n_1 n_2}\Delta\vec{k}} \quad . \tag{5.13}$$

The intensity of the scattered beam in point P is

$$I \cong E_{tot} E_{tot}^{*} \quad , \tag{5.14}$$

Using (5.13), (5.14) becomes:

$$\begin{aligned}I &= const * \frac{E_0^2}{R^2} \left|F\left(\Delta\vec{k}\right)\right|^2 \left|\sum_{n_1=1}^{N_a} \sum_{n_2=1}^{N_b} \exp\left(-i\left(n_1\vec{a} + n_2\vec{b}\right)\Delta\vec{k}\right)\right|^2 = \\ &= const * \frac{E_0^2}{R^2} \left|F\left(\Delta\vec{k}\right)\right|^2 \left|\sum_{n_1=1}^{N_a} \exp\left(-in_1\vec{a}\Delta\vec{k}\right)\right|^2 \left|\sum_{n_2=1}^{N_b} \exp\left(-in_2\vec{b}\Delta\vec{k}\right)\right|^2 .\end{aligned} \tag{5.15}$$

The sums can be easily calculated being geometric series.

$$\begin{aligned}&\sum_{n_1=1}^{N_a} \exp\left(-in_1\vec{a}\Delta\vec{k}\right) = e^{-i\vec{a}\Delta\vec{k}} + e^{-2i\vec{a}\Delta\vec{k}} + e^{-3i\vec{a}\Delta\vec{k}} + ... + e^{-Ni\vec{a}\Delta\vec{k}} = \\ &e^{-i\vec{a}\Delta\vec{k}} \frac{1 - e^{-Ni\vec{a}\Delta\vec{k}}}{1 - e^{-i\vec{a}\Delta\vec{k}}} = e^{-i\vec{a}\Delta\vec{k}} \frac{e^{-i\frac{N\vec{a}\Delta\vec{k}}{2}}}{e^{-i\frac{\vec{a}\Delta\vec{k}}{2}}} \frac{\sin\frac{N\vec{a}\Delta\vec{k}}{2}}{\sin\frac{\vec{a}\Delta\vec{k}}{2}}\end{aligned} \tag{5.16}$$

The intensity of the electromagnetic wave in the observation point P can be written as:

$$I = const * \frac{E_0^2}{R^2} \left| F\left(\Delta\vec{k}\right) \right|^2 \frac{\sin^2 \frac{N\vec{a}\Delta\vec{k}}{2}}{\sin^2 \frac{\vec{a}\Delta\vec{k}}{2}} \frac{\sin^2 \frac{N\vec{b}\Delta\vec{k}}{2}}{\sin^2 \frac{\vec{b}\Delta\vec{k}}{2}} . \qquad (5.17)$$

The intensity (5.17) is maximum if $\vec{a}\Delta\vec{k} = 2\pi l_1$, $\vec{b}\Delta\vec{k} = 2\pi l_2$ i. e. the Laue conditions for the surface. In this case the solution is $\Delta\vec{k}_{\parallel} = l_1\vec{A} + l_2\vec{B} = \vec{G}$ which represents a condition of momentum conservation, because it can be written $\hbar\Delta\vec{k}_{\parallel} = \hbar\vec{G}$. So for the peak maximum one obtains:

$$\Delta\vec{k} = l_1\vec{A} + l_2\vec{B} + \Delta\vec{k}_z = \vec{G} + \Delta\vec{k}_z \quad . \qquad (5.18)$$

which is a plane momentum conservation and not a space momentum conservation law $\Delta\vec{k} = \vec{G}$. The wave vectors of the incident and reflected waves are equal in module.

$$k'^2 = k'^2_z + \left(k_{\parallel} + G\right)^2 = k_z^2 + k_{\parallel}^2 = k^2 \qquad (5.19)$$

$\Delta k_z = k'_z - k_z$ has a continuum variation. Consequently a diffraction spot moves along a diffraction direction when the incident angle varies. The reciprocal lattice consists therefore of rods instead of points. We call these Bragg rods.

The peak intensity now becomes:

$$I\left(l_1, l_2, \Delta k_z\right) = const * \frac{E_0^2}{R^2} \left| F\left(l_1, l_2, \Delta k_z\right) \right|^2 N_a^2 N_b^2 \quad . \qquad (5.20)$$

The structure factor is calculated with the l_1, l_2 indices for the in-plane component of the momentum transfer but depends slightly on Δk_z, which is determined by the acceptance angle of the detector. The image observed on a screen consists of a series of spots for each value ($l_1 l_2$). The image formed can be described using Ewald construction shown below.

Intensity (5.20) can be seen only if a perfect crystal is exposed to a monochromatic and perfectly collimated X-ray beam which is collected by a detector with small angle (well defined $\vec{k}$, well established R). Since in practice these requirements are difficult to achieve, the intensity is obtained by rotation around the surface normal, using a detector having a large opening angle. The result is an integrated intensity that depends also on $\left|F\left(l_1 l_2\right)\right|^2$. From the intensity of each spot $\left|F\left(l_1 l_2\right)\right|^2$ is inferred, and by comparison with theoretical calculations based on structural models, one obtains information about the atomic structure.

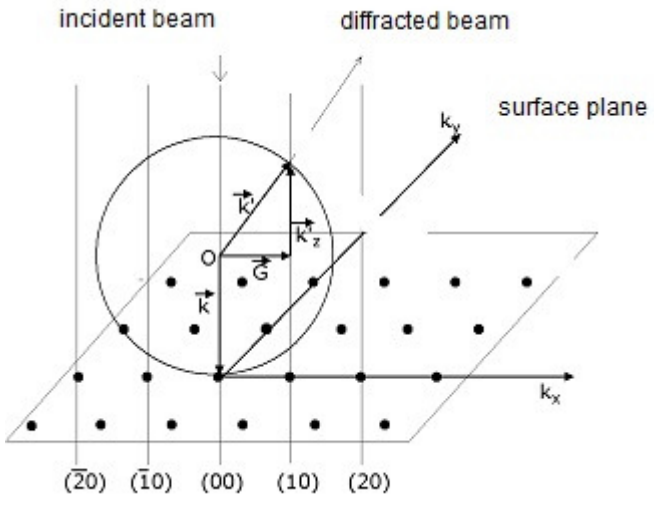

Fig. 5.5 The Ewald constructuction of the diffracted beam at normal incidence

6. LEED

6.1 Introduction

In 1920, Clinton Davisson researcher at the Research Laboratories of the American Telephone and Telegraph Company and the Western Electric Company (later Bell Laboratories) questioned about "What is the nature of secondary electron emission from grids and plates subjected to electron bombardment?" During 1921-1923, Davisson and his younger assistant Charles Henry Kunsman performed experiments on Ni, Pt and Mg foils. Soon, they observed on Ni an unexpected phenomenon that was to have crucial importance for their future experimental program: A small percentage (about 1%) of the incident electron beam was being scattered back elastically toward the electron gun.

Previous observers had noticed this effect for low-energy electrons (about 10 eV), but none had reported it for electrons of energies over 100 eV. To Davisson these elastically scattered electrons appeared as ideal probes with which to examine the extranuclear structure of the atom. After two months of experimentation, Davisson and Kunsman submitted a paper to *Science,* in which they presented a typical curve of their data, proposed a shell model of the atom for interpreting these results, and offered a formula for the quantitative prediction of the implications of the model. In the next two years they built several new tubes, tried five other metals (in addition to nickel) as targets, developed rather sophisticated experimental techniques at high vacuum ("the pressure became less than could be measured, i.e., less than 10^{-8} mm Hg,") and made valiant theoretical attempts to account for the observed scattering intensities. The results were uniformly unimpressive so the project was abandoned. In 1925, Davisson and Lester Germer restarted the same kind of experiments when the glass tube had broken during a scattering experiment and the nickel target got badly oxidized. The method of reducing the oxide on the nickel target by prolonged heating in vacuum and hydrogen produced a polycrystalline form of the nickel target and led them to new unexpected results (see fig. 5.1). Finally, when they realized what happened, they concluded that it was the arrangement of the atoms in the crystals, not the structure of the atoms, which was responsible for the new intensity pattern of the scattered electrons. In 1926, at the Oxford meeting of the British Association for the Advancement of Science, Davisson heard a lecture by Max Born in which his own and Kunsman's (platinum-target) curves of 1923 were cited as confirmatory evidence for de Broglie's electron waves. The idea of connecting the minima and maxima in the intensity patterns of scattered electrons with Louis de Broglie matter-wave hypothesis was first emitted by Walter Elsasser a student at the Franck's institute in Göttingen where Max Born was a professor at that time. (Elsasser is considered a "father" of the presently accepted dynamo theory as an explanation of the Earth's magnetism. He proposed that this magnetic field is resulted from electric currents induced in the fluid outer core of the Earth.)

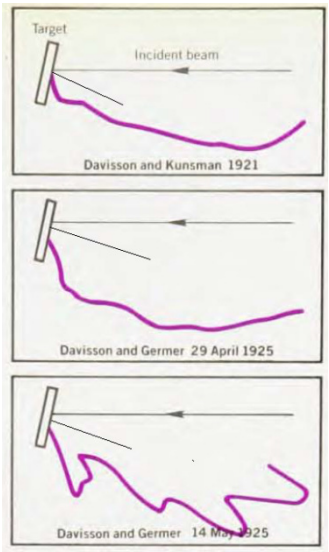

Fig. 6.1 Typical curves obtained by Davisson and Kunsman on Ni foils.

At the same time George Paget Thomson (1892-1975) conducted a different experiment and also proved the wave structure of electrons. The thin (about 100nm) gold foil was bombarding by a beam of electrons of a high velocity (of the energy of about 10^4 eV). The electrons went through the foil and gave rise diffraction rings on the screen behind the foil. The rings were the proof for the wave structure of the electrons.

The Nobel Prize in Physics 1937 was awarded jointly to Clinton Joseph Davisson and George Paget Thomson *"for their experimental discovery of the diffraction of electrons by crystals"*. Davisson retired from Bell Labs in 1946 but Germer regained his interest in low-energy electron diffraction in 1959-60. With this work Germer was able to follow up with great success the study of surfaces. The field of low-energy electron diffraction (LEED) is now widespread and very active.

5.2 LEED experimental apparatus

In LEED, the electrons are produced in a well collimated monoenergetic beam by an electron gun. Inside this gun is a heated cathode acting as an electron source. A grid that can be negatively biased by up to 30V controls current from the gun. There follow three anodes at adjustable potentials so as to focus the electrons coming out of the electron gun. Using this configuration the gun can be made to deliver more current. Beams are parallel to within 1°. The crystal to be studied is mounted at the center of a system of hemispherical grids (see fig. 5.2). The simplest arrangement has only two grids. G_1 that is connected to the ground so that once electrons are inside the region enclosed by G_1, they are in the field free region. Electrons are diffracted from the surface of the crystal and speed towards the grids. G_2 is negatively biased so that only electrons having lost less than 1-2 eV of their original energy can get through G_2 (typically only 2% from the emergent electrons). All others are turned back and collected by G_1. Finally the screen is positively biased with a large accelerating potential (about 2kV) to give the electrons enough energy to excite the phosphor efficiently. The assembly must be mounted inside an ultra high vacuum system with pressure of the order10^{-9} – 10^{-10} torr. Only at these pressures can the crystal surface be preserved from contamination by residual gases. Assuming a sticking probability of unity for molecules incident on the crystal surface, it takes of the order of a day to develop a monolayer of adsorbed gases in a vacuum of this quality.

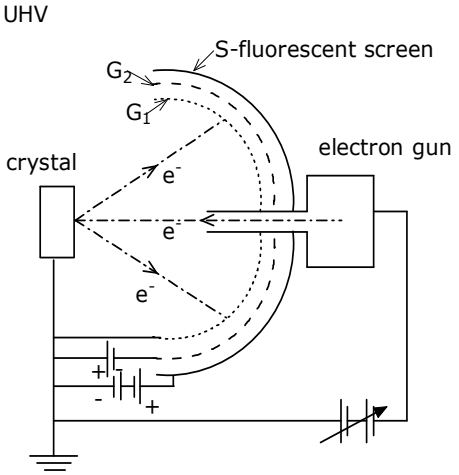

Fig. 6.2 LEED experimental apparatus

The cathode is made of tungsten heated at 2500K or a bariated material heated at only 1000K. It is this temperature that determines the basic energy resolution of the equipment because no energy filtering is applied to the emitted electrons. They have a thermal spread in energy of the order 3kT/2 or ±0.13 eV at 1000K and ±0.3 eV at 2500K. The electron gun has dimensions of the order of 10 cm/2cm. The beam is about 0.1 cm in diameter and carries a current of the order 1μA, but both quantities vary with the energy range in which the gun is operating. The energy of the electrons is chosen such that the de Broglie wavelength is comparable with the lattice constant, about 1-5Å. The electrons energies in LEED are between 20-500 eV.

The energy of the electrons is related to the de Broglie wavelength by:

$$E = \frac{p^2}{2m} = \frac{h^2}{2m\lambda^2}. \qquad (6.1)$$

For λ = 1 A° the energy is150 eV.

The crystal surface on which the experiments are to be done is usually cut outside the chamber. The proper orientation of the crystal is achieved by X-rays methods and is usually accurate to within 1°. Mechanical polishing with carbide powders and chemical etching follows, to make the surface optically smooth on the order of the wavelength of light. When the crystal goes into the diffraction chamber it almost has much surface damage and several layers of foreign molecules (oxygen, water etc.) adsorbed on the surface.

The cleaning problem is a crucial one. Two facilities are provided for cleaning inside the chamber. The specimen can be heated to a high temperature just below its melting point, as a means of removing impurities. A mass spectrometer can be used to analyze gases flashed from the surface, giving an idea of impurities present. In addition, a source of 500 eV argon ions can be used to bombard the surface, followed by flashing to high temperatures to remove impurities or any of the bombarding ions embedded in the surface. Finally the crystal is annealed to allow the atoms to arrange themselves into a well ordered surface. During the process of preparation, it is important to be able to monitor the condition of the surface. The diffraction patterns may be used to detect disorder, faceting, impurities that form a new structure (but not those impurities that adsorb without changing the basic symmetry of the surface). The next sophistication is to monitor intensities of the diffracted beams. The surface is cleaned until intensities become stable.

It could happen that impurities diffuse to the surface from the bulk. This is an effectively infinite source. Thus repeating the cleaning cycle would not change the pattern. Difficulties of this nature require some direct monitoring of surface impurities and modern equipment has an Auger facility which can detect and identify impurities on the surface. By using both the diffraction pattern and the Auger facility it is possible to do experiments in which one is reasonably certain of an atomically clean and well ordered surface.

Two types of information can be extracted with LEED:

- The diffraction patterns obtained on the fluorescent screen give information about the cleanliness, order and geometry of the surface;
- Additional information such as separation between planes of atoms near the surface or location of adsorbates relative to the substrate can be obtained by measuring the intensities of elastically scattered electron beams function of energy.

5.3 The diffraction pattern

An ideal beam of electrons infinitely wide would be described by a plane wave $B\exp\left(i\vec{k}\vec{r}\right)$ having a well definite direction and a well definite energy $E=\frac{\hbar^2k^2}{2m}$. In practice, neither the direction nor the energy are defined precisely and the beam really consists of a whole set of waves having slightly different directions and energy. The result is that whereas a simple plane wave has a definite phase at all parts of the surface on which it is incident, the more complicated beam varies in phase in a manner that is not entirely predictable because of uncertainties in the wave vector $\vec{k}$ direction. For variations of phase over a short distance on the crystal surface, these uncertainties are not important but over large distances they become so. There is a characteristic length called the coherence length such that atoms in the surface within a coherence length of one another can be thought of as illuminated by the simple wave and amplitudes are to be added. Atoms further apart must be regarded as illuminated by waves whose phase-

relationship is arbitrary. Then the intensities are added. Thus no surface structure on a scale larger than the coherence length forms a diffraction pattern.

We shall consider a monoenergetic beam, well collimated (the wave vector $\vec{k}$ direction well defined) incident on a perfectly clean, well ordered surface. The whole surface is considered to be contained in the coherence area.

The incident electron has a wave function:

$$\psi_i(r) = B\exp\left(i\vec{k}_0\vec{r}\right). \tag{6.2}$$

Our interest will center on the elastically scattered electrons, because they produce all the structure on the diffraction patterns. The motion of these electrons will be taken to be determinated by the Schrödinger equation:

$$-\frac{\hbar^2}{2m}\nabla^2\psi_T(\vec{r}) + V(\vec{r})\psi_T(\vec{r}) = E\psi_T(\vec{r}) \tag{6.3}$$

where $\psi_T(\vec{r}) = \psi_i(\vec{r}) + \psi_s(\vec{r})$ is the total wave function (incident+scattered). We want to discover what properties the scattered wave function ψ_s has in consequence of the surface symmetry. The surface has two-dimensional periodicity. Since the potential $V(\vec{r})$ acting on the electron is due to the crystal, it will have the symmetry of the surface:

$$V\left(\vec{r} + n_1\vec{a} + n_2\vec{b}\right) = V(\vec{r}). \tag{6.4}$$

Substituting in the Schrödinger equation (6.3), $\vec{r} = \vec{r}' + n_1\vec{a} + n_2\vec{b}$ and using

$$\nabla^2\psi = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} = \frac{\partial^2\psi}{\partial x'^2} + \frac{\partial^2\psi}{\partial y'^2} + \frac{\partial^2\psi}{\partial z'^2} = \nabla'^2\psi \tag{6.5}$$

because $\frac{\partial\psi}{\partial x'} = \frac{\partial\psi}{\partial x}\frac{\partial\psi}{\partial x'} = \frac{\partial\psi}{\partial x}$ etc., one obtains

$$-\frac{\hbar^2}{2m}\nabla'^2\psi_T\left(\vec{r}' + n_1\vec{a} + n_2\vec{b}\right) + V(\vec{r}')\psi_T\left(\vec{r}' + n_1\vec{a} + n_2\vec{b}\right) = E\psi_T\left(\vec{r}' + n_1\vec{a} + n_2\vec{b}\right) \tag{6.6}$$

In other words $\psi_T\left(\vec{r}' + n_1\vec{a} + n_2\vec{b}\right)$ obeys the same equation as $\psi_T(\vec{r})$. The incident wave component $\psi_i(\vec{r})$, will change when performing the transformation $\vec{r} = \vec{r}' + n_1\vec{a} + n_2\vec{b}$

$$\begin{aligned}\psi_i\left(\vec{r}' + n_1\vec{a} + n_2\vec{b}\right) &= B\exp\left(i\vec{k}_0\vec{r}'\right)\exp\left(i\vec{k}_{0\|}n_1\vec{a} + i\vec{k}_{0\|}n_2\vec{b}\right) = \\ &= \psi_i(\vec{r}')\exp\left(i\vec{k}_{0\|}\left(n_1\vec{a} + n_2\vec{b}\right)\right)\end{aligned} \tag{6.7}$$

where $\vec{k}_{0\|}$ is the $\vec{k}_0$ component, parallel to the surface. In this situation we are beaming onto the crystal an incident wave differing by a phase factor $\exp\left(i\vec{k}_{0\|}\left(n_1\vec{a} + n_2\vec{b}\right)\right)$. The amplitude of the reflected wave $\psi_s(\vec{r})$, that is always proportional to that of the incident wave $\psi_i(\vec{r})$, will be different from the old one by the same phase factor:

$$\psi_s\left(\vec{r}' + n_1\vec{a} + n_2\vec{b}\right) = \psi_s(\vec{r}')\exp\left(i\vec{k}_{0\|}\left(n_1\vec{a} + n_2\vec{b}\right)\right). \tag{6.8}$$

Equation (6.8) shows that the scattered wave obeys what is known as the Bloch theorem adapted to the surface. The form of the above equation leads us to write $\psi_s(\vec{r})$ as:

$$\psi_s(\vec{r}) = \exp\left(i\vec{k}_{0\|}\vec{r}_{\|}\right)\chi_s(\vec{r}). \tag{6.9}$$

Introducing (6.9) in (6.8)

$$\begin{aligned}&\psi_s\left(\vec{r}'+n_1\vec{a}+n_2\vec{b}\right) = \exp\left(i\vec{k}_{0\|}\vec{r}'_{\|}+i\vec{k}_{0\|}\left(n_1\vec{a}+n_2\vec{b}\right)\right)\chi_s\left(\vec{r}'+n_1\vec{a}+n_2\vec{b}\right) = \\ &\psi_s(\vec{r}')\exp\left(i\vec{k}_{0\|}\left(n_1\vec{a}+n_2\vec{b}\right)\right) = \exp\left(i\vec{k}_{0\|}\vec{r}'_{\|}\right)\chi_s(\vec{r}')\exp\left(i\vec{k}_{0\|}\left(n_1\vec{a}+n_2\vec{b}\right)\right)\end{aligned}. \tag{6.10}$$

The function has the periodic property of the potential

$$\chi_s\left(\vec{r}'+n_1\vec{a}+n_2\vec{b}\right) = \chi_s(\vec{r}'). \tag{6.11}$$

In virtue of its periodicity, we can write χ_s, as a Fourier expansion:

$$\chi_s(\vec{r}) = \sum_G \alpha_G(z)\exp\left[i\vec{G}\vec{r}_{\|}\right] \tag{6.12}$$

where $\vec{G}$ are the vectors of the reciprocal lattice. This means that ψ_s can be also written as Fourier series:

$$\psi_s(\vec{r}) = \sum_G \alpha_G(z)\exp\left[i\left(\vec{k}_{0\|}+\vec{G}\right)\vec{r}_{\|}\right]. \tag{6.13}$$

Equation (6.13) holds quite generally, but outside the crystal an even simpler expression may be found, because there the crystal potential is zero. Substituting for the total wave function ψ_T into the Schrödinger equation far from the surface:

$$\begin{aligned}&-\frac{\hbar^2}{2m}\nabla^2\left(B\exp\left(i\vec{k}_0\vec{r}\right)+\sum_G \alpha_G(z)\exp\left[i\left(\vec{k}_{0\|}+\vec{G}\right)\vec{r}_{\|}\right]\right) = \\ &E\left(B\exp\left(i\vec{k}_0\vec{r}\right)+\sum_G \alpha_G(z)\exp\left[i\left(\vec{k}_{0\|}+\vec{G}\right)\vec{r}_{\|}\right]\right)\end{aligned}. \tag{6.14}$$

Performing the derivatives:

$$\begin{aligned}&\frac{\hbar^2}{2m}k_0^2 B\exp\left(i\vec{k}_0\vec{r}\right)+\sum_G \frac{\hbar^2}{2m}\left[\left|\vec{k}_{0\|}+\vec{G}\right|^2\alpha_G-\frac{d^2\alpha_G}{dz^2}\right]\exp\left[i\left(\vec{k}_{0\|}+\vec{G}\right)\vec{r}_{\|}\right] = \\ &EB\exp\left(i\vec{k}_0\vec{r}\right)+E\sum_G \alpha_G(z)\exp\left[i\left(\vec{k}_{0\|}+\vec{G}\right)\vec{r}_{\|}\right]\end{aligned}. \tag{6.15}$$

Since the scattering is elastic, the energy is conserved $E=\frac{\hbar^2 k_0^2}{2m}$, and eq. (6.15) becomes after some rearrangements:

$$\sum_G\left[\left(\left|\vec{k}_{0\|}+\vec{G}\right|^2-\frac{2mE}{\hbar^2}\right)\alpha_G-\frac{d^2\alpha_G}{dz^2}\right]\exp\left[i\left(\vec{k}_{0\|}+\vec{G}\right)\vec{r}_{\|}\right]=0. \tag{6.16}$$

The above sum is zero only if each term is zero

$$\left(\left|\vec{k}_{0\|}+\vec{G}\right|^2-\frac{2mE}{\hbar^2}\right)\alpha_G-\frac{d^2\alpha_G}{dz^2}=0, \tag{6.17}$$

thus the $\alpha_G(z)$ functions take the form:

$$\alpha_G(z) = \beta_G \exp\left(-iz\sqrt{\frac{2mE}{\hbar^2} - \left|\vec{k}_{0\|} + \vec{G}\right|^2} \right) \tag{6.18}$$

for each $\vec{G}$. β_G is a constant. The minus sign indicates a wave that travels away from the crystal. The scattered wave outside the crystal can now be written as:

$$\psi_s(\vec{r}) = \sum_G \beta_G \exp\left(i\vec{K}_G^- \vec{r}\right) \tag{6.19}$$

where

$$K_G^- = \left(k_{0x} + G_x, k_{0y} + G_y, -\sqrt{\frac{2mE}{\hbar^2} - \left|\vec{k}_{0\|} + \vec{G}\right|^2} \right). \tag{6.20}$$

It follows from quite general considerations, independent of the detailed mechanism of scattering, that the diffracted wavefield has the form of a series of discrete beams, each with a different parallel component of momentum $\hbar\left(\vec{k}_{0\|} + \vec{G}\right)$. If for certain vectors $\vec{G}$,

$$\left|\vec{k}_{0\|} + \vec{G}\right|^2 = \frac{2mE}{\hbar^2}, \tag{6.21}$$

then the wave has $k_z = 0$ and propagates parallel to the surface. For $\vec{G}$ values greater than the critical value, the z exponential becomes complex, implying that the beam becomes evanescent, dying away exponentially in amplitude within few angstroms. Such beams can never be observed on the screen.

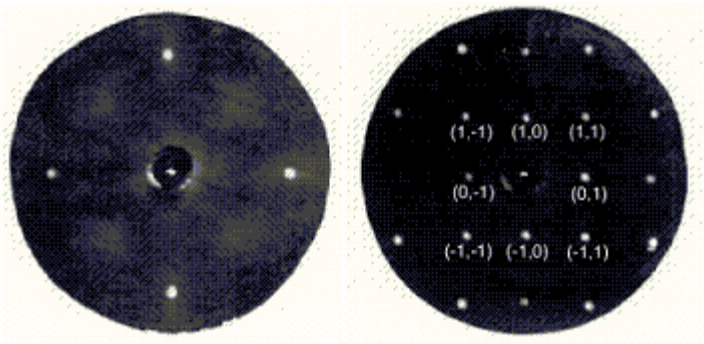

Fig. 5.3 LEED image for W(100) at 45 eV electron kinetic energy (left) and 145 eV (right).

As the incident energy is increased, more and more Fourier components produce beams that can reach the screen. The pattern seen on the screen contracts to allow room for the larger number of spots.

For a given $\vec{k}_{0\|}$ and E, $\vec{G}$ determines the diffraction pattern: there is one spot on the screen for each $\vec{G}$. By making the screen spherical and placing the crystal at the center of the sphere, the spot pattern provide us with a direct picture of the reciprocal lattice. Use is made of this fact on naming the spots. The spot produced by a beam with parallel component of momentum $\hbar\left(\vec{k}_{0\|} + \vec{G}\right)$, where $\vec{G} = l_1\vec{A} + l_2\vec{B}$, is referred as the ($l_1$, l_2) spot.

The diffraction pattern gives us the reciprocal lattice of the surface and hence $\vec{A}$ and $\vec{B}$, the vectors of the primitive reciprocal cell. Knowing $\vec{A}$ and $\vec{B}$ one can get the vectors of the primitive real cell:

$$\vec{a} = (a_x, a_y) = \frac{2\pi}{A_x B_y - A_y B_x}(B_y, -B_x)$$
$$\vec{b} = (b_x, b_y) = \frac{2\pi}{A_x B_y - A_y B_x}(-A_y, A_x) \quad . \tag{6.22}$$

The elastically scattered electrons amount to ~2% of the incident electrons. Many of the other electrons suffer inelastic collisions and cascade down in energy. The dominant loss mechanism in most material is by plasmon creation (bulk or surface plasmons). Before an electron can excite plasmons they must have a minimum energy which is about 10 eV for surface plasmons and about 15 eV for bulk plasmons. Above these energies a strong attenuation of the electron beam hold. Energies comparable with those of the inelastically scattered electrons can be transferred to electrons below the Fermi level giving rise to so called secondary electrons. Another process is the excitation of deep core levels. If an electron in a deep core level of energy E_1 is excited leaving behind a hole, another electron from a higher energy state E_2 will probably make a transition to the lower level. The energy difference can be emitted as an X ray but more probably given up to a third electron having initially energy E_3. The energy balance means that finally the third electron has energy $E_3+E_2-E_1$, not related to the incident electron energy, possibly sufficient to escape from the crystal.

The incident electrons can lose energy by scattering of a phonon. Energy losses to phonons are less than 0.1 eV. Electrons with such a small loss cannot be filtered out by the retarding grid G2 and are usually called quasi-elastic scattered electrons. At low temperatures a halo of low momentum –phonon loss electrons forms around the elastic beam. At very high temperatures when atoms have large displacements, each atom behave as an independent scatterer and a series of dark ring can be seen in the background intensity correlating with minima in the atomic scattering factors.

References

1. Ch. Kittel, *Introducere în fizica corpului solid*, Ed. Tehnică, Bucureşti, 1972.
2. M. C. Desjonquères, D. Spanjard, *Concepts in surface physics*, Springer-Verlag, Heidelberg, 1993.
3. J. B. Pendry, *Low energy electron diffraction*, Academic Press, London and New York, 1974.
4. E. W. Müller, T. T. Tsong, *Field ion microscopy*, American Elsevier Publishing Company, New York, 1969.
5. C. J. Chen, *Introduction to scanning tunneling microscopy*, Oxford University Press, New York, 1993.
6. STM Gallery, Institut für Allgemeine Physik, TU Wien, www.iap.tuwien.ac.at/www/surface/STM_Gallery.
7. IBM Almaden Research Center Visualization Lab Home Page www.almaden.ibm.com/vis/stm/gallery.html.
8. M. A. Van Hove, S. Y. Tong, *Surface crystallography by LEED*, Springer-Verlag, Berlin Heidelberg, 1979.
9. A. Zangwill, *Physics at surfaces*, Cambridge University Press , 1988.
10. N. W. Ashcroft, N. D. Mermin, *Solid state physics*, Holt-Saunders Japan, LTD , 1981.
11. T. A. Delchar, and D. P. Woodruff, *Modern Techniques of Surface Science*, Cambridge Solid State Science Series, 1990.
12. H. Ibach, *Physics of surfaces and interfaces*, Springer Verlag, Berlin Heidelberg, 2006.