Chapter 2
Introduction to the Physics of Field Ion Emitters

2.1 High Electric Field Nanoscience

Explanations of field ion microscopy and atom probe use several branches of science. In places, these need customizing in order to describe how high electric fields affect atomic-scale processes. The resulting theory forms a specialist scientific focus that is becoming known as high electric field nanoscience (HEFNS).

HEFNS has wider relevance, but a major application is to understand the preparation, operation, and failure of field ion and electron emission sources. These sources and the underlying physical processes have vital roles as components of the tools of modern nanotechnology. Aside from the role of field ion emission in FIM and APT, the liquid metal FI source [1] is used in focused ion beam machines [2] and the gas FI source [3] in helium scanning ion microscopy [4]. FI sources have also been proposed for the field emission electric propulsion of spacecraft [5]. Field electron sources [6, 7] are used in high-resolution electron microscopes [8] and the near-field emission scanning electron microscope [9], and have been considered for electronic displays [10]. The basic processes also play a role in electrical vacuum breakdown [11].

Within HEFNS, the following aspects of behavior at highly charged surfaces are part of the science behind FIM and APT:

- Relevant electrostatics and electrical thermodynamics
- Basic theory of charged surfaces
- Electron and ion behavior at charged surfaces
- Field ionization and post-field-ionization
- Atomic and molecular behavior at charged surfaces
- Charged-particle optics of field emitters
- Field evaporation and other aspects of field desorption
- Laser-induced effects
The first six topics are discussed in this chapter, and the remainder in Chap. 3. The general aim is to provide conceptual understanding and some relevant basic formulae. At present, many theoretical details are not fully understood. Some theory is included for historical completeness and/or possible future relevance.

2.2 Basic Electron Potential Energy Models

In field ion emission and atom probe (FI/AP) theory, many phenomena involve the potential energy (PE) of an atom or an ion near a charged surface. However, electron energies near surfaces need to be described first. A central concept is the total potential energy of a hypothetical point electron; this is called here the electron potential energy (EPE) and denoted by $U_e$.

2.2.1 The Sommerfeld Model

2.2.1.1 The Basic Model

For a metal field emitter, in the absence of external fields, the simplest EPE model is the basic Sommerfeld (free-electron) model [12–14], Fig. 2.1. This model disregards atomic structure and assumes a smooth, flat planar surface of large extent. The well base corresponds to the base of the metal conduction band, at EPE $E_C$. The metal surface is modeled as a sharp EPE step of height $\chi$, called the inner PE. Quantum-mechanical electron states are defined in the well, and Fermi–Dirac statistics applied. At 0 K, well states are filled to a total energy level called the Fermi level and denoted by $E_F$; the total kinetic energy $K_F (= E_F - E_C)$ of an electron at the Fermi level is called the Fermi energy.

Electron potential energies and total energies (and hence $E_C$ and $E_F$) can be measured relative to any convenient reference zero. In FI/AP theory, the most convenient reference zeros (depending on the problem under discussion) are the emitter Fermi level and the topmost orbital level of an external atom. Often, it is convenient to treat the emitter Fermi level as at the level of the laboratory Earth.

![Fig. 2.1 The basic Sommerfeld free-electron model](image-url)
In the Sommerfeld model, the work needed (in a slow thermodynamic process) to take an electron from the Fermi level, and place it (stationary) at a position “X” somewhat outside the emitter surface, is called the *local thermodynamic work function* and denoted by $\phi$. The EPE at X is called the *local vacuum level* and denoted by $E_{\text{vac}}$ or $E_o$. Thus

$$E_o = E_F + \phi.$$  

(2.1)

$\phi$ has weak dependences on temperature [15] and on field [16], but these can be neglected in FI/AP theory.

In AP theory, the main use of work functions is as part of the “thermodynamic term” (Sect. 3.2.8) used to estimate evaporation fields. For many metals, $\phi$ has an average value of approximately 4–5 eV, though some metal elements have lower values [see Appendix D]. Different crystallographic faces of a metal element usually have different $\phi$-values, for reasons discussed below. Hence, the local vacuum level varies as between points outside different crystallographic faces [17]. For alloys, allocating work function values can be problematic (Sect. 2.2.2.3).

### 2.2.1.2 Electron PE in an Applied Field and the Electrical Surface

If an electrostatic field $F$ is present, and if in addition the position dependence of exchange-and-correlation effects (Sect. 2.2.1.3) is taken into account, the EPE is as shown in Fig. 2.2. Outside the Sommerfeld model surface, the *electrostatic component* $v_{\text{ES}}$ of EPE is written

$$v_{\text{ES}} = \phi + eFx,$$

(2.2)

where $e$ is the elementary positive charge, and $x$ is the distance from the model surface; $v_{\text{ES}}$ is represented by the dashed line in Fig. 2.2.
In principle, a formula is needed that applies to real atomically structured surfaces. In the ideal case of an infinitely large, perfect crystal face, an equation with the form of Eqn. 2.2 gives correct asymptotic behavior for $v_{ES}$ if $x$ is measured from a reference surface called the electrical surface. This surface is found to be close to the outer edges of the surface atoms (Sect. 2.3.3.4). Distance $x$ measured from the electrical surface is sometimes called the electrical distance.

The Sommerfeld model is made compatible with this approach by locating the Sommerfeld model step at the electrical surface.

### 2.2.1.3 Exchange-and-Correlation Effects

By definition, electrostatic potentials relate to the force experienced by a vanishingly small test charge, which must not disturb the charge distribution that causes the force. A classical point electron does not have a vanishingly small charge. When close to a metal surface, it induces a charge distribution that may be represented in terms of a positive image charge [18–20]. A classical point electron thus experiences an “image force” that attracts it towards the surface; this gives rise to an image PE. For a classical conductor with a flat planar surface, this is given by the Schottky [20] expression $-e^2/16\pi\varepsilon_0 x$, where $\varepsilon_0$ is the electric constant (formerly called the permittivity of free space).

The equivalent quantum-mechanical (QM) effect [17, 21] generates an exchange-and-correlation (XC) component $\omega_{XC}$ in the total EPE $U_e(x)$. For simplicity, $\omega_{XC}$ is often modeled as (and is sometimes described as) an image PE. Hence, for planar emitters, the EPE $U_e(x)$ is often approximated by

$$U_e(x) = \nu_{ES} + \omega_{XC} \approx \phi + eFx - e^2/16\pi\varepsilon_0 x,$$

where the universal constant $e^2/16\pi\varepsilon_0$ has the value 0.3599911 eV nm. This total EPE $U_e(x)$ is depicted by the continuous line in Fig. 2.2. For sufficiently small values of $x$, the image-PE approximation becomes invalid; in reality, $U_e(x)$ joins smoothly to the base of the conduction band, at energy $E_c$.

Strictly, the parameter $x$ in the image PE needs to be measured from a reference surface called the image plane. Lang and Kohn [22] showed that (in simple models) the image plane and the electrical surface coincide. This result is assumed valid for electrons here and for positive ions later.

Note that the energy reference zero for $\omega_{XC}$ is taken at very large distance $x$. This is consistent with using the emitter Fermi level as the reference zero for EPE, because $\phi$ contains the whole XC component of the EPE difference between the inside and outside of the emitter. (There is also an electrostatic component of $\phi$, which is discussed in Sect. 2.2.2.1.)

Image-PE expressions are known for a sphere [23] and a hyperbola [24], but the planar approximation is often used. Sophisticated QM calculations exist for $\omega_{XC}$ (see [21]). But, at present, AP theory is not sufficiently advanced to need anything more detailed than that in Eqn. 2.3.
For semiconductor surfaces in vacuum, both electrostatic effects (Sect. 2.2.4) and XC effects are considerably more complex. Despite its age, [25, p. 122] has one of the best discussions of semiconductor XC effects. There still seems no clear consensus about the best formula to use. The view taken here is that Eqn. 2.3 is adequate for AP theory—particularly if, during ion emission in the laser pulsed atom probe, a semiconductor specimen acts like a good conductor (see Sect. 3.9.2).

2.2.2 Work Function Theory and Related Effects

2.2.2.1 The Physical Origin of Local Work Function

Smoluchowski [26] gave a simple explanation of metal-element \( \phi \)-values. He identified two main contributing effects: a “bulk” or “purely chemical” (i.e., XC) effect, which binds the electron to the metal as a whole, and a “surface” or “electric dipole” effect, which depends on the chemical nature and arrangement of surface atoms in the crystallographic face of interest.

For electrons near a metal surface in vacuum, he identified two competing tendencies. The electrons spread into the vacuum, thereby creating an electric dipole with its negative end outwards, which increases the barrier holding the electron in. Also, the electrons smooth sideways into the gaps between atoms, which has the reverse effect. For a given element, the degree of smoothing depends on the crystallography. Hence, there are differences in surface electric dipole moment, and corresponding differences in local work function, as between different crystallographic faces, and as between different chemical elements.

2.2.2.2 Patch Fields and Contact Electrification

These surface electric dipole differences lead to other small charge rearrangements. As a result, a complicated system of long-range electrostatic fields, known as patch fields [15], will exist outside neutral objects made from a pure metal element, if an object has surface regions ("patches") that exhibit different local work functions. In some places, these fields would be largely parallel to the material surface [27].

These patch fields are thought to have values of order 0.6 V/nm or less [28]; thus, they are small compared with the external fields used by FIM and AP. However, since external fields can cause strong differences in electrically induced surface dipole moment as between different crystallographic faces (Sect. 2.3.3.3), applied voltage-induced patch fields may also exist. This effect has never been investigated in detail.

A further complication is that when materials with different local work functions are in contact, then (as a result of electron thermodynamics) there can be long-range electron transfers between the materials, as well as the formation of a dipole layer at their interface. Contact electrification of this kind leads to excess charges on the
vacuum-facing surfaces of such materials: these charges also generate patch fields. Alloys may display very-short-range effects of the same kind.

A general implication is that electron behavior at metal surfaces (and the derived field distribution) is determined by electron thermodynamics, and not solely by the rules of electrostatics. In particular (contrary to the rules applying to classical conductors), charge redistributions at surfaces—even at clean metal surfaces—may create lateral components in the electrostatic fields at surfaces, particularly near the locations where emitter facets join. Conceivably, these might have some influence on atomic motion at charged surfaces and might be a further factor to be considered in discussion of AP reconstruction.

Patch fields also ensure that if an electron is formally removed from the emitter Fermi level to a large distance away (which needs work $\phi_T$ called here the total work function), then the work done is the same, whichever crystallographic face the electron path intersects. The value of $\phi_T$ depends on the shape, size, and chemical nature of the emitter, and an earlier name “absolute work function” is misleading. In practice, basic formulae in FI/AP theory always need a local work function (not $\phi_T$). Some early FI literature gets this incorrect.

2.2.2.3 Difficulties in Defining Local Work Function

For a local work function to be well defined, position “X” discussed above needs to be outside the range of image-type forces but well inside the range of patch fields. Patch fields are significant out to distances comparable with the size of the related crystal facets; thus, if a specific emitter has facets that are small in size, it may not be possible to allocate a numerically well-defined $\phi$-value to the small facet (although one could state what $\phi$ would be for a large facet of the same crystallographic orientation). Defining precise $\phi$-values for locations near the boundaries of facets can be problematic.

At metal alloy surfaces, chemically induced charge transfers may occur and cause additional local field components. In this case, the concept of local work function becomes problematic. The parameter $\phi$ that goes into field evaporation theory needs to be some sort of average of the work functions of the constituents, and this value may vary locally across the surface. However, for field evaporation of a minority-component atom from a dilute alloy, it should usually be acceptable to use the matrix work function.

2.2.3 Electron Potential Energy with an External Atom Present

When an external atom is present, with nucleus at position $x_N$, simple models assume that a point electron would also see an attractive EPE component due to a
The inclusion of the parameter $z$ allows the nucleus to be allocated an “effective charge” $ze$, which is sometimes theoretically useful. When the nucleus is located in the critical surface (Sect. 2.5.3), then the energy level of the topmost atomic orbital is equal to the emitter Fermi level. This case is illustrated in Fig. 2.3.

**2.2.4 Field Penetration and Band Bending in Semiconductors**

With semiconductors, the situation is much more complicated and is poorly understood in detail, particularly for positively charged small emitters and during laser pulsing. The physics “without pulsing” is that it is energetically difficult for electrons near surfaces to move away and expose positive charge that can screen the impressed field. The impressed field thus penetrates into the surface, sometimes by many tens of nanometers, thereby causing the electron band structure to bend upwards [29, 30], as shown in Fig. 2.4. This effect is known as band bending.

For flat planar semiconductor surfaces of large extent, and at length scales where averaging over nanoscale statistical variations is valid, the theory of static band bending is well understood in principle [29, 30]. For an intrinsic semiconductor
with no surface states (or, alternatively, with no active surface states), the thermo-
dynamic equilibrium response to an impressed positive field is clear. The bands bend upwards until, at the surface, the top edge of the valence band crosses the emitter Fermi level. At this point, holes appear in the elevated valence band, near the surface, and make the field decay with distance into the semiconductor.

Alternatively, if, at the semiconductor surface, there are active surface states located in energy in the band gap, then in zero-field conditions the bands will bend in such a way (usually upwards) as to make the surface electrically neutral. With positive applied field, the thermodynamic equilibrium response is for the surface state bands to lift in energy above the bulk Fermi level, making electrons move out of the surface states. This leaves behind positive charge on surface atoms, which screens the impressed field or part of it. If, for this reason, the impressed field can penetrate only to a limited extent, then the band bending is also limited: this effect is called “Fermi-level pinning,” or (better) band structure pinning (since, normally, it is the Fermi level, rather than the band structure, that stays constant). However, the very high fields used in field evaporation may “exhaust” the capability of the surface states to screen the impressed field.

In principle, the presence of suitable active adsorbates on the surface can quench surface states and make them inactive, but it seems unlikely that any adsorbates would remain on the surface after the evaporation of a few surface layers.

In conventional band-bending theory [30], the degree of bending is also influenced by semiconductor dopant-atom concentrations. But the shape and small size of a field emitter probably mean that surface effects, as described above, dominate for field emitters and that bulk dopant effects can usually be neglected.
In the absence of current flow, the semiconductor-specimen Fermi level \( E_F \) remains equal to that of the metallic support on which it is mounted. An effect of upwards band bending is to lift the local vacuum level \( E_\infty \) (which is at a level above the conduction band edge by the electron affinity \( \chi \)). This increases the operative local work function \( \phi_{op} \) (given by \( E_\infty - E_F \)) above its zero-field value. If \( \phi_{op} \) thereby becomes greater than the “flat-band value” \( \phi_{fb} \) deduced from the bulk band levels, Fig. 2.4, and if empty emitter states just above the Fermi level are available for electrons to tunnel into, then increase in \( \phi_{op} \) above \( \phi_{fb} \) would lower the evaporation field values predicted using \( \phi_{fb} \)-values.

For the AP, the remarks above apply in principle when field evaporation is induced by voltage pulsing—although there are also important issues about whether the electron distribution in a particular semiconductor specimen would be able to respond sufficiently quickly to a fast voltage pulse.

With laser pulsing of semiconductor specimens, further issues arise. Presumed effects of an intense laser pulse are to promote large numbers of electrons from the specimen valence band into the conduction band (and possibly some from donor surface states, if present, to acceptor surface states) and to cause specimen heating. One or other of these effects may release electrons from surface states or surface locations. There is an arguable case that, almost immediately after the laser pulse and for some time afterwards (during which field evaporation of an ion may occur), the specimen may be acting electrically more like a conductor with the chemical composition of semiconductor than like a semiconductor. At the time of writing, the physics of this nonequilibrium situation is not properly understood (see Sect. 3.9).

### 2.3 Fields, Potentials, and Charged Surface Models

#### 2.3.1 Electric Field Types

#### 2.3.1.1 Introduction

It is difficult to create reliable atomic-level models that reflect the detailed physics of realistically shaped field emitters. Hence, field ion/atom probe theory often uses simple one-dimensional models. It is also difficult to precisely relate the very local electric fields used in atomic-level models to the classical fields used in macroscopic emitter models. Hence, several different types of electric field are used. Some are awkward to define precisely, and often the relationship between the different field types is difficult to specify.

Existing literature often does not distinguish between different field types. For clarity, and with an eye to future developments, this section describes the different field types used in FI/AP theory. Field calibration is discussed in Sect. 2.9.

The basic symbol \( F \) (rather than \( E \)) is normally used for all types of electric field, with distinguishing suffices added where needed. This allows \( E \) to denote total
system energy. In fact, some “fields” in FI/AP theory are more of the nature of $D/\epsilon_0$, where $D$ is electric displacement (flux density). However, it is easier to have all “fields” in V/nm or equivalent units.

One also needs to distinguish between (a) static fields, associated with charge distributions, and (b) the oscillating electric component of an electromagnetic (EM) field, called here an optical field. Physically, these two field types cannot simply be added. This section deals with static fields, which in FI/AP theory are often positive. (A positive field pushes a positive ion in the direction of the field.)

At charged surfaces, the surface atoms are strongly polarized (Sect. 2.3.3.3). Thus, in addition to “$D$-type” fields related to (monopole) charges on individual atoms, there exist short-range fields related to the dipole components of the atomic charge distributions. These are the atomic-level equivalent of fields related to the polarization vector $P$. The sum of the fields gives the total (“$E$-type”) field.

A general convention is that local, as in “local surface field” or “local work function,” means the local value of a parameter that varies with location across the emitter surface.

### 2.3.1.2 Point Fields

Field values defined at points in space are called here point fields. These are particularly relevant when one needs to know the field “acting on an atom.” Usually the need is for the field at the position of the atomic nucleus.

The easier type to treat is the partial point field: this is the field at the position the nucleus would have, but is determined in the absence of the atom (or ion), and in the absence of any charge rearrangements that its presence would induce.

What the theory may actually require is the total point field: this is the field at the nucleus position, determined by disregarding the direct effects of the atom’s charge distribution, but taking into account the field changes due to any charge rearrangements that its presence would induce in nearby atoms (including charge transfer from the atom to its neighbors). Sometimes, but not always, the partial field value can be used to approximate the total field value.

The main applications of point fields are in theories of field adsorption, field evaporation, and surface-atom polarization. In particular, a bonding-point field is the field at the potential energy minimum of a local bonding well.

### 2.3.1.3 Impressed Field

The planar array model (Sect. 2.3.3.2) for a positively charged surface uses a specific $D$-type point field called the impressed field. This is the field, acting at the nucleus of an array atom, that is due to the distant negative charge array. More generally, the impressed field is the field, acting at the nucleus of a surface atom, that is due to the monopole components of the induced charge distributions (see Sect. 2.3.3.1) associated with all other atoms in the system.
2.3.1.4 External Field, Applied Field, and Model Fields

The critical surface (Sect. 2.5.3) for gas field ionization above an emitter surface is nearly outside the range of the short-range fields due to the surface atomic structure. Thus, a point field in the critical surface can be treated as the “typical field” above the related part of the emitter surface and is called here the external field. External fields usually vary as between emitter regions (and will usually be different above an embedded particle), but can be treated as “roughly constant” across a surface region. The term applied field refers to an external-field value that characterizes the emitter as a whole—often the external field at the emitter apex.

When a macroscopic electrostatic model is used to describe an emitter, the related fields are model fields. For practical purposes, one can identify the local model boundary field with the local external field and the boundary field at the model apex with the applied field. The critical surface moves inwards slightly as applied field increases and hence the matching is not exact, but this interpretation is considered adequate for all except extremely small emitters.

2.3.1.5 Average Field and Surface Field

The difference $\Delta \Phi$ in electrostatic potential between two points a distance $\Delta x$ apart is sometimes needed and can be written: $\Delta \Phi = -F_{av}\Delta x$, where $F_{av}$ is the average field over the distance $\Delta x$. It is often impracticable to calculate $\Delta \Phi$ accurately; thus, sometimes, one estimates $\Delta \Phi$ by estimating $F_{av}$ and $\Delta x$ separately.

FI/AP theory uses average fields in three main places. In field ionization theory (Sect. 2.7), one needs the tunneling field, $F_{tb}$, which is a field that characterizes the local electron tunneling barrier. For system potential energies (Sect. 2.5.1), one needs the work done in removing an electron from the emitter Fermi level to some point outside the emitter surface. This involves a special average field called here the surface field. Slightly different forms of surface field are needed in field evaporation theory and in the theory of the critical surface for FI imaging.

2.3.1.6 Notations for Fields

For notational simplicity, the following practice is used. If only one field type is used in a section or subsection, then $F$ will be used and its interpretation specified. If more than one field type is needed, then distinguishing notation will be used. A consequence is that the symbol $F$ can have different meanings in different places.
2.3.2 Classical-Conductor Models for Field Emitters

When applied to field emitters, classical-conductor models define smooth boundaries that represent the emitter and a counter-electrode, and are treated as electrostatic equipotentials. Laplace’s equation is then solved, analytical or numerically, to find how fields vary in the boundary representing the emitter, and/or how fields and potentials vary between the emitter and counter-electrode. In principle, the fields and potentials depend on the emitter’s whole geometrical and electrical environment, but some approximate formulae exist that involve only its apex radius \( r_a \).

In FI/AP theory, the main uses of classical-conductor models are as follows: to find the relationship between the applied field (i.e., apex field) and the applied voltage; to evaluate field stresses acting on specimens; and in ion optics. Most commonly, these models have been used to represent the overall emitter shape; however, in the last 10 years, classical models have been developed that represent the emitter shape at an atom-by-atom level (see Sect. 2.3.2.4).

2.3.2.1 Electrostatic Potential Difference and Applied Voltage

Consider a metallic emitter and a metallic counter-electrode that have local work functions \( \phi^e \) and \( \phi^c \) respectively, and let the applied voltage (i.e., the positive voltage applied to the emitter, relative to the counter-electrode) be \( V \). Because measured voltages relate to differences in Fermi level (inside metals), but electrostatic potential differences relate to differences between points in space just outside metal surfaces, the electrostatic potential difference \( \Delta \Phi \) between the emitter and the collector is given by

\[
\Delta \Phi \equiv \Phi^e - \Phi^c = V - (\phi^e - \phi^c).
\]  

(2.5)

The term \( (\phi^e - \phi^c) \) is nearly always less than 1 V. Hence, in FI/AP theory the distinction between \( V \) and \( \Delta \Phi \) is nearly always neglected: voltage \( V \) is inserted into formulae that (from the point of view of electrostatics) should contain \( \Delta \Phi \).

2.3.2.2 Simple Relationships Between Applied Field and Voltage

For a single-tip field emitter, the relationship between emitter apex field \( F_a \) and the applied voltage \( V \) can be written formally in either of the following ways:

\[
F_a = \beta_a V = V / \zeta_a,
\]  

(2.6)

where \( \beta_a \) is the apex value of the field-to-voltage conversion factor \( \beta \), and \( \zeta_a \) is the apex value of the related local conversion length \( \zeta \) \( \equiv 1/\beta \). Using \( \beta_a \) is the historical approach, but using \( \zeta_a \) now seems less confusing and usually more convenient.
\( \zeta_a \) is a conversion parameter, not a real physical length, and for field electron emitters typically has a value in the range 100–1000 nm. Other aspects of system geometry being equal, sharper emitters have smaller \( \zeta_a \)-values.

A parameter \( k_f \), called a shape factor or a field factor, is defined by \( \zeta_a = k_f r_a \), where \( r_a \) is the emitter apex radius; this leads to the well-known formula:

\[
F_a = \frac{V}{k_f r_a}.
\]  

(2.7)

For a sphere \( k_f = 1 \). For real emitter shapes \( k_f > 1 \): this is due to the effects of (primarily) the emitter shank, but also the overall system electrostatics.

Traditionally, metallic field emitters were prepared with shapes such that (in classical microscope configurations) \( k_f \) lay in the range 3–8, with a value of 5 sometimes being considered typical. This means, for example, that a field electron emitter with \( k_f = 5 \) and \( r_a = 100 \) nm would reach a typical electron emitting field of 4 V/nm at an applied voltage of 2000 V. A field ion emitter with \( k_f = 5 \) and \( r_a = 40 \) nm would reach the tungsten field evaporation field of 57 V/nm at an applied voltage of ~11000 V.

### 2.3.2.3 Macroscopic Conductor Models

Macroscopic classical-conductor models have been extensively used to investigate the charged-particle optics of field emitters [31–35]. Such models can be solved analytically if the system can be approximated as axially symmetric and the emitter shape can “to an adequate approximation” be modeled as an equipotential in a particular system of coordinates. Further, if the counter-electrode can be approximated as having “nearly the correct mathematical shape,” then usually only a single pair of terms is needed and the solution is relatively straightforward. In other cases, the models are often best analyzed via numerical solution of Laplace’s equation (although multi-term series solutions sometimes exist). The designers of field electron emission guns for high-resolution electron microscopes have much relevant numerical expertise [33].

The issue of precisely how models should be matched to a real atomically structured emitter is difficult and poorly investigated and is not discussed here. For an idealized atomically smooth real emitter, a simple approach takes the total electron PE to be the same everywhere in the boundary and equal to an average value of local work function, disregards all “patch field” issues noted in Sect. 2.2.2.2, takes the model boundary to coincide with the electrical surface (or a smoothed version of it), and interprets a local boundary field as the corresponding local external field. This should usually be adequate, except for very small emitters and for calculating fields and potentials very close to the emitter surface.

FI/AP theory has explored the use of parabolic and hyperbolic models and coordinate systems (see [34] for detailed discussion). However, there are problems with these simple models. It is often difficult to fit them to real emitter shapes, and
they may not accurately represent the electrostatics of the whole system. A better macroscopic shape model is the sphere-on-orthogonal cone (SOC) model [36], introduced into field emission in 1953 by Dyke’s group [37] and subsequently much studied by FE gun designers [33].

2.3.2.4 The Sphere-on-Orthogonal Cone Model

The SOC model, illustrated in Fig. 2.5, is axially symmetric. The model uses polar coordinates \((r, \theta)\) with their origin at the center of a core sphere of radius \(r_c\) and initially takes this sphere as the reference zero. \(\theta\) is measured from the axis, with an on-axis emitted particle having \(\theta = 0\).

If a counter-electrode “D” is at large distance \(L_D\) (measured along the symmetry axis) and has the correct shape (or approximates adequately to it), the classical electrostatic potential at point \((r, \theta)\) is adequately given by the single pair of terms:

\[
\Phi = \left(\frac{\Phi_D}{L_D^n}\right) \left[r^n - \left(r_c^{2n+1} r^{-n-1}\right)\right] P_n(\cos \theta), \tag{2.8}
\]

where \(\Phi_D\) is the electrostatic potential of a point just outside D. \(P_n(\cos \theta)\) is the Legendre polynomial of noninteger order \(n\), and \(n\) is chosen so that \(P_n(\cos \theta)\) has its first zero when \(\theta = \pi - \Theta_c\), where \(\Theta_c\) is the half-angle of the mathematical supporting cone. (\(\Theta_c\) is not necessarily equal to the half-angle determined by tip profiling in the electron microscope.) For field ion emission, \(\Phi\) and \(\Phi_D\) are negative.

For sufficiently large \(L_D\), the equation for the equipotential that the counter-electrode needs to follow is

\[
r = L_D \cdot \left[P_n(\cos \theta)\right]^{-1/n}. \tag{2.9}
\]

Dyke et al. describe this shape as “approximately parabolic.” Note that it depends on the parameter \(n\) (and hence on the emitter cone half-angle).

By suitable choice of \(\Theta_c\) (and hence of \(n\)) and of \(r_c\), and by setting \(r\) equal to the desired apex radius value \(r_a\), an equipotential of value \(\Phi_a\) [evaluated from Eqn. 2.8]
can be fitted to the tip electron microscope profile [37]. On the symmetry axis, 
\[ P_n(\cos \theta) = 1, \]
and the apex field \( F_a \) \( \equiv F(r_a,0) \) is given by

\[
F_a = -\left( \frac{\partial \Phi}{\partial r} \right)_\theta \Big|_{r=r_a,\theta=0} = \left( \frac{\Phi_D}{L_D} \right) \left[ n r_a^{n-1} - (n + 1) r_c^{2n+1} r_a^{-n-2} \right].
\]

In some contexts, one needs the relationship between the field \( F_b \), at some point “b” on the emitter model boundary, and the magnitude \( \Delta \) of the electrostatic potential difference between the counter-electrode and the emitter (as represented by the chosen equipotential). This difference is given by

\[
\Delta = - (\Phi_D - \Phi_a).
\]

Hence, if, for point “b,” a Dyke-type conversion factor \( \beta_b \) is defined by

\[
F_b = \beta_b \Delta,
\]
then its on-axis value \( \beta_a \) is found by simple algebra to be

\[
\beta_a = r_a^{-1} \left[ \left( \frac{r_a}{L_D} \right)^n \left\{ n + (n + 1) \left( \frac{r_c}{r_a} \right)^{2n+1} \right\} \left( \frac{\Phi_D}{(\Phi_D - \Phi_a)} \right) \right].
\]

The last bracketed term is close to unity, and the second term in the second bracket is significantly larger than the first. The shape factor \( k_f = 1/\beta_a r_a \). Thus, to an adequate approximation, the SOC model predicts \( k_f \) as

\[
k_f \approx \left( \frac{L_D}{r_a} \right)^n \left( \frac{r_a}{r_c} \right)^{2n+1} / (n + 1).
\]

The dependence of \( k_f \) on tip shape (cone angle and apex radius), and on the electrostatic environment (via \( L_D \)), comes out clearly in this model.

There also exists a complicated analytical formula for how \( \beta_b \), and hence the boundary field, fall off with polar angle (see [37]).

It is sometimes suggested that the parabolic and hyperbolic models are simpler than the SOC model. Certainly, the SOC model formulae look slightly more complicated. However, this is deceptive. Because the SOC model uses ordinary (planar) polar coordinates, it is easier to understand and manipulate than the other models. The only difficult feature is evaluating the nonintegral Legendre polynomials off-axis. The relationship between \( n \) and \( \Theta_c \) is given in Table 2.1; this is all that is needed for on-axis evaluation of \( \Phi \).

Because estimates of \( n \) and \( r_a \) can be obtained from simple tip profiling in the electron microscope, and \( k_f \) from either the best image voltage or experiments on the onset of field evaporation, an approximate estimate of \( r_c \) can (in principle) be found without the detailed (but more reliable) profile fitting used in [37].

For APT theory, the SOC model has two weaknesses, but also two potential advantages over the sphere-on-tangential cone (STC) model often used in APT reconstruction. One weakness (shared with the STC model) is that the SOC model
apex is nearly spherical, whereas real field evaporated end forms are slightly flattened and may change, particularly under laser-pulsed evaporation. The other is that the counter-electrode in a LEAP does not conform well to Eqn. 2.5; hence the model will become poor at large $r$-values (see [33, p. 492] for discussion of a similar problem in field electron emission). However, this may not be important for APT, if in this region the ions are effectively traveling in straight lines. Advantages of the SOC model are that (1) it may offer a better (though still imperfect) representation of actual tip shape and (2) it has analytical solutions. However, it is too early to know how useful the SOC model might be.

### 2.3.2.5 The Electrostatic Effect of Small Isolated Surface Protrusions

The electrostatic effects of small protrusions on top of field emitters are sometimes of interest. The related theory is derived for protrusions on flat surfaces, but is considered applicable to protrusions on curved surfaces, provided that the height and radius of the protrusion are less than one-tenth of the substrate radius of curvature. The theory gives the field enhancement factor (FEF) $\gamma_L$ defined by

$$\gamma_L = \frac{F_L}{F_M}, \quad (2.15)$$

where $F_L$ is the local field at some surface position on the protrusion, and $F_M$ is the macroscopic field that would be present in the absence of the protrusion. Usually, interest is in the FEF $\gamma_a$ at the protrusion apex.

Two well-known results are the following. For a hemisphere, $\gamma_a = 3$. For a hemispherical cap on a cylindrical post of height $h_p$ and radius $r_p$, for $h_p >> r_p$:

$$\gamma_a \approx \frac{h_p}{r_p}. \quad (2.16)$$

For better versions of Eqn. 2.16, and formulae for some other shapes, see [38, 39].
2.3.2.6 Models Based on Atomic-Level Geometry

It is well established [34, 35] that the precise point of impact of a field-evaporated ion onto an AP detector depends on the details of its trajectory close to the emitter surface. A problem with macroscopic emitter models is that they cannot generate this kind of detail. For an emitter assumed to have a simple cubic lattice, Vurpillot and colleagues [40] created an emitter model in which each atom was represented by a conducting cube, and the set of cubes closest to a (mathematical) hemispherical enclosing surface was regarded as defining the emitter surface. A concentric hemisphere was used as the counter-electrode. Laplace’s equation could be solved for this and similar configurations, and ion trajectories calculated. By a progressive, iterated algorithm [40], which identified which atoms were subject to the highest surface fields, and then treated these as field evaporated atoms, the shape of the emitter model was modified until its endform stabilized. Statistics relating to ion trajectories and impact points onto the counter-electrode could then be collected and used to give qualitative explanations of some features of field evaporation images (see Sect. 1.1.7). For further details and references, see [40].

This method has recently been extended by Oberdorfer and Schmitz [41] to investigate the field evaporation behavior of dielectric materials.

2.3.3 Classical Array-Type Charged Surface Models

In FI/AP theory, the main uses of array-type charged surface models relate to (a) electrical surface location, which in turn is relevant to field ionization theory and field evaporation theory, and (b) field adsorption theory. These models employ arrays of point charges and dipoles. They provide basic understanding and—for electrical surface location—surprisingly good quantitative results. In principle, quantum-mechanical (QM) models (Sect. 2.3.4) should provide more accurate results, but they are less transparent, particularly for experimentalists.

2.3.3.1 Real Charge, Excess Charge, and Induced Charge

Atomic-level charge distributions can be represented in array-type models by making distinctions between real charge (protons and electrons), excess charge, and induced charge [42, 43]. Consider a real charged body “R,” and also the formal system “F” in which each atom in R is replaced by a neutral atom with the charge distribution it would have if isolated in field-free space. The excess charge distribution of body R is the difference between the charge distributions of body R and system F (the difference “R” — “F”).
Now consider the real body in two states: a state “N” in which it is neutral overall and a state “C” in which it is charged overall. The induced charge distribution is the difference between the excess charge distributions in states “C” and “N” (the difference “C” − “N”). The induced charge distribution is also given by the difference between the real charge distributions in states “C” and “N.”

In array-type models, the induced charge distribution is modeled by placing point charges and polarizable dipoles at the positions of the atomic nuclei, and then applying classical electrostatics. The first use of point dipoles to model charged surface effects is sometimes attributed to Drechsler [44] and Becker [45]. In the context of FI/AP theory, they have been used by Tsong and Müller [46] to discuss field adsorption (Sect. 2.6) and by Forbes to discuss both field adsorption (with Wafi) [47] and more general aspects of charged surface physics (see below).

2.3.3.2 Classical Charge Arrays and Maxwell Stress

At a charged metal surface, the induced “monopole” charges are located on the surface atoms. (In reality, there may be a small amount of charge oscillation inside the surface, but this is disregarded.) The simplest useful model is the infinite classical planar array model illustrated in Fig. 2.6. This assumes a parallel-plate-capacitor situation. A superimposed point charge $q$ and a polarizable point dipole are placed at the position of each surface atom nucleus. If $A_{SL}$ is the area in the surface lattice allocated to each atom, then the surface charge density $\sigma = q/A_{SL}$, and the external field $F^{ext}$ between the plates, but well away from either, is

![Fig. 2.6 Schematic diagram of the classical planar array model for the positively charged, atomically flat surface of a good conductor. The model consists of superimposed charges and dipoles at the positions of the surface atom nuclei, together with a balancing array of distant negative charge, needed for electrostatic self-consistency.](image-url)
The field acting on the positive model charges, due to charges on the distant negative plate, is the impressed field $F^{\text{imp}} = F^{\text{ext}}/2$. (In the terminology of Sect. 2.3.3.1, $F^{\text{imp}}$ is a $D$-type field.) Hence, on the positive model charges there is an outwards force per unit area, or stress, $S_M$ given by

$$S_M = \sigma F^{\text{imp}} = \frac{1}{2}\varepsilon_0\left(\frac{F^{\text{ext}}}{2}\right)^2. \quad (2.18)$$

This is the well-known Maxwell stress.

### 2.3.3.3 Surface Atom Polarization

Now consider how the real atomic charge distributions are influenced by the distant negative charges. For each surface atom, if the electrical center of its electrons is at its nucleus, then there will be no overall resultant force on either due to the other and (for a ordered surface of infinite extent) no resultant force on either due to other atoms in the surface.

The impressed field due to the distant negative charge tries to move the surface atom electrons towards the positive plate, but this is resisted by exchange-type repulsive forces. The impressed field pulls the nuclei away from the plate. This displaces them from the electrical centers of the surrounding electrons and exposes them to electrical restraining forces. Equilibrium occurs when the two forces become equal in magnitude. The outcome is a polarized surface atom layer. The correctness of this classical argument was confirmed, later, by self-consistent quantum-mechanical calculations (see below). This polarized layer is a universal property of charged metals of any shape, but its existence is not well known.

Given the crystallographic arrangement of the surface atoms, and a value for the effective atomic polarizability $b$, the array electrostatics can be solved classically [47, 48] to give the electric dipole moment $p_{\text{el}}$ at each array site as

$$p_{\text{el}} = bF/2M_E, \quad (2.19)$$

where (here and in Sect. 2.3.3.4) $F$ is used to denote external field, and $M_E$ is a depolarization factor that recognizes the depolarizing effect that the dipoles have on each other. $M_E$ acts as a relative permittivity for the layer; values depend on surface crystallographic structure and are typically ~1.5 to ~2.

The dipole layer has moment per unit area $\Pi = p_{\text{el}}/A_{\text{SL}}$. This creates an electrostatic potential difference $\Delta\Phi_d$ across the layer (with the higher potential on the vacuum side, for a positively charged emitter) given by

$$\Delta\Phi_d = \Pi/\varepsilon_0 = p_{\text{el}}/\varepsilon_0A_{\text{SL}} = (b/2\varepsilon_0M_EA_{\text{SL}})F. \quad (2.20)$$

At metal evaporation fields, predicted values of $\Delta\Phi_d$ can be as much as 5–10 eV.
2.3.3.4 Repulsion Distance and the Location of Electrical Surface

It is important to know the location of the electrical surface, relative to the positions of the surface atom nuclei, since its location affects estimates of the energy level (relative to the Fermi level) of the electron orbitals in an external atom.

If there were no field-induced surface dipoles, then the electrical surface would be in the plane of the surface nuclei \[\Delta \Phi_d\]. When dipoles are present, the electrical surface must be shifted outwards (towards the vacuum) by the repulsion distance

\[d = \Delta \Phi_d / F = b / 2\varepsilon_0 M_E A_{SL}.\]  \hspace{1cm} (2.21)

This effect is called field repulsion. As shown in Fig. 2.7, repulsion is necessary if we are to continue to write the field-induced component of the electrostatic PE of a point positive charge in the asymptotic form \(-eFx\), and hence that of a negative point charge in the form \(+eFx\), where \(x\) is measured from the electrical surface.

As shown in Table 2.2, the predicted \(d\)-values \[42\] are comparable with the atomic radius, as given by half the nearest-neighbor distance \(r_{NN}\) in the metal space lattice. If polarizability \(b\) is independent of field, then \(d\) and the electrical surface position are also independent of field. This assumption is usually made, although in reality one expects \(b\) to decrease weakly but steadily with field \[50\].

The essential correctness of this scientific thinking has been shown by appearance energy experiments, e.g., \[51\], that clearly locate the electrical surface on the vacuum side of the surface atom nuclei. The repulsion distances deduced experimentally are of the order of an atomic radius or somewhat less, but it has been difficult to extract reliable numerical values from these experiments (see \[50\] for a detailed discussion).

In summary, for the planar array model, the electrostatic component \(\nu_{ES}\) of electron PE (measured relative to the Fermi level) is given asymptotically by
Eqn. 2.2 \[ v_{ES} = \phi + eF_x \], where the electrical distance \( x \) is measured from the electrical surface. This surface is on the vacuum side of the surface nuclei by the distance \( d \) (the repulsion distance). This distance \( d \) depends on the crystallographic structure of the emitter face, but is roughly equal to the half the nearest-neighbor distance in the space lattice. The distance parameter used in simple one-dimensional FI theory can be identified as the electrical distance.

### 2.3.4 Quantum-Mechanical Charged Surface Models

Obviously, QM charged-surface models are in principle better than classical models, but they can be very difficult to implement reliably. At present, analyzed geometries are restricted to flat planar surfaces (with or without one or more adsorbed atoms) and to relatively small atomic clusters. Kiejna and Wojciechowski [21] provide an overview of some of the related physics.

Planar-surface models fall into two broad classes: (a) those in which the ion cores are “smeared out” into a uniform positive charge distribution (with a sharp edge) called jellium (e.g., [22]) and (b) atomistic models in which the existence of ion cores and the associated localized electron-density variations are taken into account. Such models can be categorized as using self-consistent treatments (e.g., [52–55]), which attempt to solve the Schrodinger equation accurately, or as based on density functional theory (DFT), e.g., [56]. There is also a hybrid approach, namely, the “clusters embedded in jellium” model [57].

Most early QM treatments of charged surfaces used either jellium models (e.g., [22, 58]) or cluster models, e.g., [59, 60]. However, atomistic planar surface models, built using repeated cell techniques, ought to give more reliable results.

The good self-consistent treatments of Aers and Inglesfield of this type [52–54] confirm the polarization of surface atoms. Also, for Al, the repulsion distances (~150 pm) predicted by Inglesfield [52] and Lam and Needs [55] coincide with the Forbes classical estimates to within 20 pm or better [43].

<table>
<thead>
<tr>
<th>Material and face</th>
<th>( r_{NN}/2 ) (pm)</th>
<th>( d ) (pm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>W(110) [bcc]</td>
<td>137</td>
<td>157</td>
</tr>
<tr>
<td>Ir(111) [fcc]</td>
<td>136</td>
<td>144</td>
</tr>
<tr>
<td>Mo(110) [bcc]</td>
<td>136</td>
<td>160</td>
</tr>
<tr>
<td>Pt(111) [fcc]</td>
<td>139</td>
<td>141</td>
</tr>
<tr>
<td>Au(111) [fcc]</td>
<td>144</td>
<td>138</td>
</tr>
<tr>
<td>Rh(111) [fcc]</td>
<td>135</td>
<td>147</td>
</tr>
<tr>
<td>Feα(110) [bcc]</td>
<td>129</td>
<td>143</td>
</tr>
<tr>
<td>Ni(111) [fcc]</td>
<td>125</td>
<td>136</td>
</tr>
<tr>
<td>Cu(111) [fcc]</td>
<td>128</td>
<td>135</td>
</tr>
</tbody>
</table>
Field repulsion also occurs in Lang and Kohn’s seminal treatment of charged surfaces based on the jellium model [22], in the sense that the electrical surface is predicted to be on the vacuum side of the jellium surface. As noted earlier, they show that, in their model, the image plane and the electrical surface coincide. For all models, both classical and QM, the electrical surface can be identified as the centroid of the induced charge distribution [22, 42] on the relevant emitter surface.

The addition of a surface atom to a planar surface model enables investigation of (a) the field and potentials above and around the atom [61] and (b) the field evaporation and/or lateral motion [56] of a surface atom moving as a partial-ion (i.e., in conditions where part of its electron charge has moved to the substrate).

A partial-ion has a nonintegral number of electrons in its associated volume of space (or as assessed by its Mulliken charge, e.g., [62]). The existence of partial ions at charged metal surfaces follows from Gauss’ theorem. Physically, the original topmost atomic orbital still participates in the conduction band.

These “adatom” investigations employ charged surfaces density functional theory (CS-DFT). CS-DFT has also been used to investigate reconstruction of highly charged surfaces (Sect. 2.4.4.3) and is now being used to investigate field evaporation details (Sect. 3.4.4).

2.3.5 The Issue of “Field Penetration”

In some literature, particularly older literature, the statement is encountered that “fields penetrate into metal surfaces.” This effect is found in older jellium models, e.g., [63], in which Thomas–Fermi theory (rather than the approach of Lang and Kohn) is applied to the electrons. In FI/AP theory, there has been past debate (sometimes confused by failure to define the meanings of the words “field,” “penetrate,” and “surface”) as to whether field penetration occurs.

The best approach, it now seems, is to treat the question “Does a field penetrate into a metal surface?” as not scientifically helpful. Rather, it is better to ask two questions: “Where is the electrical surface?” and “How far does a significant induced charge distribution extend into the metal?”

As indicated above, the electrical surface is outside the surface nuclei, and (in newer jellium models) is normally outside the jellium surface. (The electrical surface can, however, be inside the edges of the surface-atom electron distribution.) It seems more logical to call this effect “field repulsion.” As to the induced charge distribution, most of this resides on the surface atoms, but current thinking is that there are also small and probably oscillating contributions arising from the next few layers in. These are usually neglected.

For semiconductors, of course, the situation is different. Real field penetration may occur, and the electrical surface is then inside the semiconductor surface [30].
2.4 Thermodynamics of Charged Surfaces

2.4.1 Surface Stress

The concept of surface stress is conveniently introduced by considering liquid metal behavior. To a good approximation, the difference $\Delta p \equiv p^{\text{int}} - p^{\text{ext}}$ between the local pressures inside ($p^{\text{int}}$) and outside ($p^{\text{ext}}$) the curved surface of a charged conducting liquid, located in vacuum, is given by [64]

$$\Delta p = \gamma^0 \left( \frac{1}{r_1} + \frac{1}{r_2} \right) - \frac{1}{2} \varepsilon_0 F^2. \quad (2.22)$$

Here, the first term is the classical surface tension term, where $\gamma^0$ is the surface free energy per unit area in the absence of any applied field, and $r_1$ and $r_2$ are the principal radii of curvature of the surface, at the position of interest. This term has its physical origin in the difference between the bonding environments of a surface atom and an atom in the interior and can be understood as a chemical stress term. (For the surface atom, there is no atom “outside it” that can pull it outwards, by an attractive bond, so the interior atoms “pull it inwards” and create a pressure/stress.) This applies to both liquids and solids. The second term is the Maxwell field stress term encountered in Sect. 2.3.3. Due to variations in $\Delta p$ across the surface, solids also experience shear stresses parallel to the specimen surface (see [34], p.118).

In summary, the chemical stress term acts to increase pressure within the body and the field stress to decrease it. In principle, a full expression for $\Delta p$ should include further small terms [65], resulting from tip curvature and surface atom polarization, but in non-atomic-level contexts these terms can normally be disregarded.

When $F$ is sufficiently high, as it is near the emitter apex in FIM and APT, then $\Delta p$ is negative. The physical stress is outwards (i.e., tensile in nature) and nonuniform and can be very large in magnitude. For example, at the predicted evaporation field for Fe (33 V/nm), the field stress is $-4.8$ GPa ($-48$ kbar). For Fe, $\gamma^0 \approx 2$ J/m$^2$, so for a specimen of 40 nm apex radius, the chemical stress is approximately +0.1 GPa. Clearly, the field stress term dominates.

For bulk materials, tensile stresses of this size are greater than measured tensile strengths (usually between 0.1 and 1 GPa for Fe-based materials). Field-induced stresses can cause defects, such as dislocations, to move, and this can result in specimen failure. Thus, the size and distribution of such stresses are of interest. Ringer and colleagues [66] have recently taken a new look at issues of this kind, using several different emitter electrostatic models, including the SOC model. Their article contains a useful list of references to previous investigations.

The idea that the two terms in Eqn. 2.22 can simply be added seems to originate with Maxwell in 1873 ([19], see p.190). A better derivation of Eqn. 2.22 uses a variational thermodynamic argument, as indicated below.
2.4.2 The Electrical Gibbs Function

All field emission systems have capacitance between the emitter and its surroundings. When a system contains internal capacitance, and electrical work can be done on the system by an external voltage generator, then an electrical form of thermodynamics is needed to describe changes in the system.

When appropriate thermodynamic parameters, including temperature and the applied voltage, are held constant, system behavior is determined by the change $\Delta G_{el}$ in a special thermodynamic potential $G_{el}$ introduced by Ljepojevic and Forbes [64] (their $\Psi$) and best called an electrical Gibbs function.

When a change in emitter shape causes a change $\Delta C$ in capacitance between the emitter and its surroundings, $\Delta G_{el}$ is given by

$$\Delta G_{el} = \Delta \Theta_{tot} - w_{el} = \Delta \Theta_{tot} - V \Delta q = \Delta \Theta_{tot} - V^2 \Delta C,$$  \hspace{1cm} (2.23)

where $\Delta \Theta_{tot}$ is the change in the system’s total Helmholtz free energy, $w_{el}$ is the external electrical work done on the system, and $\Delta q$ is the charge that moves round the circuit, through the voltage generator. $\Delta \Theta_{tot}$ needs to include appropriate electrical terms, including a term $(1/2)V^2 \Delta C$ giving the change $\Delta \Theta_{cap}$ in the internal capacitative energy.

When applied to the situation of a field ion emitter, taking only the most basic terms, Eqn. 2.23 yields

$$\Delta G_{el} = \Delta \Theta_{bulk} + \Delta \Theta_{surf} + \Delta \Theta_{cap} - w_{el} = \Delta \Theta_{bulk} + \gamma \cdot \Delta A - \frac{1}{2} V^2 \Delta C,$$  \hspace{1cm} (2.24)

where $\Delta \Theta_{surf}$ and $\Delta \Theta_{cap}$ are the changes in Helmholtz free energy associated with the emitter surface and the capacitance, respectively, $\Delta A$ is the increase in surface area, and $\gamma$ is the free energy per unit surface area (i.e., the “surface tension”). As is well known, $\Delta \Theta_{cap}$ can also be interpreted as the change in the “field energy” in the space between the emitter and its surroundings, with the energy density at any point in free space (where the electric field is $F$) equal to $1/2 \varepsilon_0 F^2$. It follows that, at least to a first approximation, we can also write (using $dv$ as the volume element):

$$(\Delta \Theta_{cap} - w_{el}) = -\Delta \left[ \frac{1}{2} \varepsilon_0 F^2 dv \right].$$  \hspace{1cm} (2.25)

Equations 2.24 and 2.25 are basic equations of charged surface thermodynamics and yield several results of interest to FI/AP theory. In particular, by developing an expression for $\Delta \Theta_{bulk}$ in terms of the pressure difference $\Delta p$ across a curved, charged conducting surface, and considering a formal small change in emitter shape, Ljepojevic and Forbes [64] were able provide a variational thermodynamic proof of Eqn. 2.22. Other applications of these ideas are now outlined.
2.4.3 Field Dependence of Metal Atom Bonding Energy

The bonding energy of an atom to the emitter surface is the work needed (on the system comprised by the emitter and its electrical surroundings) in order to remove the atom, as a neutral, to remote field-free space. In principle, this bonding energy is a system parameter, in the sense that it is defined by the difference in system configuration and system energy before and after atom removal.

In the absence of any applied field, bonding energy is denoted by $\Lambda^0$. In field evaporation theory for solid chemical elements, $\Lambda^0$ is taken as equal to the tabulated sublimation energy $\Delta H_{\text{sub}}$. This may not be exactly true for a small sharply curved emitter, but the overall effects of emitter curvature are usually disregarded. However, when a surface atom is in a special position, such as isolated on top of a surface facet, then $\Lambda^0$ will be different from $\Delta H_{\text{sub}}$. The physical effects that cause $\Lambda^0$ can be described as zero-field chemical effects.

In the presence of a local field, the bonding energy of a surface atom increases by $\Delta \Lambda$ to the value $\Lambda^F$. The simplest (and usual) approximation writes

$$\Lambda^F = \Lambda^0 + \Delta \Lambda \approx \Lambda^0 + \frac{1}{2} c_0 F^2,$$

where $F$ is the external field above the atom position and $c_0$ is a coefficient that might in principle vary with the atom’s environment but in a first approximation can be treated as independent of field.

It was long thought that $c_0$ was determined by the polarizability of the relevant surface atom or molecule. But it now seems clear, from the general thermodynamics arguments above, that the main electrical effect must normally be the change in system capacitance caused by the removal of the atom.

Following Forbes [65], consider a charged, electrically isolated, parallel-plate capacitor. Let one plate “E” be atomically flat and represent the emitter, and suppose that the internal atomic layers are parallel to the emitter surface. Remove a layer of atoms from plate E by the following formal steps. (1) Pull plate E away from the opposing plate by a distance equal to an interlayer spacing. This increases the stored energy in the capacitor by an average amount $w_1$ per atom given by

$$w_1 = \frac{1}{2} \varepsilon_0 \Omega F^2,$$

where $\Omega$ is the atomic volume. (2) Create an electrically shielded enclosure inside E by breaking the bonds between two atomic layers and pulling the two sections of E slightly apart; remove a layer of atoms, one by one, from one of the planar surfaces of the enclosure; and eliminate the enclosure by moving the two parts of E together again. The average work $w_2$ done per atom in this process is $w_2 = \Lambda^0$.

After these steps, the surfaces of E and its counter-electrode are in exactly the same state of ionization and polarization as they were before, and (because the changes were made under constant charge conditions) the field $F$ in the capacitor has exactly the same value as before. Therefore (on average), the total work $\Lambda^F$ per atom done in the process is simply $w_1 + w_2$, and
\[ c_0 = \varepsilon_0 \Omega, \quad \Delta \Lambda = \frac{1}{2} \varepsilon_0 \Omega F^2. \]  

(2.28)

\( \Delta \Lambda \) can also be interpreted as the work needed to fill (with electrostatic field energy) the “hole in space,” of volume \( \Omega \), left by removing an atom.

A comparison of this theoretical result with experiment is possible for rhodium. For Rh, \( \Omega \approx 0.0138 \text{ nm}^3 \), thus theory predicts \( c_0 \approx 0.763 \text{ meV V}^{-2} \text{ nm}^2 \). From the experiments of Ernst [67], Forbes and Chibane [68] deduced the empirical result \( c_0 = 1.05 \pm 0.3 \text{ meV V}^{-2} \text{ nm}^2 \). Thus, agreement is satisfactory. For a field of 41 V/nm (the Rh evaporation field derived from Müller’s formula, Eqn. 3.23) \( \Delta \Lambda \approx 0.64 \text{ eV} \). This value is significant in comparison with the assumed zero-field bonding energy for Rh (\( \Lambda^0 = 5.75 \text{ eV} \)) and will affect evaporation field predictions, Sect. 3.5.1. Similar correction terms need to be included for all materials.

2.4.3.1 Commentary

Older FI/AP literature takes the primary physical origin of the coefficient \( c_0 \) to be the partial ionization and associated polarization (PIP) of the surface atom. Certainly there is energy associated with these effects, but this energy forms part of the energy associated with charging the emitter’s capacitance in the first place.

There may be some changes in PIP effects when the atomic arrangement after field evaporation is very different from what it was beforehand (e.g., when the last few atoms on a plane are removed). However, for most evaporated atoms the main contribution to \( c_0 \) is now thought to come from field-energy effects.

For field evaporation from kink sites on moderate to large crystal facets, the field configuration close to the surface will be similar before and after FEV, but is “shifted along by one atom.” This implies that the energy density (energy per unit volume) that appears in \( w_1 \) is normally best taken as the energy density in free space slightly above the surface, namely as \( \frac{1}{2} \varepsilon_0 F^2 \), where \( F \) is the external field.

As noted above, the last few atoms in an evaporating layer are expected to have different evaporation characteristics. This is sometimes observed. Discussion of related \( \Lambda^F \) values is very difficult, because detailed comparison of the “before” and “after” situations is needed, for both chemical bonding effects (both zero-field and PIP effects) and field-energy effects. The physics is very complicated and has never been fully explored.

2.4.4 Thermodynamically Driven Effects

Thermodynamically driven charged surface phenomena affect field emitters in several ways. At the macroscopic level they provide the driving force for emitter shape changes—a process known as thermal-field (TF) shaping. At an atomic level, surface reconstruction may occur. The effects noted in Sect. 2.4.3 also determine
the potential energy structure in which the TF surface migration of “atoms” (really partial ions) occurs. The physics of TF shaping is discussed first, because it is more transparent and is useful background to the atomic-level effects.

2.4.4.1 Thermal-Field Shaping

In 1600, Gilbert found that applying a high electric field to a conducting liquid causes it to rise into a cone-like shape [69], and in 1731 Gray found that this cone develops a jet at its apex [70]. These effects are exhibited by water-based jets [71–73], by the liquid metal ion source [1, 74], and by solid metal field emitters hot enough for the surface atoms to be mobile, e.g., [6, 75, 76]. On the other hand, at low fields, liquid bodies tend to become spherical and solid bodies to “ball up.”

These TF-shaping effects are driven by Eqn. 2.24, using the principle that a system not in equilibrium tends to change in a way that decreases its Gibbs function. In TF shaping, the term $\Delta \tilde{F}_{\text{bulk}}$ is disregarded, on the grounds that surface atom motion will not significantly change internal elastic strain energy. Predictions then depend on the applied voltage $V$.

If $V$ is sufficiently small, then the surface energy term in Eqn. 2.24 dominates, and $G^{\text{el}}$ is made more negative by reducing emitter area. In this case, the emitter tends to “ball up” and become blunted. Faceting can occur when variations in surface energy exist as between different crystallographic faces [77].

Conversely, if $V$ is sufficiently large, then the capacitance term in Eqn. 2.24 dominates, and $G^{\text{el}}$ is made more negative by increasing the capacitance between the body and its surroundings. In this case, the body changes its shape to “reach out” towards its surroundings. Individual atoms tend to move towards locations of higher electric field, and body shape changes, by overall sharpening and/or by growth of one or more protrusions or nanoprotrusions.

Thermodynamic arguments give the direction of thermodynamic drive. Detailed atomic-level kinetics are, of course, determined by the activation energy barriers to motion experienced by individual atoms (or groups of atoms), by the effects of local electric fields on these barriers, and by local temperatures.

The condition for blunting to change to sharpening is of interest. It is often assumed, e.g., [6], that sharpening is dominant at local field strengths $F$ such that

$$\frac{1}{2} \varepsilon_0 F^2 > \gamma^0 (1/r_1 + 1/r_2).$$

(2.29)

This “stress-focused” formula is considered not to be of general applicability, because it has not been mathematically derived as a general stability criterion (see [1], Sect. 2.4.2), but it seems to work in practice [6].

For illustration, consider an Fe emitter of tip radius 60 nm. For Fe, $\gamma^0 \sim 2 \text{ J/m}^2$; Eqn. 2.29 then gives the “change-over” field as $\sim 3 \text{ V/nm}$. Field evaporation takes place at much higher fields ($\sim 30 \text{ V/nm}$ for Fe). More generally, field evaporation fields are much greater than “change-over” fields. Thus, at APT operating fields, the
thermodynamic tendency is for atomic migration to take place statistically in such a way that TF sharpening occurs—if the emitter temperature is high enough to allow migration. Obviously, during laser pulsing, an important basic requirement is to keep temperatures low enough to avoid TF surface migration.

2.4.4.2 Field-Induced Surface Relaxation and Reconstruction

A general effect of the Maxwell stress on a field emitter is to cause slight relaxation of the crystal lattice in the region of the emitter affected, normally considered to be around 1–2% [34, 78]. In addition, surface atoms (particularly protruding atoms) become partially ionized, and with increase in the applied field their bonding points move outwards relative to the underlying atomic layers. Sánchez et al. [56], for a particular model of the field evaporation of an Al atom, found an outwards movement of its bonding point by ~20 pm just prior to evaporation, i.e., a change of ~10% from the zero-field bonding distance.

The electrically induced changes in atomic bonding energies discussed above can in principle lead to atomic-level surface reconstruction effects, in particular situations where lateral changes occur in the position of individual bonding sites. The characteristic bright zone-line decoration observed in W FIM images is thought to be an example of this. However, the original attempt [79] to explain this, by modeling PIP effects by an “internal” field-energy term, was made too long ago to be credible by modern standards.

Currently, the best theoretical approach to bonding-site modification issues of this general kind (and, more generally, to the problem of calculating the potential energy structures in which surface partial-ions move) appears to lie in CS-DFT approaches, similar to the supercell approach used by Sánchez et al. [56] to investigate field evaporation.

In their approach, impressed field-induced effects are created by placing planar layers of negative charge (equivalent to counter-electrodes) on either side of a central slab, with the system made neutral overall. Changes in total system energy are calculated as a function of the position of the nucleus of an adsorbed atom, which is made a partial-ion by the impressed field. This approach produces an integrated calculation of the effects described above as zero-field chemical effects, PIP effects, and field-energy effects—which in reality cannot be cleanly separated.

A present limitation of CS-DFT approaches is that it is difficult to accurately simulate situations where the three-dimensional environment of a kink site seems important. There may also be issues as to whether this “constant counter-electrode charge” approach sufficiently well simulates the experimental situation, which is one of constant applied voltage [80, 81].

CS-DFT theory can also be used to address wider issues concerning field-induced surface reconstruction [81]. However, it is not clear how effective current codes would be in investigating “cluster reconfigurations” [82, 83], such as the zig-zag chains sometimes observed on W(110) facets, which have been classically interpreted [82] as associated with local dipole–dipole repulsions.
2.4.4.3 Field-Induced Surface Atom Migration

The general driving force for the field-induced surface migration of surface atoms (as partial ions) is the drive to increase capacitance between the emitter and its surroundings. This is achieved by the movement of individual atoms from lower field sites to higher field sites, as has been demonstrated experimentally [84]. However, at an atomic level, for solid emitters, this movement is impeded by activation energy barriers that must be overcome by thermal activation. From the argument that the \( \Delta \Lambda^F \) term must be higher at the top of the activation energy barrier than at the lattice sites either side, it may be concluded that the energy barrier must be reduced by a high applied electric field, but quantitative details remain unclear.

From aspects of the discussion in [56, 85], it is clear that CS-DFT theory is able to address some issues of this kind, but again it is currently difficult to investigate situations where the three-dimensional environment of a kink site seems important. It seems clear that (subject to this limitation) CS-DFT theory is also potentially able to investigate reaction path geometry in field evaporation (Sect. 3.3.5).

2.5 Basic System Energetics of Field Ion Emission

In FI/AP theory, the energetics of ion and atom motion are issues of system energetics. Two approaches exist. The less general approach uses the concept of electron orbital level defined below. The more general system-PE approach works with functions \( U(r) \) that—for various charge-states of an “external atom” and for a given applied field—represent the PE of the atom + emitter + surroundings system when the nucleus of the external atom is at position \( r \). When both approaches are valid, the first is energetically equivalent to the second, and can be derived from it.

In what follows, “atom” means an atomic entity in any state of charge—neutral, fully ionic, or partially ionic.

2.5.1 Standard System Potential Energy Curves

The approach based on standard system potential energies was introduced into FI/AP theory by Gomer and Swanson [86, 87]. The simplest application is to FI imaging, taking the external atom as a noble gas atom. A more sophisticated version is used in field evaporation theory, taking the external atom as a metal atom.

The system energy reference zero is taken as the situation where (a) the emitter is notionally at zero temperature (i.e., all emitter electrons are considered to be in states below the Fermi level) and (b) the external atom is neutral, and in field-free vacuum distant from the emitter. Relative to this zero, the PE of a neutral (or “quasi-neutral”) atom at position \( r \) is denoted by \( U_0(r) \). The provision “quasi-neutral” is made because even a field-adsorbed noble gas atom may effectively carry a very small fractional charge [88].
The standard PE $U_n(r)$, of an ion of charge $ne$ at position $r$, is defined by

$$U_n(r) - U_0(r) = w(r),$$

(2.30)

where $w(r)$ is the work done in a slow, atomic-level, electrothermodynamic cycle [89]. In the cycle, the atom is formally removed from position $r$ to remote field-free space, $n$ electrons are removed (one by one) from the atom and placed at the emitter Fermi level, and finally the ion of charge $ne$ is returned to position $r$.

In the discussion above, it is implicit that the “neutral” and ionic states are both well defined, with both the neutral atom and the ion in their electronic ground states. For $n = 1$, Fig. 2.8 is a one-dimensional (1D) standard system-PE diagram that plots the PE curves $U_0$ and $U_1$ schematically along some line normal to the emitter surface. For simplicity, the ionic curve is represented as linear. The designation “standard” implies that all removed electrons are placed at the emitter Fermi level. In this 1D section of the three-dimensional (3D) distribution, $U_0$ and $U_1$ intersect at a crossing point. The 3D equivalent is that they intersect in a crossing surface that functions as an escape surface for the atom-to-ion transition: that is, if the ion finds itself on the “outer” side of the escape surface, then in all normal circumstances it is free and can move away from the emitter. A more quantitative discussion is presented below.

In principle, the Schrödinger equation may be solved for the nuclear motion in each PE curve and a set of eigenstates defined (which will be a continuum for the ionic curve). These standard PE diagrams are particularly useful for discussing issues of energetics. However, diagram variants (in which the ionic curves are...
adjusted vertically when a transferred electron enters a state with energy different from the Fermi level) may be needed for specialized purposes, such as the calculation of ion tunneling probabilities [87, 89].

In standard PE diagrams, so-called quasi-classical (or “vertical”) transitions, in which the nuclear motion is treated classically, and ionization is assumed to take place without any change in nuclear position or velocity, can be represented by vertical arrows, as shown in Fig. 2.8.

For an imaging gas atom, away from the immediate vicinity of the emitter surface, the 1D approximations often used are

\[ U_0(x) = -\frac{1}{2} \alpha_G F_N^2, \]
\[ U_1(x) = (I_1 - \Phi) - eF_x - e^2/16\pi\varepsilon_0 x, \]

Here, \( \alpha_G \) is the gas atom polarizability in free space, and \( I_1 \) is the first ionization energy (i.e., the energy needed to remove the topmost electron from the neutral gas atom, when this atom is in remote field-free space). \( F_N \) is the field at the position of the gas atom nucleus, and \( F \) is the relevant (average) surface field. In a 1D model, \( F_N \) and \( F \) are allocated equal values, but the two types of field are in principle different (and in real 3D situations would have different numerical values).

For a metal, the PE \( U_n(r) \) of a metal ion of charge \( ne \) can be written formally as

\[ U_n(r) = (H_n - n\Phi) + U_n^{\text{ES}}(r) + U_n^{\text{CHEM}}(r), \]

where \( H_n \) is the sum of the first \( n \) ionization energies, and \( U_n^{\text{ES}}(r) \) and \( U_n^{\text{CHEM}} \) are the electrostatic and chemical components of the ion PE. A common (but not necessarily accurate) basic 1D approximation for \( U_n(r) \) is

\[ U_n(x) = (H_n - n\Phi) - neF_x - n^2e^2/16\pi\varepsilon_0 x. \]

The issue of how to formulate an expression for the PE of a metal atom bound to a charged metal surface is not straightforward and is considered in Chap. 3.

### 2.5.2 Electron Orbital Level

For \( n = 1 \), the energy \( -w(r) \) in Eqn. 2.30 can also be interpreted as the electron orbital level \( \epsilon(r) \) of the topmost electron orbital in the neutral atom, measured relative to the emitter Fermi level [89]. This yields the definition

\[ \epsilon(r) \equiv U_0(r) - U_1(r), \]

which is illustrated in Fig. 2.8. The dependence of \( \epsilon \) on distance \( x \) in the 1D section is shown schematically in Fig. 2.9, which is an orbital-level diagram. The level \( \epsilon \) is
positive outside the crossing point, and negative inside it. At the crossing point, which (by definition) is at the critical distance $x_c$, $\epsilon = 0$.

For an imaging gas atom, away from the immediate vicinity of the emitter surface, Eqns. 2.31 and 2.32. yield the 1D approximation

$$\epsilon(x) = \phi + eFx - I_1 + e^2 / 16\pi\varepsilon_0 x - \frac{1}{2}\alpha_G F_N^2.$$

(2.36)

Equation 2.36 is sometimes simplified by defining an effective ionization energy $I_{\text{eff}}$ by Eqn. 2.37 below, and rewriting Eqn. 2.36 as Eqn. 2.38:

$$I_{\text{eff}} = I_1 - e^2 / 16\pi\varepsilon_0 x + \frac{1}{2}\alpha_G F_N^2.$$

(2.37)

$$\epsilon(x) = \phi + eFx - I_{\text{eff}}.$$

(2.38)

This simplification is used in labeling Fig. 2.9.

Close to the surface, additional terms would appear in these expressions. The last two terms in Eqn. 2.36 are smaller than the first three. Often either the polarization term or both terms are disregarded; obviously, if both are disregarded, then this is equivalent to putting $I_{\text{eff}}$ equal to $I_1$.

### 2.5.3 Critical Distance, Critical Surface and Forbidden Zone

The He-on-W system is often used to discuss FI imaging. Relevant data are $\phi \approx 4.5$ eV, $I_1 = 24.6$ eV, BIF $\approx 45$ V/nm. Calculation using the first three terms in Eqn. 2.36 shows that $\epsilon(x)$ changes from negative to positive at a distance $x^c \approx 450$ pm. In 1D theory, this distance $x^c$ is called the critical distance.

More generally, there is a surface in space, approximately 400–500 pm above the outer edges of the surface metal atoms, where the conditions $U_1 = U_0$, and hence $\epsilon(r) = 0$, hold. This crossing surface between $U_1$ and $U_0$ is called the critical
surface, Fig. 2.10. Normally, field ionization is not possible when a gas atom nucleus is inside the critical surface, because its topmost orbital level is below the emitter Fermi level and there are no empty emitter electron states for the gas atom electron to tunnel into. This region inside the critical surface is called the 

Forbidden Zone.

2.5.4 Static Energy Deficits and Appearance Energies

In APT, an ion may arrive at the detector with kinetic energy less than that corresponding to the voltage difference between the counter-electrode or detector and the emitter. With voltage-pulsed APT, part of this deficit may be related to dynamic effects associated with changing voltages in the system during the pulsing process, but there is also a static effect, which is discussed in this section.

To discuss static energy deficits, it is convenient to consider the experimentally derived appearance energies defined by Eqn. 1.1. In normal circumstances, a minimum observed value for energy deficit would exist, and some onset criterion could be used to define an experimental onset energy deficit $D_{on}$ and hence an experimental onset appearance energy $A_{on}$. In the APT context, this deficit manifests itself as the ion having less kinetic energy than might naively be expected from the voltage difference between the emitter and extraction electrode. Physically, the reason for the deficit is that the high-voltage generator has to provide the work needed to create the ion.

Predictions of onset energy deficit were first made in the context of noble gas field ionization, and there were initially some discrepancies in the theory. A more general theory of appearance energies, which applies to all field desorption processes (including field evaporation), was given by Forbes [90]. This theory predicts...
the value of the critical (or standard) appearance energy $A_{cr}^r$. This quantity applies to an ion (of charge $re$) formed near the emitter in such a way that the removed electrons make energy-conserving transitions to the emitter Fermi level. That is, in gas field ionization the ions are formed, with zero kinetic energy (KE), in the critical surface; in field evaporation they behave as if formed, with zero KE, at the pass at the top of an activation energy hump, where the reaction path intersects the escape surface.

For an ion of charge $re$, the general formula for $A_{cr}^r$ is [87]

$$A_{cr}^r = H_r - U^{pa},$$

(2.39)

where $H_r$ is the sum of the first $r$ ionization energies, and $U^{pa}$ is the standard PE for the desorbing entity at the position of the pass, as just defined. When ionization takes place by desorption from a bound state, as in field evaporation, this formula becomes

$$A_{cr}^r = H_r + \Lambda^F - Q_n,$$

(2.40)

where $\Lambda^F$ is the bonding energy (for point field $F$ at the bonding site) prior to desorption, and $Q_n$ is the activation energy associated with the escape process. Often $Q_n$ is small and can be disregarded.

Appearance energies have been measured accurately by retarding potential techniques, e.g., [91], and have provided unique information about bonding-site position [92], the location of the electrical surface [50], and about the energetics of the surface bonding and emission processes, e.g., [67, 93, 94].

### 2.6 Field Adsorption

A neutral gas atom or molecule near a field ion emitter moves in a long-range polarization potential energy (PPE) well, with the PE $U_0$ given by Eqn. 2.32. The well boundary can be defined by a suitable criterion, such as $|U_0| = k_B T_{gd}$, where $k_B$ is Boltzmann’s constant and $T_{gd}$ is the gas temperature distant from the emitter. At applied fields typical of imaging, this PPE well extends several or many tip radii into space above the emitter, and also significantly down the emitter shank.

The well captures neutral gas entities (both imaging gas atoms and impurity molecules) and creates a loosely bound field-adsorbed gas phase around the emitter tip, in which the gas concentration is higher than at large distances from the emitter. This long-range (or weak) field adsorption is a purely polarization effect.

Effects of this kind determine the operating gas supply to a gas field ion emitter/source. Forbes [3] discusses past work and provides a corrected version of Southon’s [95] theory of gas supply. More recently, Sujiyama et al. [96] have performed numerical calculations that examine the effect of emitter shank angle.

Close to the emitter surface, stronger short-range forces come into play and give rise to the phenomenon of short-range field adsorption (SRFA) (also called strong
At applied fields typical of imaging, individual atoms or molecules may be locally bound to specific sites above protruding metal surface atoms, including to surface atoms in relatively open crystal facets such as W(111). On closed-packed faces, field-adsorbed atoms may be bound into a mobile near-surface layer. Because (the nuclei of) these adsorbed atoms are well inside the forbidden zone, strongly field-adsorbed atoms are not field ionized.

SRFA bonding energies (i.e., the energy needed for a gas atom to be released from a local bonding site into the loosely bound field-adsorbed gas phase) are relatively small, typically of order 0.1 eV, and strongly field-adsorbed atoms are present only at sufficiently low emitter temperature [97, 98]—for He probably only at temperatures below ~100 K.

Becker [45] was first to suggest that gas atoms bound by polarization forces could be present on the surface during imaging, especially at high-field sites above protruding metal atoms. Panitz, in early AP experiments [99], found that field evaporation products included complex ions involving a metal atom and a noble gas atom (He or Ne). This suggested that, at sufficiently high applied field, the gas atoms were locally bound. A hopping-bright-spot phenomena [100–102], observed when a small amount of Ne was present in He, suggested that—on open crystal facets—neon atoms could be temporarily bound to a surface site, in the “apex” position above the underlying metal atom. AP experiments by Tsong and Müller [103] suggested that He (and hence all) noble gases could be locally bound, certainly to kink sites, with SRFA bonding energies typically of order 0.2 eV for He.

The theoretical origin of SRFA, although now largely resolved, was at one stage controversial. Tsong and Müller [46, 63] initially modeled it as a polarization effect, due to an interaction between field-induced dipoles in the field-adsorbed atom and the underlying metal surface atom. Forbes [104] confirmed that (for metal elements) the surface dipole in the metal atom seemed much more important than its monopole charge. (However, this may not be true for alloys, if strong charge-transfer effects occur.)

The simple dipole–dipole model neglects the potentially depolarizing effect of adjacent metal atom dipoles, but the array models then introduced, e.g., [47], could not predict plausibly high SRFA bonding energies.

Subsequently, Kreuzer and colleagues noted [105] that imaging fields of order 45 V/nm would lift the topmost He atom orbital into the metal conduction band: hence overlap-integral effects might give rise to a form of chemical bonding, and an inwards shift [92] of the adsorbed-atom bonding point. Later work [106] suggested that both chemical and polarization effects operated, possibly with polarization the stronger effect at fields near BIF. More recent work (see [88]) has confirmed that noble gas atoms adsorbed on metals may carry an effective charge.

With H, field adsorption effects of the same general kind can occur, and there is also the possibility that an adsorbed hydrogen molecule may dissociate into atoms [107] and bond chemically. More information about SRFA of noble gases, H and some other species can be found in [106–109].
The relevance of SRFA to FI/AP theory is as follows. In FIM, it appears that strongly field-adsorbed gas atoms provide intermediate collision partners and help to cool incoming imaging gas atoms: FIM image resolution is noticeably worse at emitter temperatures near 200 K, and ion energy distributions are slightly broader, than at temperatures near 80 K [97].

In APT, when there is a strongly field-adsorbed atom on the surface during field evaporation, then (as noted above) the evaporated entity may be a complex consisting of a noble gas atom or a H atom bound to a metal ion. Such complexes may break up in flight after desorption (see [109]). The possibility of resulting complications in APT analysis is one of the reasons why modern APT analysis is performed in ultrahigh vacuum conditions.

Strongly field-adsorbed atoms also alter the details of field evaporation theory, particularly values of bonding energies and evaporation fields (Sect. 3.2.1). As most classical FEV theory was developed from experiments made in the presence of He or another imaging gas, this fact may need to be remembered when applying classical FEV theory to APT analysis performed on adsorbate-free surfaces.

In principle, it is also possible for a weakly field-adsorbed atmosphere of neutral emitter atoms to be created near an emitter by overheating and thermal evaporation with the field applied. This happens with a liquid metal ion source [1].

2.7 Field Ionization and Post-Field-Ionization

2.7.1 Introduction

Field ionization (FI) is a radiationless energy-conserving process in which an electron tunnels through a barrier and out of a neutral atom or molecule, or out of an integrally charged ion, leaving behind an entity with one more elementary positive charge than before. FI of neutral atoms occurs in FI imaging; FI of metal ions occurs in field evaporation, where it is called post-field ionization (PFI) or (more usually) post-ionization (PI). Details differ, but the theory is similar in both cases.

FI is classified as (1) near-surface FI if the atom or ion nucleus is close to the emitter and the tunneling barrier is influenced by its closeness, or (2) free-space FI if the atom or ion nucleus is well away from the surface. In FI imaging and in FEV theory, near-surface FI occurs. However, if the applied field is high enough, impurity atoms and molecules approaching the emitter tip from the vacuum space are ionized by free-space FI: in a He FIM, this helps keep the emitter tip clean.

FI theory here (and in most FI/AP discussions) is quasi-classical. This means that electron behavior is treated quantum-mechanically, but the nuclear motion is treated classically. Strictly, this is not appropriate for the near-surface FI of gas atoms, as FIM conditions make full molecular-type quantum-mechanical analysis more appropriate. However, a quasi-classical approach is simpler and can be
justified (see Sect. 21.6 in [110]). In a quasi-classical approach, the electron tunneling rate–constant $P_e(r)$ measures how quickly an atom or ion would become field ionized if its nucleus were stationary at some position $r$.

The detailed quantum-mechanical theory of field ionization, initiated by Oppenheimer [111], is notoriously tricky and difficult [112], even for the ionization of atoms in steady electrostatic fields. From the beginning [113], FI/AP theory has always used simplified approximate treatments, based on the Jeffreys–Wentzel–Kramers–Brillouin (JWKB) approach. These are analogous to those used [114, 115] for tunneling problems in field electron emission and appear to derive ultimately from a mathematical method introduced (in astronomy) by Carlini in 1817 [116]. The first experiments on a discharge phenomenon from a positively electrified point are even earlier [117].

### 2.7.2 Surface Field Ionization Theory for Imaging Gas Atoms

In FI imaging, a strongly field-adsorbed atom usually lies between the imaging gas atom and the emitter, but tunneling treatments usually disregard this. For a field-adsorbed He atom, which has no electron resonance levels anywhere near the emitter Fermi level, this neglect is probably justified.

In older treatments of the near-surface FI of an imaging gas atom, the emitter was modeled by a simple Sommerfeld-type PE box and image-type interactions were included when defining the barrier. The barrier that exists along a line that passes through a gas atom nucleus at the critical distance and is normal to the model surface was illustrated in Fig. 2.3. Older treatments then approximated the barrier in some mathematically convenient way and applied a 1D, simple-JWKB approach. Older treatments of this kind are surveyed in [110]. The resulting predicted electron tunneling rate-constant $P_e$ can be written in the generalized form:

$$P_e = A_e \exp \left( -\nu b \cdot \left( \frac{I_{eff}}{F_{tb}} \right)^{3/2} \right),$$

where $b$ is the Second Fowler–Nordheim Constant [118] [$\cong 6.831 \text{eV}^{3/2} \text{V nm}^{-1}$], $I_{eff}$ the effective gas atom ionization energy defined earlier, $F_{tb}$ the surface field that defines the tunneling barrier, and $A_e$ a pre-exponential that was originally taken as a classical approach/attempt frequency [119] and is usually put equal to $10^{15}$ to $10^{16} \text{s}^{-1}$. $\nu(\nu)$ is a correction factor associated primarily with the particular shape assumed for the tunneling barrier.

Expressions for $\nu$ depend on model details. Alternative methods exist of determining both $F_{tb}$ and $\nu$, but in good models the parameter $\nu$ is a sensitive function of position, both across and normal to the surface. The variations in $\nu \cdot \left( \frac{I_{eff}}{F_{tb}} \right)^{3/2}$ determine both the $P_e$ variations in the critical surface (Sect. 2.5.3), and the falloff in $P_e$ with distance outwards from the critical surface. In turn, this falloff determines or influences the energy spread of emitted ions.
A slightly better approach [120] carries out an integration over the different directions radiating from the gas atom nucleus. This introduces a correction factor into $A_e$. Later, Lam and Needs [121] found errors in the details of earlier treatments and went on to develop an improved approach [122, 123] that used a self-consistent quantum-mechanical technique to calculate electron PEs above charged atomically structured Al surfaces. The potential due to an external atom nucleus was superimposed, and $P_e$-values calculated by using a simple-JWKB approach, and integrating over different radial directions, as in [120].

Before the early 1980s, there was no good theoretical explanation of the FIM atomic resolution capability (see [110]). Later models [110, 122–124] now coincide in predicting that, for an imaging gas atom with its nucleus in the critical surface, the $P_e$-variations are sufficient to explain atomic resolution, with the Lam and Needs work being the most convincing. Essentially, the FIM resolves atoms because: (a) the surface barrier is thinner, and the surface field and $P_e$ are higher, over partially ionized surface atoms than over points between them; and (b) on slightly open facets, at BIF, the ripple in $P_e$ is high enough to overcome blurring due to gas kinetics [110].

### 2.7.3 Post-Field-Ionization of Metal Ions

For PFI of metal ions, the $I_{\text{eff}}$ in Eqn. 2.41 is the effective ionization energy for the ion. Tunnelling barrier details are different, but the rate-constant formula can be put into the same general form. In PFI, the total probability of ionization ($\Pi_e$), as the ion moves away, is of interest. This is obtained by integration of $P_e$ along a representative ion path, taking ion speed into account. The best known calculations are those of Kingham [125], who calculated PFI probabilities as a function of “model field” for many metal elements, in various different charge states, and presented results in graphical form. For each ionic transition, a parameter of interest is the field $F_{m,m+1}$ at which the probability of PFI from charge-state $m$ to charge-state $m+1$ is 50%. These fields are listed in Appendix E. The Lam and Needs criticisms [121] imply that improvements could also be made to Kingham’s PFI theory, but the effects are expected to be small.

Strictly, all existing PFI theories are charge-hopping-type theories (Sect. 3.1), in which the PFI event is pictured as a single sharp “hop” of a complete electron out of the departing ion into the substrate. This event takes place on the vacuum side of the activation energy hump over which the ion escaped. However, if the departing ion is still sufficiently close to the emitter surface, it is also possible to envisage charge-draining-type theories in which an electron charge drains (relatively) slowly out of the departing ion, with PFI probability that may approach unity. Experimentally, it might be difficult to distinguish between such an event and direct charge-draining into a 2+ or 3+ state. The theory of charge-draining-type PFI mechanisms has never been investigated in detail.
2.8 Field Ion Imaging

2.8.1 Introduction

This section builds on Sect. 1.1.2, by giving further details of FI imaging theory. The treatment is based on the ideas in [104, 110]; these are summarized in [3, 126]. As the length scale for significant ionization variations is much smaller than an atomic radius, reference to gas atom position relates to the position of its nucleus.

As shown in Fig. 2.10, at BIV most ionization takes place in a thin ionization layer just outside the critical surface. Layer thickness is of order 10 pm. In the layer, strong local variations exist in the ionization density (count of atoms ionized per unit volume). Most ionization occurs in disc-like ionization zones centered above nuclei of protruding metal surface atoms. As noted above, a field-adsorbed atom may lie between the ionization zone and the underlying metal atom.

2.8.2 Contributions to Emission Current Density

For BIV and nearby voltages, the ion arrival map at the detector (and hence the FI image) is a blurred map of the distribution of the ion flux density and hence the emission current density (ECD) (current per unit area) leaving the ionization layer.

The ECD $J_A$ for a point “A” in the critical surface is determined by the ionization at A and at related positions close above A and is given by

$$J_A = \left(\frac{e}{n_1}\right) C_{G,A} P_{e,A} \delta_A.$$  (2.42)

$C_{G,A}$ (called the gas concentration at point A) relates to the probability per unit volume of finding the gas atom nucleus very near A and is measured in “atoms per unit volume”); $P_{e,A}$ is the rate-constant for an atom with its nucleus at A; and $\delta_A$ is a decay length associated with how quickly the product $C_{G}P_e$ falls off with distance outside the critical surface, for point A. The constant $n_1$ is included for dimensional consistency and is best read as “1 atom.”

The decay length $\delta_A$ varies little with position in the critical surface, so the ECD variations are mainly determined by the variation, across the emitter surface, in the critical-surface values of the product $C_{G}P_e$.

Gas-kinetic and ion-optical effects (Sect. 2.10) cause the emission associated with each point in the critical surface to become blurred into a disc at the ion detector. The disc radius depends on the mean lateral kinetic energy ($\kappa_{av}$) of the gas atom population subject to ionization, at the instant of ionization; an effective gas temperature $T_g \sim \kappa_{av}/k_B$ can be defined by $\kappa_{av}$. This temperature $T_g$ will be influenced by the emitter temperature $T_e$, but usually $T_g > T_e$. This blurring effect is illustrated in Fig. 2.11: for $T_e$ near 80 K, the image spots are bigger than for $T_e$ near 5 K.

The primary influence on the relative intensities and “resolution” of neighboring emission sites (usually related to individual surface atoms) are the short-scale,
across-surface variations in ECD. Except at very low gas temperature $T_g$ (well under 20 K), the dominant cause of short-scale ECD variations is the variation in $P_e$ with position in the critical surface, as discussed in Sect. 2.7.2. At very low temperatures, the short-scale variation in $C_G$ also plays a role. Changes in local $C_G$ values are responsible for some of the spot intensity changes seen in Fig. 2.11.

Differences in ECD on larger image length scales (more than a few atomic diameters) are—at all temperatures—mostly associated with the imaging gas dynamics, which cause across-surface variations in the local mean value of $C_G$.

### 2.8.3 Imaging Gas Behavior

With a FIM, the current–voltage characteristics often have the “two-regime” shape shown in Fig. 2.12. A FIM is operated in the upper regime, sometimes called the supply- and capture (SAC) regime. In the SAC regime, the typical history of gas atom motion has three main stages—capture, accommodation, and diffusion, as follows. (1) The gas is captured on the emitter shank and then moves to its tip, heating up as it does so, because it gains kinetic energy from the polarization PE well. More of the gas supply is initially captured by the shank than by the tip (see [3], Appendix 1). (2) This hot trapped gas then cools, by transferring kinetic energy to the substrate when the atoms bounce, and accumulates into the higher field regions above the emitter tip. (3) As the gas becomes fully accommodated to emitter temperature $T_e$, across-surface diffusion takes place close to the surface, and gas concentrations build towards those characteristic of a thermodynamic equilibrium across the emitter tip as a whole. This description derives from detailed analysis of voltage and temperature dependences in FIM images [104, 126, 127].

The applied field (i.e., the emitter apex-field value) $F_a$ controls the electron tunneling rate-constant values and hence determines the point (on average) in the gas atom history at which FI occurs. Thus, $F_a$ determines the distribution of the gas, in space and in energy, at the instant of ionization.
A working rule proposed by Forbes in 1971 [127] implies that regional (and some smaller scale) image intensity differences are determined by the gas fluxes trapped into these areas during capture and accommodation, but that short-scale intensity variations are influenced by a quasi-equilibrium distribution of gas within relatively small areas of the emitter. This rule follows physically if BIV corresponds to a point near the end of the accommodation stage. In this case, the gas temperature at ionization might be expected to be slightly above emitter temperature. The temperature issue is very complex and has never been fully decided.

The dim-ring phenomenon shown in Fig. 2.13 is a good illustration of this rule. The very bright emission site in Fig. 2.13b (associated with a deposited O₂ molecule) has “turned off” the imaging of the nearby rings, which were visible in its absence (Fig. 2.13a). The presumed explanation is that the dim net-plane edges each constitute an extended potential well (a confine) into which gas atoms become trapped during the accommodation stage, with enough time to “run around the rings” several or many times. A very high ionization rate-constant near the O₂ molecule means that the ionization probability for a gas atom approaching the molecule is very high compared to other locations in the ring, and consequently this emission site draws off most of the gas supply captured into the confine.

Support for the idea that both gas effects and ionization effects are involved in the physics of best image field (BIF) is provided in Table 2.3. If gas distribution effects are important, the polarization PE \((1/2)\alpha G(F_{BIF})^2\) might be expected to be nearly constant for the various imaging gases (this has been called the assumption of corresponding potential structures). If rate-constant effects are important, the (approximate) tunneling exponent \(bf^{3/2}/F_{BIF}\) might be expected to be nearly constant for different imaging gases. In fact, as shown in Table 2.3, when measured \(F_{BIF}\) values are used, both these quantities have fairly similar values for all gases.

The characteristic FIM image intensity patterns for different metal elements seem to be largely a consequence of how the field evaporation endform for the
material influences field distribution details, and hence gas distribution effects, although ionization effects probably influence the sharpness of intensity changes at the boundaries of bright regions.

More general consequences of the working rule are that any emitter shape modifications that change the field distribution in which gas atoms move, and any surface condition changes that alter rates of gas accommodation, are likely to alter the distribution of emission current density. In particular, one can prepare emitter shapes that guide nearly all the gas atoms to the emitter apex, as in the ALIS™ gas FI source [4]. The theory of field ion imaging is also, of course, the theory of the gas field ion source. For further discussion and references, see [3, 128, 129].

Table 2.3  Comparison of measures of the polarization PE well and the ionization rate-constant, for the noble gases

<table>
<thead>
<tr>
<th>Gas</th>
<th>I (eV)</th>
<th>(a_G) (meV V(^{-2}) nm(^2))</th>
<th>(F^{\text{BIF}}) (V/nm)</th>
<th>(a_G(F^{\text{BIF}})^{2/2}) (meV)</th>
<th>(bf^{3/2}/F^{\text{BIF}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>He</td>
<td>24.6</td>
<td>0.143</td>
<td>44</td>
<td>138</td>
<td>18.9</td>
</tr>
<tr>
<td>Ne</td>
<td>21.6</td>
<td>0.275</td>
<td>35</td>
<td>168</td>
<td>19.6</td>
</tr>
<tr>
<td>Ar</td>
<td>15.8</td>
<td>1.14</td>
<td>18.5</td>
<td>197</td>
<td>23.1</td>
</tr>
<tr>
<td>Kr</td>
<td>14.0</td>
<td>1.75</td>
<td>14</td>
<td>171</td>
<td>25.6</td>
</tr>
<tr>
<td>Xe</td>
<td>12.1</td>
<td>2.78</td>
<td>11</td>
<td>168</td>
<td>26.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.6</td>
<td>0.09</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>1.41</td>
<td>1.38</td>
</tr>
</tbody>
</table>

For consistency, the old measured value of helium BIF is used

![Fig. 2.13](image)

Micrographs showing a “dim-ring effect” that illustrates the plausibility of the Forbes (1971) working rule for explaining the formation of FIM image contrast. Image (a) is a platinum endform imaged with helium, near 80 K and at BIF. Image (b) is a similar endform, but with a deposited oxygen molecule that has caused brilliant emission and has “turned off” the imaging of two net-plane rings. The whole gas supply captured (during gas accommodation) into the polarization PE wells associated with the net-plane edges is being emitted in the vicinity of the oxygen molecule, due to the creation there of an emitting feature with a very high ionization rate-constant. Image (a) was taken after the oxygen molecule seen in image (b) had been field desorbed. These micrographs were taken in the late 1960s by Father C. Schubert, S. J.
2.9 Field Calibration

2.9.1 Introduction

It is useful to discuss field calibration at this stage. Close above an operating field ion emitter, point fields (however defined) vary sensitively with position. Accurate measurement or calibration of the various fields used in FI/AP theory poses immense difficulties, and only limited precision has been achieved. For a given metal-element emitter of moderate radius, the best image voltage (BIV) and the field evaporation onset voltage at a given temperature can be measured relatively accurately (typically to 2%). The problems are to convert these measurements to absolute values of “field,” and to ensure that calibrations performed on one emitter can be applied to others.

In FI/AP theory and practice, most (though not all) field estimates rely on the following principles: (1) that, for a given emitter, the “field” in question (“$F$”) can be taken as proportional to the applied voltage, according to the formula $F = \beta V$, where $\beta$ is the relevant voltage-to-field conversion factor; and (2) that, for any given imaging gas (for emitters of moderate radius), the BIF is the same for all metal elements.

It follows that if, for a chosen imaging gas, the “field” corresponding to a chosen best imaging situation can be determined, then the values of this type of field for other applied voltages and emitters can be found. Also, by imaging a given emitter with different imaging gases, the BIFs for these gases can be found from voltage ratios. These principles guide the first two calibration methods discussed below. Historically, the chosen system was the He imaging of a (110)-oriented W emitter system.

There is no known satisfactory way of relating calibrations of nanoscale-level electrostatic fields to calibrations of macroscopic fields. Rather, one has to use a nanoscale phenomenon that has a well-established (or, at worst, “adequate”) theory containing a “field” as a parameter. There are three obvious candidates: field electron emission, field ionization, and PFI. In principle, field-evaporation escape theory might be thought a fourth option, but currently there are too many uncertainties about escape theory for this to be a useful approach.

2.9.2 Calibration via Field Electron Emission

All older (pre-1973) field calibrations in fact rely on a 1961 paper by Müller and Young (M&Y) [130]. M&Y measured the field electron emission (FE) current density at a specified applied voltage and then determined $\beta$ for their emitter via the Fowler–Nordheim-type (FN-type) equation developed by Murphy and Good [131, 132]. M&Y used the current-density tables of Good and Müller [133] to find the characteristic local surface field for FE and obtained $\beta$ from Eqn. 2.44.
The value M&Y obtained for the He BIF was 44 V/nm, but (because the estimated accuracy was ±15 %) this has often been rounded to 45 V/nm, e.g., [134]. M&Y’s estimate of accuracy derives from Dyke and Dolan’s [135] estimate of the accuracy of their FN-type equation, which is trivially different from that in [131]. The estimate is based on comparisons with earlier experiments [136] that used electron microscope profiles of field emitters. The original 1953 comparison [136] is slightly flawed, due to undetected errors then present in FE theory [137].

There is a question as to whether the field in the theoretical model of the FE tunneling barrier is the same physical quantity as the average field \( F_{tb} \) in the tunneling barrier between the helium atom and the emitter. Since FE tunneling barriers are typically 1 nm in width, it is arguable that the value 45 V/nm is best taken as a measure of the “external field” \( F_{ext} \) as defined earlier (i.e., the point field in the critical surface). If so, as assumed here, then the ±15 % needs to be taken as an estimate of the accuracy of \( F_{ext} \).

On the basis of charged-surface modeling, it has been argued [110] that, directly over emitting atoms, \( F_{tb} \) for He FI is somewhat greater than \( F_{ext} \), perhaps by 10 % or more. This is qualitatively compatible with the physical existence of SRFA. The results together indicate the possibility of systematic underestimation of \( F_{tb} \)-values.

For field evaporation from W emitters of apex radius 20 nm or greater, at 77 K, with He present, M&Y also derived the experimental onset evaporation field for W FEV (with He present) as 54 V/nm. Under the same conditions, Van Oostrom [138] derived experimental values of 57 V/nm for a thermally annealed tip and values between 67 and 73 V/nm from different regions on a tip with a field-evaporated endform, where different endform regions have different regional radii. Müller and Tsong [128] thought Van Oostrom’s evaporation fields to be too high, but conceded that a new experimental BIF estimation might result in a value as high as 50 V/nm [128, p. 156]. As before, these evaporation-field values may be estimates of \( F_{ext} \), rather than surface field as used in FEV theory. It is also now known [139] that correction factors are missing from the Murphy and Good FN-type equation. This historical situation well illustrates the difficulties of field measurement and calibration in field ion emission.

### 2.9.3 Calibration via Free-Space Field Ionization

A second, and hopefully more accurate, field calibration method was developed by Sakurai and Müller [134, 140]. This used measurements of the energy deficits associated with free-space field ionization (FSFI), above some defined region of the emitter. Deficits were measured at the position where the FSFI distribution peaked. Derived formulae enabled a “local shape factor,” analogous to \( k_f \) in Eqn. 2.7, to be found. A “regional emitter radius” was determined by ring counting (Sect. 1.2.6.1), and an equation similar to Eqn. 2.7 was then used to determine a regional value of field above that surface region. We can assume that this field
2.9 Field Calibration

adequately coincides with the mean critical surface field (i.e., mean $F_{\text{ext}}$-value) for this region.

This approach has been used [140] to determine the $F_{\text{ext}}$-values relevant to BIF and to field evaporation onset (defined as an evaporation flux of 1 layer/s [141]), for various emitter regions, for several metal elements at 21 and 78 K. In particular, the values obtained for W at 78 K lie in the range 55 V/nm [for the (110) face] to 62 V/nm [for the (111) face]. The complete list of onset evaporation fields found [140] is shown in Table 2.4. A misprint in the table in [140] has been corrected by using the corresponding table in [141].

This 1977 work [140] is the origin of the sometimes used values of 45 V/nm for the He-on-W BIF and 57 V/nm for the W onset evaporation field. BIF values for other noble gases were recorded in Table 1.1.

Castilho and Kingham [142] have suggested small improvements and have also proposed an alternative approach based on attempting to calculate BIFs from first principles; more generally, their calculations tend to support the basic validity of Sakurai and Müller’s method.

### Table 2.4

<table>
<thead>
<tr>
<th>Metal</th>
<th>Gas</th>
<th>$T_e$ (K)</th>
<th>FEV onset field (V/nm) for:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(001)</td>
</tr>
<tr>
<td>W</td>
<td>He</td>
<td>78</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>57</td>
</tr>
<tr>
<td>Mo</td>
<td>He</td>
<td>78</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>50</td>
</tr>
<tr>
<td>Ir</td>
<td>He</td>
<td>78</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>54</td>
</tr>
<tr>
<td>Rh</td>
<td>He</td>
<td>78</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>48</td>
</tr>
<tr>
<td>Ni</td>
<td>He–Ne</td>
<td>78</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>35</td>
</tr>
<tr>
<td>Pt</td>
<td>He</td>
<td>78</td>
<td>48</td>
</tr>
</tbody>
</table>

For field evaporation theory, a problem with all the above methods is that they estimate external fields rather than the surface field used in FEV escape theory (Sect. 3.2.3). The “field” in PFI theory is expected to be numerically closer to this surface field than to $F_{\text{ext}}$. Thus, in principle, an alternative method exists when pure metal emitters field evaporate in a mixture of adjacent charge states, with the higher charge state formed by charge-hopping-type PFI. By measuring the proportion of the more highly charged ion, a field value can be estimated from Kingham’s
diagrams [125], such as those reproduced in Appendix E. A small error may exist, for the reasons discussed by Lam and Needs [121], but one might reasonably hope that this is less significant than making a proper distinction between external field and FEV surface field. This PFI method has been applied to liquid metal ion sources [143], though with mixed results. It has had some use in APT, for example [144], and may merit more attention.

2.10 The Charged-Particle Optics of Field Emitters

2.10.1 Introduction

The theory in this section is a summary of the basic “optical” theory that applies to charged-particle (CP) emission from well-behaved field emitters. It applies to both electrons and ions and follows the approach used by Hawkes and Kasper (HK) [33] for electrons, as developed further by Forbes [3, 145] for CPs in general. Common theory is possible because, in basic CP optics, the trajectories of the emitted particles do not depend on their charge-to-mass ratios. Common theory is useful because it allows results obtained with electrons to inform procedures using ions, and vice versa. Fuller accounts of parts of the theory presented here, and more CP optical background, may be found in [3, 33, 145].

A real emitter has an optical surface, where an emitted CP is deemed to emerge before it picks up any kinetic energy from the electric field. In the HK approach, the real emitter is modeled optically by combining a spherical charged-particle emitter (SCPE) with a formal weak converging lens. The SCPE models the emitter’s optical surface; the lens compresses the emitted beam. For real ion emission, the optical surface is the escape surface discussed in Sect. 2.5.1. In the model this is smoothed out into a spherical surface. The discrepancy between this and the real optical surface causes aberrations in the model results.

If CP emission is regarded as incoherent, then CP motion can be treated as that of an electrified point, and basic discussion can use ray-like (rather than wavelike) theory. Each point $P_0$ on the SCPE then has an associated trajectory (shown by a bold line in Fig. 2.14) that describes the motion of a CP that leaves with zero kinetic energy parallel to the optical surface. This principal trajectory (for point $P_0$) starts normal to the surface.

Emitted CPs in fact emerge with a distribution of lateral kinetic energy (KE) $\kappa$, with the nature of the distribution dependent on the emission mechanism. There is a related distribution of starting angles $\alpha'$ (measured relative to the normal). To characterize the lateral KE distribution, a critical lateral kinetic energy $\kappa_c$ is defined (for example, by specifying that 50% of ions have $\kappa < \kappa_c$, or that $\kappa_c = k_BT_{lat}$ where $T_{lat}$ is the effective lateral temperature of the escaping particle at ionization).

Four aspects of field emitter optics need discussion: (1) the behavior of principal trajectories, and related issues of focusing and image magnification; (2) the
relationship between optical theory and the projection methods used to index field emission images, and also in APT reconstruction formulae; (3) issues relating to image spot size and to the resolving ability of the projection techniques (particularly FIM and APT); and (4) issues relating to aberrations. These are covered in the following subsections, though not on a strict one-to-one basis.

2.10.2 Operation of the Spherical Charged-Particle Emitter

The SCPE, Fig. 2.15, has no analogy in photon optics, but plays a key role in field emitter optics. SCPE optical behavior has been analyzed by Ruska [146] and many others, e.g., [31–33, 147]. At large radial distances from an SCPE of radius \( r_a \), CPs emitted from \( P_0 \) appear to be diverging from point \( P_1 \) on a sphere of radius \( r_a/2 \), and the starting angle \( \alpha_{B'} \) that corresponds to \( \kappa_c \) is

\[
\alpha_{B'} = 2(k_c/neF_a r_a)^{1/2} = 2(k_a \kappa_c/neV)^{1/2},
\]  

(2.43)
where \( ne \) is the magnitude of the particle charge, \( F_a \) is the magnitude of both the apex field for the real emitter and the surface field for the SCPE, and \( V \) is the magnitude of the voltage applied to the real emitter. For consistency with other optical theory \([3]\), the shape factor (field factor) is here denoted by \( k_a \) rather than \( k_f \). The angle \( \alpha_{B0} \) has been called \([3, 145]\) the blurring falloff half-angle (blurring FOHA). The single prime indicates that the primed parameter relates to the optics of the SCPE.

In reality, in FI emission contexts, an emission source on sphere \( S_0 \) has a finite size. For a source on the emitter axis, this can be quantified by a small distance \( \rho_0 \) in the surface of \( S_0 \), as shown in Fig. 2.15. In FIM, \( \rho_0 \) relates to the distribution of emission current density associated with the critical surface (see Sect. 2.5.3); in APT it relates to the atomic-level statistics of small differences in the path followed by a field evaporating atom as it escapes. (Little is known about this, but one might estimate \( \rho_0 \) from the atomic vibration amplitude at the bonding site.) As before, a specific criterion is needed; for example, the criterion for FIM spot-size analysis might be that a disc of radius \( \rho_0 \) contains 50% of the site emission current.

In Fig. 2.15, disc edges would be represented by the bold points either side of \( P_0 \). The related source FOHA \( \alpha_{S'} \) is

\[
\alpha_{S'} = \rho_0/r_a.
\]  

(2.44)

Optically, the real source is on \( S_0 \) and the SCPE creates a virtual Gaussian image of this source on sphere \( S_1 \); this virtual image lies between points \( P_1^- \) and \( P_1^+ \).

Since each point on \( S_0 \) creates a cone of emitted particles, the effect at sufficiently large distance from the emitter is that the original source FOHA \( \alpha_{S'} \) has been blurred into the larger value \( \alpha_{T'} \). Strictly, the value of \( \alpha_{T'} \) needs to be determined by

![Fig. 2.15 Schematic diagram illustrating of the optical operation of a spherical charged-particle emitter (SCPE). (See text for details.) (Copyright 2009 from Handbook of Charged Particle Optics by R. G. Forbes. Reproduced by permission of Taylor and Francis Group, LLC, a division of Informa plc.)](image-url)
a convolution, but for blurring of a circular emission site, if \( \alpha_S' \) and \( \alpha_B' \) are comparable in size, then it is thought acceptable to use the formula:

\[
(\alpha_T')^2 \approx (\alpha_S')^2 + (\alpha_B')^2.
\] (2.45)

A blurring ratio (also called an objects size ratio) \( m_B \) is defined and given by

\[
m_B \equiv \frac{\alpha_B'}{\alpha_S'} = \frac{(2/r_0)(\kappa c r_a/neF_a)^{1/2}}{(2/r_0)(\kappa c k_a/neV)^{1/2}},
\] (2.46)

and a blurring magnification \( m_T \) by

\[
m_T \equiv \frac{\alpha_T'}{\alpha_S'} \approx (1 + m_B^2)^{1/2}.
\] (2.47)

### 2.10.3 Operation of the Weak Lens

For simplicity in what follows in Sect. 2.10, it will be assumed that the emitter axis is parallel to the optical axis of a more complete system, that this system contains an ion arrival plane that is normal to the system axis, that ions travel between the emitter and the arrival plane in time-independent electrostatic fields, and that there are no extra electrodes that provide extra focusing effects in the region between the emitter and the arrival plane. In the arrival plane, the pattern of arrival of the emitted electrons or ions creates an arrival map, which may be a field electron, field ion, or field evaporation “projection image.” (It is not, of course, a focused image, in optical system terms.)

It is well known that, with a field emitter, the effect of the emitter shank is to compress the arrival map, by bending the trajectories inwards. Thus a principal trajectory begins, at the emitter, at an angle \( \theta_P \) to the emitter axis; \( \theta_P \) is the launch angle. When it reaches the arrival plane, it is traveling at a smaller angle \( \theta_E \) to the emitter axis; \( \theta_E \) is the arrival angle. Most of the bending takes place relatively close to the emitter (e.g., [33]).

Hawkes and Kasper, HK [33], argue that this effect can usefully be represented as the action of a weak converging lens. This takes the Gaussian image near \( P_1 \) as a virtual object and creates a virtual image near a point \( E \), some distance behind the emitter apex. Using an SOC model for the emitter, numerical calculations by Wiesner and Everhart [32] located \( E \) as 0.3–0.6 mm behind the center of the core sphere.

There is an issue of how to define the (radial) angular magnification of this lens. In principle, three alternatives exist: the Helmholtz formula of photon optics, namely, \( \tan \theta_E = m_H \cdot \tan \theta_P \); a formula proposed by HK, namely, \( \tan \theta_E = m_{HK} \cdot \theta_P \); and the angle-ratio formula \( \theta_E = m_{AR} \cdot \theta_P \). Choice between them needs to be decided on the basis of which formula best represents actual field emitter behavior.

There exists extensive evidence that, out to about 50°, radial distance as measured from the center of an arrival map is, in practice, linearly related to the launch
angle \( \theta_P \). This evidence comes from the analysis of field electron \([148]\) and field ion \([\text{e.g.,} [149–151]]\) micrographs, from numerical computations performed on the trajectories of electrons \([\text{see} [33]]\) and ions \([152]\), and from field evaporation maps (Sect. 7.1.6 in \([35]\)). Although no analytical explanation has yet been found/devised for it, this effect is certainly real and confirmed.

The implications are that the HK formula and HK angular magnification are the correct ones to use and that (for a field emitter with its axis normal to the arrival plane) radial distance \( R_P \) in the arrival plane is given by

\[
R_p = L_E \tan \theta_E = L_E m_{\text{HK}} \theta_P,
\]

where \( L_E \) is the distance from E to the arrival plane.

Since, in relation to Fig. 2.14, the radial distance \( \rho_P \) of \( P_0 \) from the emitter axis is \( \rho_P = r_a \sin \theta_P \), a displacement magnification \( \lambda \) is defined and given by

\[
\lambda \equiv R_P / \rho_P = m_{\text{HK}} (L_E / r_a) (\theta_P / \sin \theta_P) = m_{\text{HK}} L_E / r_a \sin(\theta_P).
\]

Clearly, in the HK model, this displacement magnification is a function of the launch angle. Equation 2.48 also yields the reverse formula:

\[
\rho_P = r_a \sin (R_P / m_{\text{HK}} L_E).
\]

Since a small change \( \delta \theta_P \) corresponds to a small arc of length \( \delta l_S \) \( = r_a \delta \theta_P \) in the SCPE surface, the related small radial distance \( \delta R_S \) in the arrival plane is

\[
\delta R_S = L_E m_{\text{HK}} \delta \theta_P = (m_{\text{HK}} L_E / r_a) \delta l_S \equiv \mu_{\text{rad}} \delta l_S,
\]

where the local radial projection magnification \( \mu_{\text{rad}} \) is defined and given by Eqn. 2.51. In the basic discussion here, \( \mu_{\text{rad}} \) is independent of angle \( \theta_P \).

In APT, compression effects are usually described by an image compression factor (ICF) \( \xi \) (or, in older literature, \( \beta \)) that is defined as the reciprocal of the angle-ratio magnification \( m_{\text{AR}} \). Consequently, the ICF is expected (from arguments above) to vary with radial position in the arrival plane. However, in the limit of small angles:

\[
m_{\text{HK}} \approx m_{\text{AR}} = 1 / \xi;
\]

hence, stated typical ICF values can be used to make estimates of \( m_{\text{HK}} \).

Conventional wisdom, e.g., \([34]\), is that the ICF is typically \(~1.5\), which suggests that \( m_{\text{HK}} \) is typically \(~2/3\). However, recent work \([144]\) found a range of values between 1.3 and 1.9 and also that—during prolonged field evaporation of a given emitter—the ICF dropped slightly, for example, from \(~1.6\) to \(~1.4\) as apex radius increased from \(~40\) to \(~80\) nm. This suggests that HK angular magnification might normally lie in the approximate range 0.5–0.8. There was also interesting evidence \([144, 153]\) that changes in ICF were linked to changes in the shape factor.
It needs to be pointed out that current AP reconstruction procedures, Sect. 6.4, are not based on the CP optical science formulae set out above. Rather, current state-of the-art algorithms derive from early work [154, 155] that itself derives from the essentially phenomenological point projection methods originally introduced into FIM in the context of indexing field ion micrographs, as discussed earlier. In many cases the existing reconstruction algorithms appear to be sufficient for purpose, but a longer term hope is that a blend of existing practice and CP optical science arguments might lead to improvements.

2.10.4 The Link Between Optical and Projection Methods

In Fig. 2.16, L₁ is the back-projection, towards the emitter axis, of the final part of the principal trajectory from P₀, after compression has occurred. L₂ is a line drawn parallel to L₁, in such a way that L₂ passes through P₀; L₂ intersects the system axis at P₄, a distance N_rₐ behind the SCPE center at P₂. Because Fig. 2.16 is schematic, the positions of E and P₄ are not realistic: in reality, L₁ straightens out much further away from the emitter, in such a way that E is much further away from P₂, and P₄ is slightly closer to P₂, than shown. Southworth and Walls [149] suggested that (for the metal emitters in use in the 1970s) N was typically ~0.8; in such cases, point P₄ is inside the dotted circle.

As is well known, e.g., [34], in the field ion micrographs of metal elements the dark centers of net-plane-ring systems are often called poles, because the normals to the corresponding crystal facets represent defined crystal lattice directions. The arrangement of poles in a field electron or field ion micrograph is topologically similar to those in a stereographic projection. Due to the known relationship between radial distance in the micrograph and angle in the emitter, it was tempting in the early days of field ion microscopy (especially for materials scientists) to try to explain FIM image formation by a projection relationship that used a fixed point on the specimen axis.
It was eventually agreed that no such relationship exists and that $N$ in Fig. 2.16 is a function of launch angle $\theta_P$ [34, p. 113]. The implication is that if naïve methods, based on fixed-point projection, are used to generate a relationship between $P_P$ and $R_P$, then this relationship will not be compatible with the known facts about field emitter optics, particularly for large angles $\theta_P$.

From this viewpoint, the continued use of fixed-point projection arguments in APT reconstruction procedures (rather than formulae based on CP optical arguments) would seem a longstanding but rectifiable scientific mistake—although it is certainly not the most critical problem in reconstruction theory (see Chap. 6).

### 2.10.5 Spot Size and Resolving Power

#### 2.10.5.1 Basic Spot-Size Formulae

Spot-size issues relate both to the size of “image spot” formed in the arrival plane and to the use of observed spot sizes and/or characteristics to make deductions about the size or apparent size of emission sites at the emitter surface. In FIM, the relevant current-density distributions are real; in APT, the spots are probability distributions for where a point-like ion from a given lattice site may arrive, and for making estimates of precisely where it has come from. Resolving power relates to whether emission from adjacent or neighboring emission sites (or, for APT, lattice bonding sites) can be detected as coming from separate sites.

If there were no blurring effects, then the radius $\delta R_S$ of an image spot in the arrival plane would be given by Eqn. 2.51, with $\delta \theta_P$ set equal to $\alpha_S'$. When blurring due to lateral velocity effects occurs, then the image spot radius $\delta R_T$ is given by

$$
\delta R_T = L_E m_{HK} \alpha_T' = m_T m_{HK} L_E \alpha_S' = m_T \mu_{rad} \delta l_S \equiv \mu_{rad,T} \delta l_S, \tag{2.53}
$$

where, as before, $\delta l_S$ is the length of arc relating to the actual size (here radius) of the emission site; $\mu_{rad,T}$ is defined by Eqn. 2.53 and can be called the (radial) spot magnification. From Eqns 2.47 and 2.46 it can be seen that spot radius is expected to increase with temperature and decrease with applied voltage or field. Effects of this kind are observed in FIM images, but interpretation is complicated because, for a given emission site, $\delta l_S$ may also be a function of field and/or temperature.

What is often of more interest is the apparent size $\delta l_T$ (here radius) of the emission site, as a result of blurring, and/or the radius $\delta l_B$ of an optical blurring disc defined in the following way (assuming incoherent ray-like emission): looking back at the emitter from the detector, the emission from each point appears to be smeared out into a disc of radius $\delta l_B$.

It is unclear whether the exact CP optics of determining $\delta l_B$ has ever been fully investigated, but it seems that (to an adequate approximation) one can take $\delta l_B \approx \alpha_B' r_a$ and use Eqn. 2.43 to obtain
\[ \delta l_B = \alpha_B' \cdot r_a = 2(\kappa_cr_a/neF_a)^{1/2} = 2r_a(k_a\kappa_c/neV)^{1/2}. \] (2.54)

The apparent total radius \( \delta l_T \) should in principle be obtained by convolution, but if \( \delta l_S \) and \( \delta l_B \) are comparable then it is thought acceptable to use the formula:

\[ (\delta l_T)^2 = (\delta l_S)^2 + (\delta l_B)^2. \] (2.55)

For ion emission, the most usual approximation is to put \( \kappa_c = k_B T_c \), where \( k_B \) is Boltzmann’s constant, and \( T_c \) is: for APT, the emitter temperature at the instant of ion field evaporation; for FIM, the effective gas temperature at the instant of ionization (which, at BIV, is slightly above the emitter temperature [3]). However, for FIM, Forbes [104] has argued that during emission the motion of the ion nucleus ion may need to be described wave-mechanically, and consequently there may be a lateral zero-point energy \( \kappa_{\min} \) associated with the lateral component of the ion wave function. This is equivalent to the existence of an effective minimum temperature \( T_{\min} \). The primitive model used in [104] suggested that \( T_{\min} \) might be less than 20 K. A similar argument may apply to field evaporation, but with a different minimum temperature, but this issue has never been investigated.

### 2.10.5.2 Application to Resolving Power Issues

The issues of resolving power are different for APT and FIM. For APT the issue of bonding site identification is considered to be much more strongly influenced by the variations in emitter endform shape (and resulting variations in ion trajectories) that occur during the evaporation of a single atomic layer and can be conceptualized as aberrations (Sect. 2.10.6). Consequently, thermal blurring effects have been disregarded. If they were taken into account, then they would generate additional uncertainty in reconstructing the original lateral coordinates of a detected ion.

For the illustrative values \( r_a = 40 \text{ nm}, z = 1, F_a = 40 \text{ V/nm}, \) Eqn. 2.54 reduces to

\[ \delta l_B = CT^{1/2}, \] (2.56)

where \( C = 18.6 \text{ pm/K}^{1/2}. \) This generates values of \( \delta l_B \) lying between 166 pm at 80 K, 322 pm at 300 K, and 415 pm at 500 K. These are values of the radius of a disc of uncertainty, defined by requiring that the lateral error in bonding-site location be less than this distance for about 50 % of ions. In general terms, uncertainties of this size are neither negligible nor serious, when compared with other reconstruction uncertainties [153], but longer-term recognition of thermal blurring may improve the reconstruction process and/or understanding of its limitations.

For FIM, the issue of whether separate centers of intensity can be detected in an image is strongly influenced by the signal transfer properties of the detector and the availability of contrast-enhancing software. Nevertheless, both the radius \( \delta l_B \) of the optical blurring disc and the trends in Eqn. 2.54 are of interest.
A prerequisite for “resolution” of separate emission sites (and, where relevant, their identification as “atoms”) is that there be sufficient local variation in the emission current density, as discussed in Sect. 2.4.2. The requirement then is that this potential image contrast must not be destroyed by thermal blurring. Some criterion is needed: for example, one might require that $\delta l_B$ be less than some specified fraction of the separation of the atomic lattice sites that one hopes to resolve.

For an emitter of apex radius 60 nm, and at the He BIF 45 V/nm, the value of $C$ in Eqn. 2.56 is 21.4 pm/K$^{1/2}$. This yields $\delta l_B$-values of 192 pm at 80 K and 263 pm at 150 K. The gas temperature at BIF has never been definitively decided, but is thought to lie between these temperatures [3, 156]. Optical blurring discs of these radii are compatible with the experimental facts of FIM resolving ability.

The trends exhibited by Eqn. 2.54 are of interest, because it indicates that (other things being equal) blurring is reduced by reducing tip apex radius, reducing gas temperature, or increasing the apex field. The discovery in 1955 [113, 157] that the FIM could be made to resolve atoms was attributed at the time to emitter cooling by refrigerants. The view of Melmed [158] is that this discovery was strongly due to the use of an emitter of lower apex radius by Bahadur, a graduate student in the laboratory at that time, who carried out the relevant experiment. Prior to discovery of FIM atomic imaging, it was their standard laboratory practice to clean emitters by thermal annealing: this blunts the relatively sharp emitter prepared by electro-chemical etching. Contrary to this standard practice, Bahadur used a sharp emitter cleaned by field evaporation; when this was combined with emitter cooling, FIM imaging of atoms was discovered.

FIM and APT literature contain formulae that claim to predict the minimum object size that can be imaged in an FIM, e.g., Eqn. 2.83 in [34]. These formulae appear to contain multiple errors and are compatible neither with CP optics nor with quantum mechanics, nor with modern signal processing technology; it is considered that they should be discarded.

2.10.6 Aberrations

Obviously, real emitters do not have the ideal spherical shape of an SCPE. Rather, the field evaporation endform of a defect-free, single-crystal, elemental emitters has a flattened apex and a tendency to exhibit facets. The result is that different parts or regions of the endform have different local radii of curvatures [159, 160]. Since principal trajectories depart normal to the emitter surface, these radii differences cause angular differences in optical behavior as between different regions. In turn, these lead to differences in local projection magnification $\mu$, as defined by

$$\mu = \frac{\delta R_S}{\delta l_S}, \quad (2.57)$$

where $\delta l_S$ is a small length in the surface of the emitter, and $\delta R_S$ is the corresponding length in the arrival map. These magnification variations can be thought of as aberrations in the optics of the SCPE.
Aberrations can also be caused by other forms of irregularity in the overall emitter shape, in particular by precipitate particles protruding from the surface of multiphase material, and by voids in the surface [34].

There is qualitative understanding of many of the CP optical effects seen in field ion images [34]. However, with the exception of some old work by Rose [161], of uncertain applicability, there have been relatively few attempts to put these effects onto a satisfactory quantitative basis. In general, problems of this kind are analytically intractable, and detailed numerical analyses have been very few until recently [40, 41, 144, 162, 163].

Systematic aberrations also occur on a smaller scale, when the emission from net-plane edges and from kink sites has a principal trajectory that is not normal to the smoothed surface that describes the overall emitter shape. This occurs to some extent with gas field ionization, but is a much stronger effect for field evaporation (and for field desorption in general). A further feature of field evaporation is that, because the emitting surface changes continuously, the features seen in FEV arrival maps are the result of integrating the optical effects associated with many different emitter shapes that differ in fine detail.

For the future, the greater use of numerical trajectory modeling, for more realistic emitter shapes, is expected to increase our understanding of the detailed CP optics of real field emitters and thereby enable improvement to APT reconstruction algorithms.

References

12. A. Sommerfeld, Naturwissenschaften 41, 825 (1927)
15. C. Herring, M.H. Nichols, Rev. Mod. Phys. 21, 185 (1949)
17. E. Wigner, J. Bardeen, Phys. Rev. 48, 84 (1935)
20. W. Schottky, Phys. Z. 15, 872 (1914)
26. R. Smoluchowski, Phys. Rev. 60, 661 (1941)
36. R.N. Hall, J. Appl. Phys. 20, 925 (1949)
44. M. Drechsler, Z. Elektrochem. 61, 48 (1957)
77. C. Herring, Phys. Rev. **82**, 87 (1951)
References

Atom-Probe Tomography
The Local Electrode Atom Probe
Miller, M.K.; Forbes, R.G.
2014, XVIII, 423 p. 182 illus., 62 illus. in color., Hardcover
ISBN: 978-1-4899-7429-7