Microscopy and interferometry with an atomic-size electron source

M. P. Silverman

An ultrasharp tungsten point with an emission region of subnanometre size produces a bright, self-focused, quasi-monochromatic, highly coherent electron beam — indeed the brightest particle source known to science and arguably the closest particle analogue to the laser. Used in an electron projection microscope, the 'nanotip' source provides high magnification and resolution without the use of lenses — in fact, a realization of the low-energy electron holography first envisioned by Gabor. Atomic-sized electron sources may make possible the investigation of new types of quantum processes involving the interference of correlated particles.

One of the most important practical legacies of de Broglie's wave–particle duality, the electron microscope is a sensitive probe of the structure of materials and the fundamentals of quantum mechanics. From the relatively crude instrument developed by Ernst Ruska and Max Knoll in the 1930s, electron microscopy has since evolved into a bewildering (for the uninitiated) variety of instrumental techniques characterized by an 'alphabet soup' of acronyms (such as FEM, LEEM, LEED, STM, STEM, SPM, SP, etc.) that designate specific features of emission, scanning, diffraction, tunnelling, transmission, probing and energy, among other things. At root level, however, an electron microscope—indeed any microscope—performs two basic, although not unrelated, functions: imaging and interferometry.

With regard to imaging, an oft-stated advantage of an electron microscope over a corresponding light microscope is that the electron wavelength — which can be one-thousandth of a nanometre for a high-voltage (~100 kV) microscope — is many times smaller than that of visible light. Thus, the electron microscope has a much smaller far-field diffraction limit and is capable of much higher resolution than its light optical counterpart. Nevertheless, the imaging of an electron beam is accomplished with electromagnetic lenses, which, like optical lenses, give rise to unavoidable aberrations. For example, the heart of a state-of-the-art field-emission microscope (FEM) is the objective lens, a thick magnetic lens into which the sample is inserted — with the attendant inconvenience that chromatic and spherical aberrations are minimized at different specimen locations.

By contrast, a scanning electron microscope (SEM), where the probe tip is positioned sufficiently close to a sample surface so that quantum tunnelling may ensue, operates in the near-field domain and can achieve lensless imaging with atomic-scale resolution. However, a SEM can probe only a minute portion of the sample at a time, and it is not always clear how the structure in the resulting composite image relates to the actual structure of the specimen. Moreover, there are many applications, concerning both imaging and interferometry where it is advantageous, or even necessary, to have a macroscopically extended beam of particles.

To be useful for interferometry an electron source must produce a coherent beam — in effect, electron 'waves' that give rise to stationary interference fringes upon linear superposition. Until recently, the high-voltage FEM contained the most coherent electron source available, the only source suitable for making holographic images with massive particles. But high-energy electrons can damage fragile specimens. Indeed, damage caused by this electron impact has been estimated to be comparable to that arising in close proximity to a nuclear explosion!

Clearly a bright, collimated, coherent, quasi-monochromatic, moderate-energy source of electrons would find innumerable applications in basic and applied physics. One has only to think of the stunning advances in modern optics wrought by the laser.

In this paper are discussed some of the features and potential applications of a remarkable electron source, designated a nanotip for its nanometre-scale emission region, or a teton tip for its nipple-like shape. Produced by field emission from an atomic-size tip of a thin tungsten wire, the resulting electron beam is considered the brightest particle source currently known to science — orders of magnitude brighter than a synchrotron or any extant laser. Brightness, as will be discussed later, is related to virtually all the principal attributes of a beam — intensity, divergence, coherence and degeneracy — and is perhaps the single most significant

M. P. Silverman is in the Department of Physics and Astronomy, Trinity College, 300 Summit Street, Hartford, CT 06106, USA

CURRENT SCIENCE, VOL. 68, NO. 10, 23 MAY 1995
Figure 1. Electron field emission at a metal–vacuum surface. For \( z < 0 \), the effective potential energy inside the metal is \(-V_0\). For \( z > 0 \) the potential energy \( V \) is the sum of a contribution \( V'_1 \) from the external electric field and \( V'_2 \) from the image charge. The potential barrier of \( V \) is lower by \( \Delta \phi = (e^2 F)^{1/2} \), where \( F \) is the applied field, than the work function \( \phi \). The Fermi energy, \( E_F = -\phi \), is the energy of the highest occupied level.

Figure 2. Schematic diagram of a nanotip field emission source. The rearward extension of rays from the sample to the optic axis defines a virtual source of effective radius \( r \), of subnanometre size.

Figure 3. Geometric configuration of a projection microscope with inchworm positioner (A), nanotip (B), specimen stage (C), and microchannel plate detector (D). \( Z_1 \) and \( Z_2 \) are, respectively, the source-sample and sample-image distances.

determinant of the overall performance of microscopes and interferometers.

Employed in the configuration of a projection microscope, the electron nanotip permits low-energy, lensless, holographic imaging of fragile objects. In addition, the nanotip source may well make it possible to realize in the laboratory new types of quantum-interference experiments involving correlated electrons.

Characteristics of nanotip emission

The field emission of electrons from a cold metal cathode, first observed by Lilienfeld in 1922 and interpreted by Fowler and Nordheim in 1928, was one of the first confirmed examples of quantum-mechanical tunnelling. Figure 1 shows a schematic diagram of this process, whereby application of an extraction potential of a few kilovolts, in conjunction with the image potential of the tunnelling electron, results in an effective decrease in the metal work function. The larger the applied field, the lower is the energy required for electrons to tunnel into the vacuum. Applying Fermi–Dirac statistics to this simple electrostatic model, Fowler and Nordheim (FN) derived an equation relating the total emission current \( I \) to the extraction potential \( V \). A key feature of the FN theory of field emission is that a plot of \( \ln(I/V^2) \) against \( I/V \) is linear for a given tip geometry.
The tip of a standard field emission gun consists of a fine tungsten point, about 100 nm in diameter, welded to a tungsten hairpin filament. Although of microscopic size, these points are still large compared with the scale of an atom. Nanotip field emitters, by contrast, comprise an emission region at least ten times smaller—a nipple-like protrusion resting upon a much larger base, as schematically represented in Figure 2. The protrusion is actually a surface asperity terminating in one or a few tungsten atoms. The effective source size, deduced experimentally from measurement of the lateral coherence width and application of the van Cittert—Zernike theorem, has been shown to be a fraction of a nanometre, e.g. 0.73 nm in a recent experiment\(^\text{10}\) with a nanotip operating at 90 V. (One of the most important results in modern optics, the van Cittert-Zernike theorem\(^\text{11}\), relates by a Fourier transformation the coherence of waves at an observation plane to the intensity distribution across the source.) The atomic-scale size of these new electron sources gives rise to a number of important consequences.

First, the size of a radiation source ultimately determines the resolution of the (far-field) images it produces. Thus, an electron microscope employing atomically sharp tips should, in principle, be able to attain atomic-scale resolution of the surface features of material samples.

Second, as a consequence of the very small radius of curvature of the tip, a low extraction potential—less than 100 V as illustrated above, rather than several kV—creates an electric field sufficiently high (~10\(^7\) V/cm) in the vicinity of the tip for electron tunnelling to occur. Low-energy electrons not only can lead to reduced damage to fragile specimens, but are experimentally convenient to produce. A low-voltage projection microscope (Figure 3) is in essence a relatively inexpensive ‘desktop’ instrument, compared with a high-voltage state-of-the-art FEM with its long (>1 m) acceleration column and large high-voltage generators.

Third, the beam from a nanotip is highly coherent, as one might have expected for waves from an approximate ‘point’ source. Early theoretical modelling\(^\text{12}\) of electron emission from ultrasharp tips indicated a transverse coherence width of the order of centimetres at the viewing screen (15 cm distant from the source), a result in accord with experiment\(^\text{10}\). This degree of coherence represents a major advance in the direction of low-energy electron holography and interferometry.

Fourth, recent experimental work seems to show significant differences between the properties of beams derived from tips of atomic size and the broader standard field emission tips. For example, beams from nanotips exhibit much higher intrinsic brightness, strong quantum-mechanical self-focusing effects\(^\text{11}\), current-voltage characteristics different from those described by the FN equation, and multiply peaked emission electron energy spectra\(^\text{14}\). Such departures from standard field emission theory are perhaps to be expected, for the electron wavelength inside the tip is comparable in size to the source. Strong quantum effects of analogous nature have also been observed in the conductance of electrons through narrow constrictions in mesoscopic quantum systems\(^\text{15}\). Thus, besides facilitating potentially new ways to probe matter, the nanotip is interesting in its own right for the potentially unusual quantum processes involved in electron emission.

The high brightness of a nanotip is an especially virtue for both microscopy and interferometry. Although the concept can be more rigorously delineated\(^\text{16}\), for the present discussion it is sufficient to define the mean axial brightness \(B\) of a beam as the current density \(j\) emitted into a (differentially small) solid angle about the forward direction. Divided by the relativistically corrected acceleration potential, the resulting ‘reduced brightness’ is an invariant along the axis of a cylindrically symmetric electron optical system. Thus, for a nonrelativistic electron source, \(B\) effectively scales with applied potential. From considerations of the volume of phase space occupied by the beam, with account taken of the Pauli principle, one can deduce\(^\text{18}\) the maximum theoretical brightness \(B_{\text{max}}\) of electrons with charge \(e\), mass \(m\), wavelength \(\lambda\), energy \(E\) and energy dispersion \(\Delta E\).

\[
B_{\text{max}} = \frac{2e\Delta E\hbar \lambda^2}{h^2 \sqrt{2me}} \rightarrow 4me\Delta E\hbar \frac{1}{\lambda^2},
\]

where \(\lambda\) and the acceleration potential \(V\) are related as follows:

\[
\lambda = \sqrt{\frac{h}{2me\left[V + \left(\frac{e}{2mc^2}\right)^2\right]}} \rightarrow \frac{1226 \text{ nm}}{\sqrt{V} \text{ (volt)}}.
\]

The first equality in these equations is relativistically exact; the second is a nonrelativistic reduction in which \(E\) is specifically the kinetic energy \(p^2/2m = h^2/2m\lambda^2\), with \(\lambda\) in nm and \(V\) in volts in (1b).

Although brightness is one of the most critical parameters determining the performance of an electron optical instrument, it is also one of the most difficult to measure. In the standard (or minimum probe size) method, one can try to assess directly the experimental quantities (total current, focused probe cross-section, and beam divergence) that enter into the definition. Alternatively, one can deduce brightness from the lateral coherence of the beam. The two methods are not equivalent if the beam is coherent, for in that case the true size of the source cannot be inferred by direct measurement of the focused beam intensity since the Gaussian image of the source is hidden in the coherent pattern of the condenser lens aperture. The results of the standard method would then give an overestimate of the source size and a lower limit to the beam brightness.
GENERAL ARTICLES

Recent experiments\textsuperscript{17} employing the standard method in a lensless projection configuration, like that of Figure 3, fitted with an ultrasharp tungsten single-crystal tip terminating in a single atom (as verified by field ion microscopy), yielded an axial brightness of $3.3 \times 10^8 \text{A cm}^{-2}\text{sr}$ at an energy of 470 eV. Scaled to the energy of 100 keV, the equivalent brightness would have been $7.7 \times 10^{10} \text{A/cm}^2\text{sr}$, a figure two orders of magnitude greater than that of a state-of-the-art FEM. For purposes of comparison, it is useful to consider the actual particle flux per unit of energy. A nanotip with the cited brightness and a typical energy dispersion of 0.2 eV yields approximately $2.4 \times 10^{10}$ electrons/cm$^2$-ss-\text{eV}, which is at least four orders of magnitude greater than a synchrotron source and a billion times the reduced brightness of the Sun\textsuperscript{18}.

The ratio of the actual brightness to the theoretical maximum in eq. (1)

\begin{equation}
\delta = \frac{B}{B_{\text{max}}} \tag{2a}
\end{equation}

yields the degeneracy parameter $\delta$ of the beam. Interpretable as the mean number of particles per cell of phase space, $\delta$ plays a seminal role in prospective new kinds of quantum physics experiments involving multi-particle interference\textsuperscript{19-21}. In an alternative, but essentially equivalent, formulation, the 'corpuscular' degeneracy parameter can be related to the 'undulatory' coherence parameters as follows:

\begin{equation}
\delta = \frac{(j/e)A_cT_e}{\lambda} \tag{2b}
\end{equation}

where $A_c$ is the coherence area and $T_e$ the coherence time. To within a numerical factor of order unity, $A_c$ is essentially the square of the transverse (or lateral) coherence length $\ell$, which characterizes the extent (in a plane normal to the beam) over which superposing waves manifest interference fringes. $T_e$, which is effectively Planck's constant divided by the dispersion $\Delta E$, characterizes the extension of a wavepacket along the beam, i.e. the longitudinal coherence length $L_e = vT_e$, where $v$ is the mean electron speed. According to eq. (2b) the degeneracy of a beam is the mean number of particles passing through a coherence area during the period of a coherence time.

For fermions the Pauli principle restricts the degeneracy parameter to a maximum value of unity. The degeneracy of a high-voltage FEM beam is typically of the order of $10^{-6}$, which makes multiparticle quantum interference experiments marginally possible\textsuperscript{22} if at all. For a nanotip source, however, $\delta$ is expected to be at least two orders of magnitude greater. While still low compared to optical sources— which emit massless bosons whose degeneracy is unrestricted by the Pauli principle—this is nonetheless a significant improvement.

Imaging and Interferometry

Conceptually, as well as instrumentally, the point projection apparatus with field emission nanotip, illustrated in Figure 3, is perhaps the world's simplest electron microscope and interferometer\textsuperscript{23}. By means of a piezoelectric (inchworm) positioner, the tip can be placed at a microscopic distance $Z_1$ from the surface of a thin specimen. Electrons drawn off by a low extraction potential propagate through the specimen to a microchannel viewing screen at a macroscopic distance $Z_2$ from the sample. If the sample is narrower than the beam—for example, a mask-like object comprising opaque regions perforated by holes—this simple configuration generates without the use of lenses a highly magnified projective or shadow image. The magnification $M$ is given geometrically by the relation

\begin{equation}
M = \frac{(Z_1 + Z_2)/Z_1}{\frac{Z_2}{Z_1} \Rightarrow Z_2 = Z_1/M,} \tag{3}
\end{equation}

which for $Z_1 \sim 100 \text{nm}$ and $Z_2 \sim 10 \text{cm}$ leads to $M \sim 10^6$.

There is a close connection between the standard configuration of a transmission electron microscope (TEM), in which the specimen is illuminated by plane waves from a distant source, and the configuration of a point projection microscope (PPM), in which the specimen is illuminated by spherical waves from a proximate source. This connection is illustrated in Figure 4. If the sample is sufficiently thin so that multiply scattered electrons do not contribute to the image, then it can be shown\textsuperscript{10} that the resulting PPM image is equivalent to a TEM image "defocused" by the tip-to-sample distance $Z_1$. An example of a highly magnified image of a perforated carbon foil is shown in Figure 5. Because of the high coherence, the image reveals interference fringes resulting from Fresnel edge diffraction. Although the point projection field emission microscope was first demonstrated\textsuperscript{24} in the 1930s, it is only recently, with the use of a nanotip source, that the instrument has the necessary resolving power and coherence to produce Fresnel fringes\textsuperscript{25}.

Slightly defocused to produce Fresnel fringes, the microscope serves as an interferometer providing detailed information about the source of illumination and the coherence properties of the beam. For example, by modelling the electron transmission through perforated carbon film as a problem in scalar diffraction, one can predict the lateral positions of the fringes in terms of $\lambda$, $Z_1$ and $Z_2$. Since $Z_2$ is a directly measurable macroscopic length, comparison of theory with experiment leads to the nanoscopically small and initially unknown tip-to-specimen separation $Z_1$ and, therefore, to the magnification $M \sim Z_2/Z_1$. In the case of Figure 5, with $\lambda = 0.13 \text{nm}$ and $Z_2 = 14 \text{cm}$, one obtains $Z_1 = 461 \text{nm}$ and $M = 304,000$. 

1006

CURRENT SCIENCE, VOL. 68, NO 10, 25 MAY 1995
Figure 4. Comparison of (a) PPM and (b) TEM configurations. In (a) the tip S generates a (vertically descending) spherical wave. Since the field emitter itself forms a small lens, the virtual source within it may be considered to be formed by a preceding virtual lens (dashed lines). In (b) the source S' generates a (vertically ascending) plane wave. The respective images at A and B are identical if the specimen is weakly scattering.

From the pattern of fringes one can also determine the effective source size, which, as a measure of the attainable instrumental resolution, is particularly important. Since interference occurs only if the superposing beams are coherent, the lateral extent of the set of Fresnel fringes corresponds to the transverse coherence length \( L_t \). Application of the van Cittert-Zernike theorem \(^\text{11}\) then leads to the effective radius of emission (of a cylindrically symmetric source)

\[
r_e = \frac{\lambda}{nL_t/Z_1}.
\]

(4)

Measurements taken from Figure 5 yield \( L_t = 26 \text{ nm} \) (in the sample plane) and \( r_e \approx 0.73 \text{ nm} \).

Knowledge of \( L_t \) provides the coherence area \( A_c \), which, together with the current density \( j \) at the viewing screen, yields the beam brightness

\[
B = 2jA_c/\lambda^2 = 2nj(ML_t/\lambda)^2
\]

(5)

(deducible from eqs (2a) and (2b)). For the experiment leading to Figure 5, \( j = 20 \text{ pA/cm}^2 \) at the image plane and \( B = 4.6 \times 10^5 \text{ A/cm}^2\text{-sr} \), which is equivalent to a brightness of approximately \( 5 \times 10^8 \text{ A/cm}^2\text{-sr} \) for a 100 kV beam. Brightness depends sensitively on the quality of tip preparation. In the present experiment, the nanotip was simply formed electrochemically with a KOH solution, then cleaned briefly in vacuum at room temperature by field evaporation. The tip was not subjected to ion sputtering or observed by field ion microscopy to ensure that it ended in a single tungsten atom.

For the purpose of high-resolution imaging, rather than interferometry, it is desirable to 'focus' the PPM to remove the Fresnel fringes. How can this be done if, as recounted above, perfect focus corresponds to \( Z_1 = 0 \) and, therefore, \( M = \infty \)? This adjustment cannot be made physically, for it requires direct contact between the source and the specimen (in which case field emission cannot occur). In this regard it is important to note that the PPM configuration shown in Figure 3 corresponds exactly to the geometry for in-line holography envisaged by Gabor\(^\text{26}\). Since PPM imaging is an example of in-line holography, one can effect holographic reconstructions by computer – in effect, to bring the image 'back into focus' – by treating the image as if it were produced by plane wave illumination from a distant source (as illustrated in Figure 4). Recent efforts along these lines have proven quite successful\(^\text{27}\). The 'twin-image' problem attendant to in-line holography can be effectively circumvented if, at exact focusing of the real image, the virtual twin is sufficiently defocused so that it largely constitutes a weak and undifferentiated background.

**New directions in quantum interferometry**

In recent years the field emission electron microscope has played a seminal role, not only in probing the
structure of materials, but also in fundamental investigations of quantum mechanics. Experiments have been performed that show, for example, the build-up of a two-slit electron interference pattern one electron at a time, or that confirm the counterintuitive and controversial prediction of Aharonov and Bohm (AB) that static magnetic and electric fields can have nonlocal influences on charged particles. For want of space, detailed accounts of these fascinating investigations must be sought elsewhere.

A feature common to virtually all quantum-interference experiments performed with massive particle beams – whether electron, neutron, or any other – is that they involve the ‘self-interference’ (à la Dirac) of single-particle states. This means that the interference patterns could be produced over a period of time with only one particle at a time passing through the interferometer. Despite the diversity of experimental motivations and techniques, these experiments are all characterized by the first-order correlation function of the particle field, i.e. the functional equivalent of the degree of first-order coherence in optics.

One can conceive, however, of a ‘quantum optics of particles’ that goes beyond simple two-slit interference to consider higher order interference phenomena of correlated massive fermions and bosons. The second-order (in field amplitude) coherence, for example, characterizes the correlated fluctuations of particle intensities, as demonstrated in the domain of visible optics by the experiments of Hanbury Brown and Twiss (HBT) that can only be briefly mentioned here.

There are experiments, for example, that combine features of both the AB and HBT effects, whereby the correlated fluctuations at two detectors, and not merely the single-particle probability of arrival at one detector, are influenced by an excluded region of confined magnetic flux.

Another set of experiments represents a hybrid of the AB effect and the Einstein–Podolsky–Rosen (EPR) paradox, whereby two electrons whose quantum states are ‘entangled’ (to use the word employed by Schrödinger) each diffract around a distinct and well-separated region of confined magnetic flux. The joint detection probability can be shown to depend on the difference in magnetic flux of the two regions; were the flux in one region to be null, then to a ‘local’ observer near that region the probability of detecting an electron would seem to depend on the flux of the other arbitrarily distant region of magnetic flux through which and around which the detected electron did not pass! (A double dose of quantum-mechanical nonlocality!)

The cited examples manifest one or more of the three distinct kinds of quantum interference: (i) interference resulting from wave-like propagation of particles and dependent upon optical path length difference; (ii) interference resulting from particle charge and spatial topology and dependent upon confined magnetic (or electric) flux; and (iii) interference resulting from particle indistinguishability under exchange and dependent upon quantum statistics. What distinguishes these processes from the by now ‘familiar’ brand of quantum interference is that they are all intrinsically multiparticle processes, the interesting consequences of which are often lost or unobserved if one is counting single-particle arrivals at a single detector.

At first glance, the existence of such processes may seem disturbing, for how are they to be reconciled with Dirac’s fundamental dictum – expressed originally for photons, but equally applicable to all particles – that ‘each (particle in a split beam) interferes only with itself. Interference between different (particles) never occurs’. The new kinds of particle interferometry do not necessarily refute Dirac, but rather force the physicists to reinterpret the meaning of ‘different’ particles.

To do experiments that involve multiparticle interference, one must have a source for which there is a reasonable probability of obtaining correlated particles. Heuristically speaking, the wavefunctions of the emitted particles should in some sense overlap at the time of production. It is, therefore, no surprise that the signal-to-noise ratio in a particle correlation experiment increases with the beam degeneracy parameter δ – hence the great interest in the high brightness of an ultrasharp field emission source.

As a first step in the direction of higher-order electron interferometry, experiments are underway to probe the statistical properties of electron emission from a nanotip, in essence to perform a HBT experiment with electrons. This can be implemented in at least three ways by measuring (i) the variance in electron counts about the mean count at a single detector, (ii) the conditional probability of electron arrival at a single detector as a function of the time interval between two consecutive detections, and (iii) the correlated count rate of electrons at two detectors as a function of time delay. Theoretical modelling of a field emission source as a chaotic source, i.e. with density operator diagonal in energy–momentum–spin states, predicts that the emission statistics should exhibit antibunching, a kind of particle avoidance behaviour. One stark manifestation of antibunching is that the conditional probability of receiving two electrons with the same spin component simultaneously at the same detector is rigorously zero. Given the short coherence times of even the most monochromatic electron sources, Tc = 10−14 s for ΔE ~ 0.2 eV, it is presently very difficult to probe electron emission directly for time intervals less than Tc. Nevertheless, small departures from Poissonian statistics should be demonstrable even for time intervals longer than Tc.

At this point a definitive quantum description of an atomic-size electron emitter has not been established.
Interestingly, despite the wide belief that Fermi–Dirac statistics necessarily implies electron antibunching, it has been demonstrated recently that electrons, like photons, can under appropriate circumstances exhibit a variety of statistical effects depending upon the specific ensemble composition\(^5\). For example, correlated electrons entering both input ports of an interferometer can produce quantum ensembles with statistical properties vastly different from those of a chaotic ensemble, despite the fact that all state vectors are antisymmetrized in accordance with the Pauli principle. One of these ensembles, in fact, would—if realizable in the laboratory—lead to the highest phase sensitivity of any electron interferometer. It is, therefore, of particular interest to examine electron emission from a nanotip with two (or more) atomic-scale asperities located within a lateral coherence length. Should such a system exhibit correlated electron emission, one would indeed have an extraordinary source for electron interferometry.

Of this, though, one can be certain: the brightest, most coherent particle source in science will assuredly find many uses and stimulate new ideas.


ACKNOWLEDGEMENT. I thank Prof. S. Ramaseshan for the invitation to submit this article.