© *Journal of Microscopy, Vol.* 98, *Pt* 3, *August* 1973, *pp.* 417–435.
Received 3 *January* 1973; *revision received* 5 *April* 1973

Resolution limits in the surface scanning electron microscope

by C. J. D. CATTO *and* K. C. A. SMITH, *Cambridge University Engineering Department, Trumpington St, Cambridge*

SUMMARY

An approximate theory is developed for estimating the limiting topographical resolution of the scanning electron microscope operating under certain idealized conditions. Limitations imposed by electron beam shot noise and electron diffusion effects within the specimen are considered for the case of an instrument incorporating a field emission source and in which there is ideal collection of the secondary electron signal. The specimen is assumed to be an homogeneous, isotropic solid with the beam incident normally to its surface.

It is estimated that, under these ideal conditions, the limiting resolution for a metallic specimen lies in the region of 1 nm. The possibilities of realizing a resolution of this order in a practical instrument are discussed.

INTRODUCTION

Until very recently the performance of the surface scanning electron microscope operating in the secondary emission mode has been determined fundamentally by the limited brightness obtainable from the commonly used tungsten hairpin cathode. This has imposed a resolution limit of the order of 10 nm, being determined by the smallest probe diameter obtainable consistent with reasonably noise-free pictures and moderate picture recording times (Smith, 1956, 1972). This order of resolution has also coincided with the order of the limit which has been predicted on the basis of diffusion of the secondary electrons in the specimen (Wells, 1957). Various workers have, however, reported resolutions better than 10 nm under certain operating conditions (Pease & Nixon, 1965) and recently experiments with instruments employing either a lanthanum hexaboride thermionic cathode (Broers, 1969) or a tungsten field emission cathode (Komoda & Saito, 1972) have demonstrated resolutions substantially below this figure.

In one type of scanning transmission electron microscope utilizing a field emission source, a probe diameter of less than 0·5 nm has been achieved (Crewe & Wall, 1970), but this order of performance requires that the specimen be immersed deep in the magnetic field of the objective lens in order to minimize electron-optical aberrations. Considerable practical difficulties lie in the way of achieving a comparable performance in the surface scanning microscope mainly because of the difficulties of achieving efficient collection of the secondary electrons from a specimen immersed inside a magnetic objective lens. Several commercial

electron microscopes now exist with which it is possible to form the surface scanning secondary electron image from a specimen situated within the normal transmission objective, but these instruments so far have not demonstrated probe diameters much below 3–5 nm. However, it is reasonable to assume that with the further development of this type of instrument, coupled with the application of the field emission source, the resolving power of the surface scanning microscope lies potentially below 1 nm. It is therefore pertinent to enquire whether the limits imposed by beam shot noise and diffusion effects within the specimen allow the possibility of a resolution comparable with the resolving power now potentially available.

In the present paper an approximate theory is developed which allows an estimate to be made of these limits for the case of an instrument incorporating a field emission source and operating under certain idealized conditions. It is assumed that the specimen is an homogeneous, isotropic solid with the scanning beam incident normally upon one flat surface. It is also assumed that all secondary electrons emitted are detected by an ideal noise-free detector. Contrast is assumed to arise solely from the differential emission of secondary electrons arising through geometrical changes in surface topography the dimensions of which are of the same order as the scanning probe. Large-scale contrast effects such as those caused by penetration and breakthrough by the beam at sharply contoured features in the specimen, and contrast effects such as those arising because of channelling etc., are assumed to be negligible.

Further details of the theory presented in this paper may be found in Catto (1971), and a general discussion of contrast and resolution in the SEM is contained in Everhart (1970).

OUTLINE OF THE SIGNAL PRODUCTION PROCESS

Three major components of the secondary electron current constituting the signal in the SEM may be distinguished. First, there is the component which arises when the primary beam is incident on a perfectly smooth region of the surface (Fig. 1a). As the beam penetrates beneath the surface, secondary electrons are released, some of which diffuse to the surface and escape. The ratio of the resulting secondary current to the current in the primary beam may be defined as the secondary emission ratio δ_f for the flat surface. Secondly, when the beam is incident on a small asperity (Fig. 1b) secondary electrons generated within the asperity may escape to give rise to an additional secondary current associated with the presence of the asperity. However, some of the secondaries generated in the region below the asperity and which would, in its absence, contribute to δ_f are now absorbed within the asperity; thus the incremental change in the secondary current signal arises as a difference between the current produced solely within the asperity and the current absorbed by it. It is convenient for the purposes of the theory which follows to define a secondary emission coefficient δ_p describing the fraction, relative to the primary beam current, of the secondary current generated solely within the asperity and which escapes.

Both of the secondary emission coefficients defined above arise as a result of secondaries generated within the first 10 nm or so of the surface on entry of the primary beam; secondaries generated much deeper than this are absorbed before reaching the surface. A proportion η of the primary beam, however, is backscattered and emerges from the specimen with appreciable energy (Fig. 1c). As the backscattered primaries pass again through the surface they release secondaries which form a third secondary current component the magnitude of which is

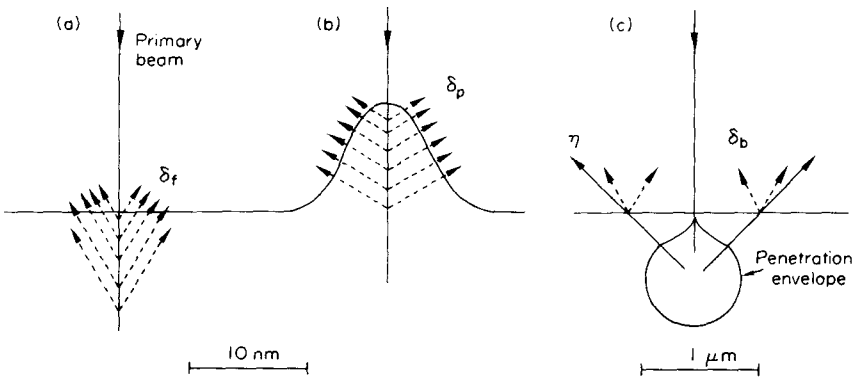

Fig. 1. Illustrating schematically the production of the secondary electron signal in the SEM. Scale markers indicate a scale change of two orders of magnitude between (a), (b) and (c). (a) Beam entering flat surface: secondary emission coefficient δ_f. (b) Beam entering asperity: secondary emission coefficient δ_p. (c) Beam backscattered: secondary emission coefficient δ_b, back scattering coefficient η.

described by δ_b, the ratio of the secondary current due to backscattered primaries to the current in the primary beam.

It should be noted that the scale of Fig. 1c is some two to three orders of magnitude greater than that of Fig. 1a and b. Whereas the secondaries contributing to δ_f and δ_p emerge over a region within about 10 nm of the primary beam, secondaries contributing to δ_b emerge over a diameter of the order of 1 μm. To a first order, therefore, small scale variations in surface topography do not effect any differential changes in the coefficient δ_b as the beam scans across the surface. Image contrast, for the small scale detail considered here, consequently occurs mainly because of differential changes in the coefficient δ_p and, to a lesser extent, changes in the effective magnitude of δ_f. These changes constitute the signal which must be detectable above the noise associated with all three secondary electron components.

In the above account of the signal production process it has been assumed that δ_p arises as a result of an infinitely thin primary beam. Clearly the effective value of the secondary emission coefficient for the asperity will fall with increase of beam diameter. The magnitude of this effect is considered in the sections which follow.

DEPTH OF ORIGIN AND DIFFUSION OF SECONDARY ELECTRONS

The production and escape mechanisms for secondary electrons have been described in general terms by Bruining (1954) and, in relation to the SEM, by Wells (1957) and Everhart (1958).

Consider the case of an infinitely thin primary beam penetrating the surface layers of a specimen and generating secondary electrons along its path. Following Bruining (1954), if the probability of escape of an electron produced at the surface is $P(o)$, then the probability of escape of an electron produced at depth x below the surface is

$$P(x) = P(o)e^{-x/L} \tag{1}$$

where L is the mean-free path.

If n_s is the secondary production per unit distance (considered constant), the

number of electrons reaching the surface per incident primary electron, originating from within a depth x, is given by

$$N(x) = P(o)\int_0^x n_s e^{-x/L}dx$$
$$= P(o)n_s L(1 - e^{-x/L}). \quad (2)$$

Thus, the total number of secondaries reaching the surface per incident primary is

$$N(\infty) = P(o)n_s L.$$

The probability of escape at the surface can always be made unity by the application of a small positive extracting field at the surface of the specimen. Hence we may write

$$\delta_f = N_s L. \quad (3)$$

It may be deduced from the above equations that 95% of the secondary current contributing to δ_f originates from within a depth $3L$. The mean-free path L has been measured by many workers and for a wide range of materials, but the results for any particular material differ considerably (see collected data for metals in Catto, 1971). In the following theory a value of $L = 1{\cdot}5$ nm has been taken for the purposes of numerical evaluation, this being the average of values given for copper.

If the energy of the primary incident beam is more than a kilovolt or so, the beam energy can be regarded as being substantially constant over the depth in which the secondaries contributing to δ_f are produced. The contribution from depths greater than 10 nm is negligible for most metals.

Also of importance in the estimation of contrast and resolution is a knowledge of the radial density distribution of the secondary electron current escaping at the surface. This has been considered by Wells (1957) who gives a curve of current density distribution about an infinitely thin primary beam. To a good approximation this curve is found to fit an expression of the form

$$\delta_{(r)} = \frac{\delta_f}{GL} e^{-r/GL}dr \quad (4)$$

in which $\delta_{(r)}$ is the current in an annulus of width dr at radius r, and G is a constant having a value very nearly 0·5. From this expression it is deduced that 95% of the total secondary current is contained within a radius of $r = 1{\cdot}5L$. An escape cone of half-angle about 30° may therefore be postulated (Fig. 2).

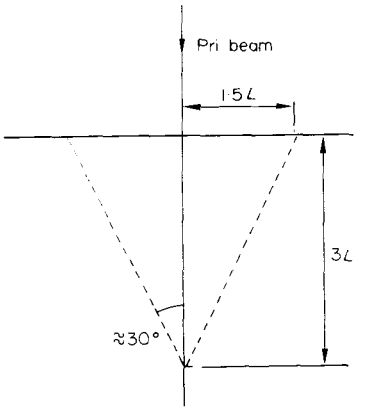

Fig. 2. Escape cone for secondary electrons.

The above theory concerns the production of secondaries occurring on entry of the primary beam which retains substantially its initial direction and energy over the region in which the secondaries contributing to δ_f are produced. The backscattered electrons on the other hand will emerge from the specimen at varying angles and with varying amounts of energy, although it is still permissible to regard the energy as constant over the production depth L. If the angle at which a backscattered primary emerges is θ, then, from Fig. 3, it can be seen that the magnitude of the secondary current escaping is enhanced by a factor sec θ. Secondary production by backscattered electrons is also enhanced, compared to production by incident primaries, because of their lower mean energy. Consequently, although only a small fraction of the primary beam is backscattered, the coefficient δ_b is generally comparable in magnitude with δ_f (Kanter, 1961) and must be taken into account in any assessment of contrast and resolution.

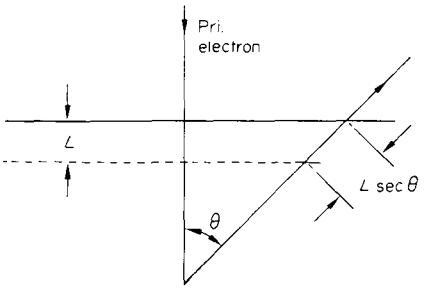

Fig. 3. Secondary production by backscattered primaries.

The sum of the coefficients δ_f and δ_b is equal to the total low-energy secondary emission coefficient δ for a flat surface, i.e.

$$\delta_f + \delta_b = \delta. \tag{5}$$

Experimental values of δ are given in the literature and some information is available concerning the coefficients δ_f and δ_p separately (Seiler, 1967). In the following section expressions for δ_f and δ_p are derived and their sum scaled according to equation (5) to fit experimentally determined values for copper.

SECONDARY EMISSION COEFFICIENTS FOR A FLAT SURFACE

The production of secondary electrons is related to the rate of energy loss of the primary and backscattered electrons. Assuming that the production is proportional to the rate of energy loss with distance (Bruining, 1954) equation (3) may be written

$$\delta_f = Lk\left(-\frac{dE}{dx}\right) \tag{6}$$

where k is the secondary production per unit energy loss.

The rate of energy loss for electrons of energy greater than about 10 keV may be derived from a modified Thomson-Widdington law (Cosslett & Thomas, 1964) of the form:

$$E_0^2 - E^2 = b'x \tag{7}$$

hence

$$-\frac{dE}{dx} = \frac{b'}{2E} \tag{8}$$

where E is the energy at depth x and b' depends on the initial energy E_0 and the density ρ. Data for copper relating to b' and initial energy, derived from Cosslett & Thomas (1964), have been used together with (8) to obtain Fig. 4. An approximate fit to the curve in this figure is given by

$$-\frac{dE}{dx} = (25 - 0{\cdot}1\, E_0^{1{\cdot}2}) eV\ \mathrm{nm}^{-1} \qquad (E_0 \geqslant 10\ \mathrm{keV}) \tag{9}$$

with E_0 in volts.

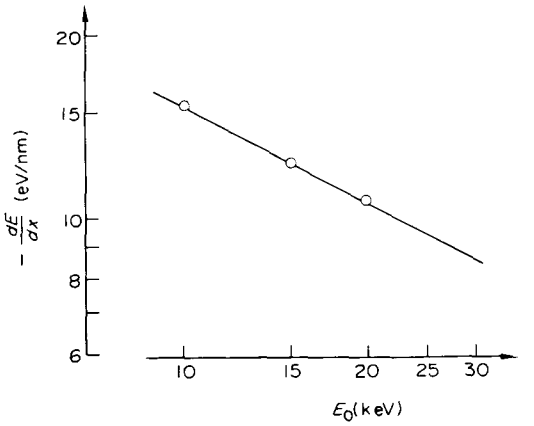

Fig. 4. Energy loss rate in copper. (Derived from data in Cosslett & Thomas, 1964).

It will be noted that in equation (9) the rate of energy loss is expressed as a function of the initial energy E_0 only, and not of the energy E. This is because it is assumed, implicitly, that the energy loss, in relation to E_0, is very small over the depth in which secondaries are produced.

At energies below 10 keV a better relation is the 'constant loss' law given by Lye & Dekker (1957):

$$-\frac{dE}{dx} = \frac{E_0}{R(E_0)}$$

where the range $R(E_0) = CE_0^{1{\cdot}35}$ (with C = constant).

This expression may be made consistent with equation (9), at an energy of 10 keV, by putting $C = 380$, which value will be used in subsequent expressions. Thus

$$-\frac{dE}{dx} = \frac{380}{E_0^{0{\cdot}35}} eV.\ \mathrm{nm}^{-1} \qquad (E_0 \leqslant 10\ \mathrm{keV}.) \tag{10}$$

with E_0 in volts.

Using equations (9) and (10), depending on the value of the primary beam energy, an expression for δ_f may be derived according to equation (6) in which the only unknown is k.

An expression similar to (6) is applicable for the number of secondaries released at the surface per backscattered primary, except for the sec θ factor discussed above. For complete evaluation of δ_b, therefore, it is necessary to know not only the number of backscattered primaries but also their distribution with energy and

angle of emergence. This information may be derived from energy-spread data given by Bishop (1966) and from the diffusion model for backscattering discussed by Archard (1961).

Figure 5, which is derived from Bishop's data, shows the radial energy distribution of the backscattered electrons for one value of the primary beam energy. This curve also may be fitted approximately by the exponential relationship:

$$E_r = E_p e^{-r/r_b} \tag{11}$$

where E_p is the primary beam energy, E_r is the energy at radius r, and r_b has the value 230 nm for $E_p = 30$ keV. At energies of 20 keV and 10 keV, estimated values for r_b are 140 nm and 50 nm respectively (Catto, 1971).

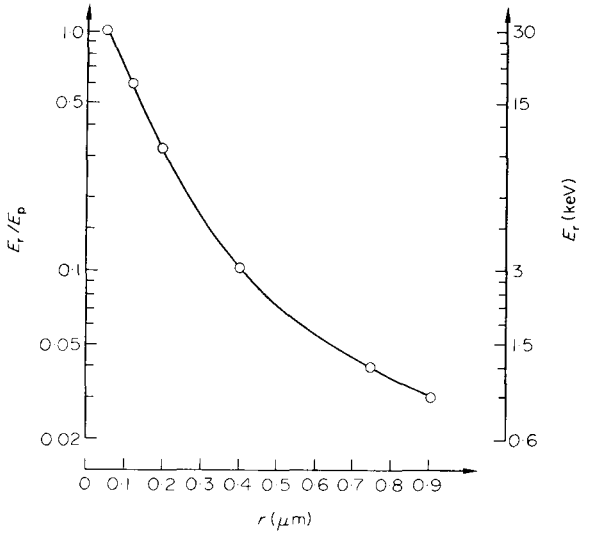

Fig. 5. Radial energy distribution for backscattered electrons in copper. Primary beam energy $E_p = 30$ KeV (after Bishop, 1966).

The diffusion model for backscattered primaries is illustrated in Fig. 6. The backscattered primaries are considered to diffuse uniformly from the centre O of a sphere of radius R_d at depth uR_d ($0 \geqslant u \geqslant 1$). The solid angle subtended at 0 by the intersection of the sphere with the surface is

$$\frac{1}{R_d^2}\int_0^{\theta_0} 2\pi R_d \sin\theta \; R_d \; d\theta = 2\pi(1 - \cos\theta_0).$$

If the total diffusion current is equal to the primary beam current, the backscattering coefficient η is given by:

$$\eta = \frac{2\pi(1 - \cos\theta_0)}{4\pi} = \tfrac{1}{2}(1 - u). \tag{12}$$

From this expression the parameter u may be derived using the experimental values for η given for copper by Bishop (1966) (see Table 1). Some of the other parameters defined in Fig. 6, and which are used later, are also tabulated.

Values of R_d are estimated from the calculations and equations given by Bishop (1966) and Lewis (1950).

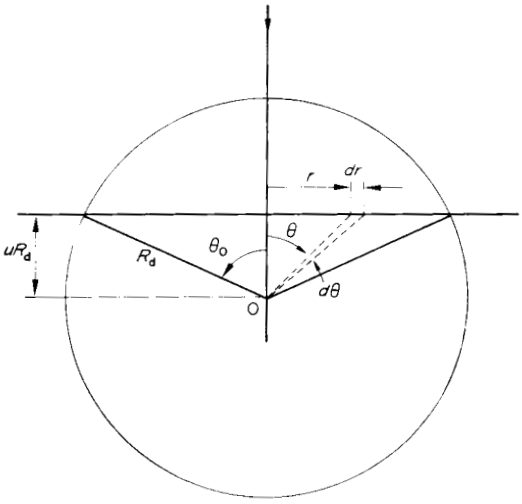

Fig. 6. Diffusion model for backscattered primary electrons.

Let the primary beam current have unit value and let $\mathcal{J}_{br}$ be the diffusion current of backscattered primaries through an annulus of width dr at radius r (Fig. 6) where $r = uR_d \tan\theta$. A solid angle calculation then gives

$$\mathcal{J}_{br} = \frac{\frac{1}{2}uR_d r dr}{[(uR_d)^2 + r^2]^{3/2}}.$$

The secondary current in the same annulus will be, using (6),

$$\begin{aligned}\mathcal{J}_{sr} &= Lk\left(-\frac{dE}{dx}\right)\mathcal{J}_{br}\sec\theta \\ &= Lk\left(-\frac{dE}{dx}\right)\frac{\frac{1}{2}rdr}{(uR_d)^2 + r^2}. \qquad (13)\end{aligned}$$

For a primary beam energy of 10 keV, equations (10) and (11) may be combined to give the rate of energy loss for backscattered primaries issuing from the annulus, i.e.

$$-\frac{dE}{dx} = \frac{380}{E_p^{0.35}} e^{-0.35r/r_b}.$$

Hence

$$\mathcal{J}_{sr} = Lk\frac{380}{E_p^{0.35}} e^{-0.35r/r_b}\frac{\frac{1}{2}rdr}{(uR_d)^2 + r^2}. \qquad (14)$$

Therefore

$$\delta_b^{(10)} = \frac{Lk380}{E_p^{0.35}}\int_0^{r_1}\frac{e^{-0.35r/r_b}rdr}{(uR_d)^2 + r^2} \qquad (\text{for } E_p = 10\text{ keV}). \qquad (15)$$

The limit r_1 in (15) may be taken as the value of the radius at which the mean energy E_1 of the backscattered primaries is so low that all the energy is absorbed within the length L, i.e. from (10) and (11).

$$\frac{E_1}{L} \approx \frac{380}{E_1^{0.35}}$$

and

$$E_1 = E_p e^{-r_1/r_b}$$

therefore

$$r_1 \approx r_b ln(E_p/100) \qquad \text{(for } L = 1{\cdot}5 \text{ nm)}. \tag{16}$$

In a similar fashion, for a primary beam energy of 20 keV both energy loss equations (9) and (10) are combined with (11) and (13) to give

$$\delta_b^{(20)} = Lk \int_0^{r_c} \frac{(25 - 0{\cdot}1E_p^{0{\cdot}5} e^{-0{\cdot}5r/r_b})\frac{1}{2} r dr}{(uR_d)^2 + r^2} + \frac{Lk380}{E_p^{0{\cdot}35}} \int_{r_c}^{r_1} \frac{e^{-0{\cdot}35r/r_b} r dr}{(uR_d)^2 + r^2} \qquad \text{(for } E_p = 20 \text{ keV)} \tag{17}$$

The limit r_c in (17) is the value of the radius at which the change-over occurs between the energy loss relations (9) and (10), i.e. from (11).

$$E_p e^{-r_c/r_b} = 1 \times 10^4 \text{ eV}$$

or

$$r_c = 0{\cdot}69\, r_b,\ 1{\cdot}1 r_b \text{ for } E_p = 20, 30 \text{ keV respectively.}$$

A similar relation to (17) may be derived for a primary beam energy of 30 keV.

The above expressions have been evaluated numerically to obtain the ratio (δ_p/k) for three different primary beam energies (Table 2). The ratio (δ_f/k), obtained from equations (6), (9) and (10), is also tabulated. Experimentally determined values of $\delta(=\delta_f + \delta_b)$ given by Reimer, Seidel & Gilde (1968) have been used to derive δ_f and δ_b separately. Also given in Table 2 is the ratio $\beta = (\delta_b/\eta)/\delta_f$ which gives a measure of the proportion of low-energy secondaries produced per backscattered primary in comparison with that per incident primary. These values are comparable with the value of 5 for copper at 4 keV obtained by Seiler (1967) using a different method.

The relative proportions of δ_f and δ_b in δ are shown in Fig. 7.

Table 1. Backscattering coefficients and parameters relating to the diffusion model for backscattering

E_p (keV)	10	20	30
η*	0·34	0·33	0·32
$u(=1-2\eta)$	0·32	0·34	0·36
R_d (μm)	0·19	0·55	0·9
uR_d (μm)	0·06	0·19	0·32

* Experimental values after Bishop (1966).

Table 2. Secondary emission coefficients and related parameters

E_p (keV)	10	20	30
r_1 (μm)	0·23	0·74	1·3
r_c (μm)	—	0·1	0·25
δ_b/k (eV)	37	33·5	31
δ_f/k (eV)	22·5	15·8	12·8
δ*	0·22	0·135	0·095
δ_b	0·137	0·092	0·067
δ_f	0·083	0·044	0·028
β	4·8	6·4	7·5

* Experimental values after Reimer, Seidel & Gilde (1968).

ESTIMATION OF THE SECONDARY EMISSION COEFFICIENT FOR AN ASPERITY

For the purposes of the calculations which follow, it is convenient to assume that the surface asperities are of Gaussian shape with height and radius always approximately equal, also that the primary beam contains a Gaussian distribution of current. The parameters defining the volume and current profiles are shown in Fig. 8. The diameter $2r_j$ contains 98% of the total beam current.

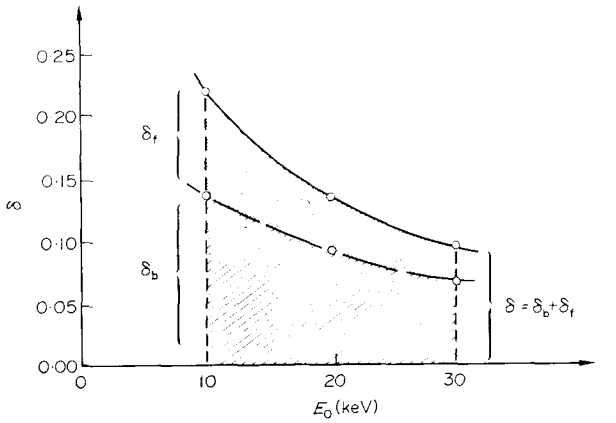

Fig. 7. Secondary emission coefficients for copper.

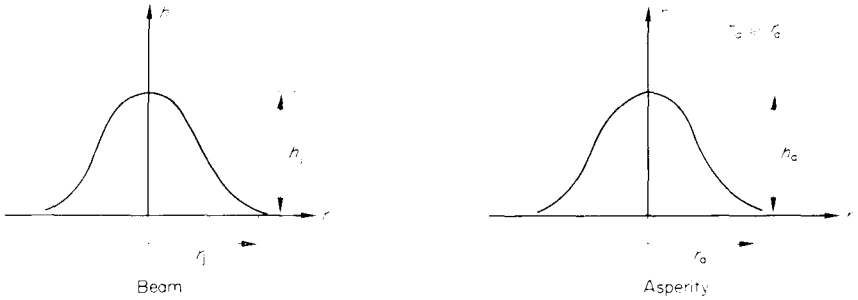

Fig. 8. Gaussian profiles for electron beam and surface asperities. For beam: current density $h = h_j e^{-(2r\ r_j)^2}$. For asperity: height $h = h_a e^{-(2r\ r_a)^2}$.

Consider first an infinitely thin beam entering an asperity as illustrated in Fig. 1b. If the asperity has dimensions of the order of L, the secondaries produced have a greater chance of escape and the signal is increased compared with that for the flat surface. Exact expressions for δ_p may be derived by numerical solution of a transcendental equation (Catto, 1971), but in view of the approximate nature of the calculations made elsewhere and the spread in some of the experimental values used, this approach is hardly justified. Instead an estimate of δ_p is made as follows: referring to Fig. 9, the probability of escape of secondaries depends on the escape cone angle and the escape path length. At point A the escape cone half-angle is about 135°, and the path length varies from zero to a maximum of about r_a. At point B the half-angle is 90°, and the path length varies between about $0{\cdot}7r_a$ and r_a. Average values are about 110° and $0{\cdot}7r_a$. From equation (1), the attenuation over this mean path is about $e^{-0{\cdot}7r_a/L}$. Comparing the situation with the flat-surface case in which the escape cone is 30° (Fig. 2), the increase in solid angle is about a factor of 10. Hence, we may write for δ_p:

$$\delta_p \approx 10k\left(-\frac{dE}{dx}\right)h_a e^{-0{\cdot}7r_a/L}$$

$$\therefore \delta_p \approx 10\delta_f \frac{r_a}{L} e^{-0{\cdot}7r_a/L} \qquad (18)$$

using equation (6) and putting $r_a = h_a$.

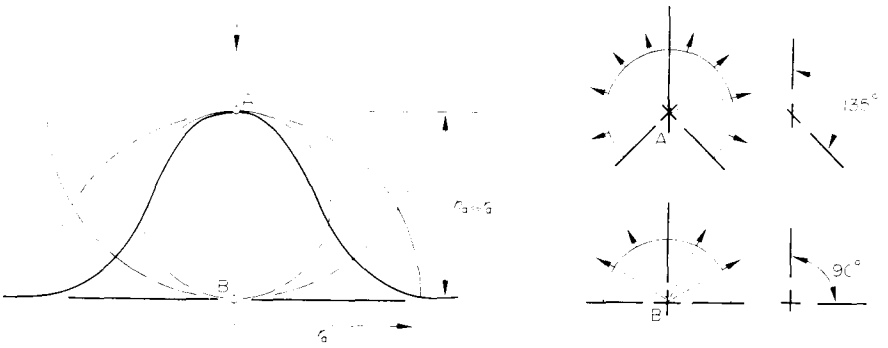

Fig. 9. Estimation of δ_p: approximate escape cone angles and path lengths for secondaries.

This expression for δ_p gives an estimate of the secondaries which escape from the asperity and add to the signal. The presence of the asperity also modifies the background signal from the flat surface. This is illustrated in Fig. 10. As the beam approaches the asperity and the escape cone and asperity overlap, some of the secondaries formerly contributing to δ_f are intercepted and consequently the signal level falls until the point is reached at which the primary beam actually starts producing secondaries within the asperity. The resulting dip in the signal should enhance contrast, but it can be shown that under the limiting resolution conditions considered in this paper the effect is small and has a negligible effect on contrast (Catto, 1971).

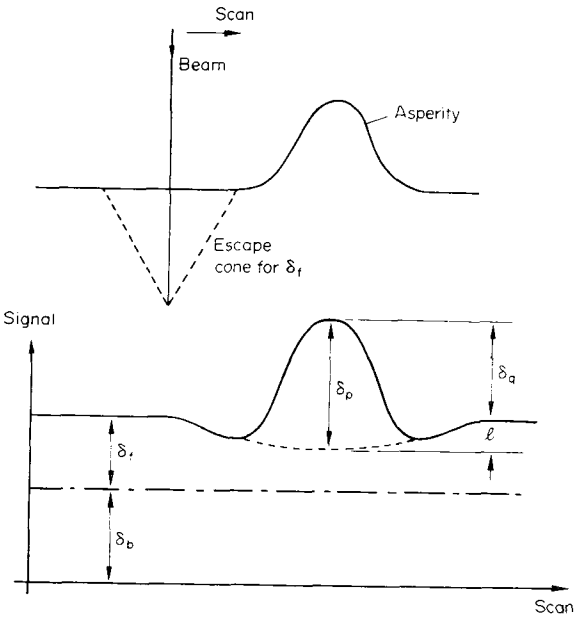

Fig. 10. Signal levels at an asperity (unit-primary beam current). l is the occlusion loss; δ_q is the peak signal above background.

When the beam is directly incident on the asperity, however, occlusion of the δ_f component of the signal may be appreciable. If l is the reduction in secondary current caused by occlusion (unit primary beam current assumed), then the nett secondary emission coefficient for the asperity will be $(\delta_p - l)$. From equation (4) the occlusion loss l may be written approximately as:

$$l \approx \frac{\delta_f}{GL}\int_0^\infty e^{-r/GL}(1 - e^{-h/L})dr \tag{19}$$

where $h = h_a e^{-(2r/r_a)^2}$ for a Gaussian-shaped asperity.

So far the case of an infinitely thin beam only has been considered. For a beam of finite diameter scanning across an object, the signal will rise earlier and fall later than for the infinitely thin beam case, resulting in a larger apparent object size. Also, for the same total beam current, the current density will be less and the peak signal will be lowered.

The factors by which the signal is reduced and broadened are obtained from the two-dimensional convolution of beam with asperity (Schade, 1955). The result of such a convolution for the Gaussian profiles of Fig. 8 is a reduction of the peak signal by a factor F_p and a broadening by a factor F_r where

$$F_p = [1 + (r_j/r_a)^2]^{-1} \tag{20}$$

and

$$F_r = [1 + (r_j/r_a)^2]^{1/2}. \tag{21}$$

The factors F_p and F_r are plotted in Fig. 11.

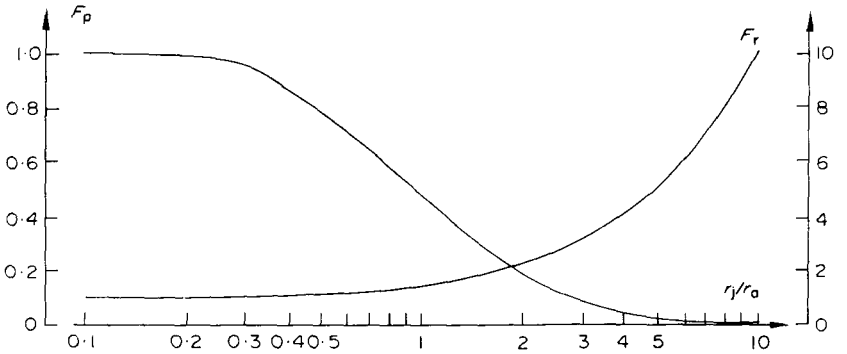

Fig. 11. Peak and radius factors derived from the convolutions of beam and asperity.

For a beam of non-zero diameter, therefore, the coefficient δ_p must be multiplied by the peak factor F_p, and in calculating the occlusion loss from (19) the generalized value G_m must be used instead of G where

$$G_m = [G_0^2 + (r_j/L)^2]^{1/2} \qquad (G_0 \approx 0{\cdot}5). \tag{22}$$

The nett coefficient for the asperity is then

$$\delta_q = F_p\delta_p - l_m \tag{23}$$

where l_m is the modified occlusion loss derived from (19) and (22). Using equation (18) δ_q may be written:

$$\delta_q = 10\frac{F_p r_a}{L}e^{-0{\cdot}7r_a/L} - l_m. \tag{24}$$

Some values of δ_p, l_m and δ_q for copper are given in Table 3. These values have been calculated for three different beam energies and beam diameters $2r_j$ of 0, 0·5 nm and 5 nm. It will be noted that variation of the diameter $2r_a$ of the asperity at constant beam diameter and energy gives rise to a maximum in δ_q.

NOISE

In the determination of the signal-to-noise ratio in the SEM three sources of noise must be taken into account: primary beam shot noise, secondary electron production noise and secondary electron escape noise. Secondary emission noise has been discussed by deHaan (1960) whose treatment is followed below.

Suppose the primary beam contains n electrons per unit area and unit time, and each of these electrons meets a large and constant number p of electrons in the material. Let there be a small and constant chance q that one of these electrons is excited and a small and constant chance r that an excited electron is emitted. If on average each incoming primary electron gives rise to s emitted secondaries, then

$$s = pqr$$

and for n primaries the number of emitted secondaries will be

$$z = ns = npqr.$$

The analysis applied by deHaan (1960) allows Gaussian fluctuations in n, q and r in turn, with the result that the total mean-square noise may be written:

$$\overline{(\Delta z_{\text{total}})^2} = \underset{\substack{\text{(escape}\\\text{noise)}}}{\bar{n}s} + \underset{\substack{\text{(production}\\\text{noise)}}}{\bar{n}sr} + \underset{\substack{\text{(beam}\\\text{noise)}}}{\bar{n}s^2}.$$

In the secondary emission processes considered in relation to SEM there are three distinguishable sources of secondaries represented by δ_b, δ_f and $F_p\delta_p$. Consequently, deHaan's (1960) expression must be used repeatedly to give the total mean-square noise:

$$(\Delta z_s)^2 = \bar{n} \sum_{i=b,f,p.} (s + sr + s^2)_i \qquad (25)$$

where $s_b = \delta_b$; $s_f = \delta_f - l_m$; $s_p = F_p\delta_p$.

Table 3. Secondary emission coefficient for an asperity

		E_p (keV)								
		10			20			30		
		$2r_j$ (nm)								
r_a/L		0	0·5	5	0	0·5	5	0	0·5	5
0·1	δ_p	0·077	0·077	0·077	0·041	0·041	0·041	0·026	0·026	0·026
	l_m	0·001	0·001	—	—	—	—	—	—	—
	δ_q	0·076	0·02	—	0·04	0·01	—	0·026	0·007	—
0·2	δ_p	0·144	0·144	0·144	0·076	0·076	0·076	0·049	0·049	0·049
	l_m	0·003	0·003	0·001	0·002	0·001	—	0·001	0·001	—
	δ_q	0·141	0·08	0·001	0·074	0·041	0·001	0·048	0·026	—
0·5	δ_p	0·293	0·293	0·293	0·153	0·153	0·153	0·099	0·099	0·099
	l_m	0·011	0·011	0·005	0·006	0·006	0·002	0·004	0·004	0·002
	δ_q	0·282	0·252	0·02	0·147	0·132	0·01	0·015	0·085	0·007
1·0	δ_p	0·411	0·411	0·411	0·215	0·215	0·215	0·139	0·139	0·139
	l_m	0·024	0·024	0·014	0·013	0·013	0·007	0·008	0·008	0·005
	δ_q	0·387	0·374	0·095	0·202	0·195	0·05	0·131	0·127	0·032
1·5	δ_p	0·435	0·435	0·435	0·228	0·228	0·228	0·147	0·147	0·147
	l_m	0·03	0·031	0·025	0·015	0·016	0·013	0·01	0·01	0·008
	δ_q	0·406	0·398	0·17	0·213	0·21	0·088	0·137	0·135	0·058
2·0	δ_p	0·408	0·408	0·408	0·214	0·214	0·214	0·138	0·138	0·138
	l_m	0·028	0·03	0·034	0·015	0·016	0·018	0·009	0·01	0·012
	δ_q	0·38	0·375	0·207	0·119	0·196	0·108	0·128	0·127	0·07
4·0	δ_p	0·202	0·202	0·202	0·105	0·105	0·105	0·068	0·068	0·068
	l_m	0·01	0·013	0·052	0·005	0·007	0·027	0·004	0·004	0·018
	δ_q	0·192	0·189	0·12	0·10	0·062	0·062	0·065	0·064	0·04

The useful signal is $\delta_q = F_p\delta_p - l_m$ (Table 3). Thus, using equation (25), the signal-to-noise ratio may be calculated.

The situation is complicated by the fact that there is some correlation between the noise components (e.g. a backscattered primary that has diffused towards the surface far away from the point of entry of the primary beam may have energy and direction such that it can produce as many as ten or more secondaries together). A correction to allow for this has been made in the signal-to-noise calculations presented here (Catto, 1971).

In the calculations it has been assumed that the mean number of electrons falling on each picture point $\bar{n} = 10^4$, which implies a beam current of about $2{\cdot}5 \times 10^{-12}$ A for a picture of 400^2 points and a frame time of 100 sec. It has been demonstrated by Crewe & Wall (1970) that this order of current can be provided in a 0·5 nm probe in an instrument employing a field-emission source and operating at 20 kV. In a conventional SEM employing a tungsten thermionic source, the probe must be of the order of 5 nm to achieve a similar current. These two cases are considered below.

The results of the calculations are presented in Fig. 12 in which signal-to-noise ratio is plotted against the normalized radius (r_a/L) for three different beam energies (10, 20, 30 keV) and for three beam diameters $2r_j$ corresponding to 0, 0·5 and 5 nm. Since the r.m.s. noise is proportional to $\sqrt{\bar{n}}$, increasing the beam

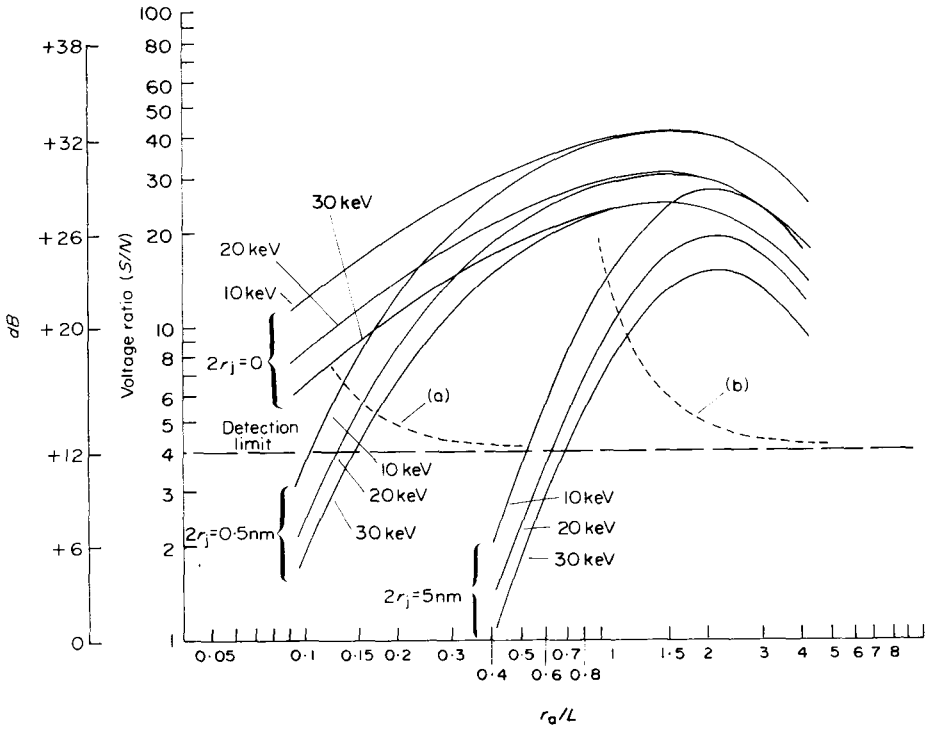

Fig. 12. Signal-to-noise ratio versus normalized radius of asperity r_a/L for various primary beam energies and beam diameter $2r_j$. Curves are derived on the basis of 10^4 primary electrons per picture point. The dashed curves indicate loci for $(S/N)_{Resln}$ parameter defined by equation (30); (a) for Beam diameter $2r_j = 0{\cdot}5$ nm, (b) for beam diameter $2r_j = 5$ nm.

current from $2{\cdot}5 \times 10^{-12}$ A by a factor F_j results in an improvement in signal-to-noise ratio of a factor $\sqrt{F_j}$ and the curves in Fig. 12 are shifted up by $10 \log_{10} F_j db$.

There will be a limiting signal-to-noise ratio below which an asperity on a flat surface cannot be detected reliably. Making the usual assumption that the noise amplitude follows a Gaussian distribution with standard deviation σ (=r.m.s. value), an error in detection of the asperity can arise from: (a) the instantaneous noise amplitude being so large and negative that the sum (noise + signal) is below display black-level, even when the beam is incident directly on the asperity; (b) the instantaneous noise amplitude being so large and positive that, even in the absence of the asperity, the sum (noise + signal) exceeds display black-level. These cases are illustrated in Fig. 13a, b. If the signal S has an amplitude twice the back-bias B, the probability of error P_e is given by

$$\begin{aligned} P_e &= P(|\text{ noise-amplitude }| > B) \\ &= 2P(\text{noise-amplitude} > B) \end{aligned}$$

by symmetry; e.g. for $P_e = 5\%$, using tables of the normal distribution,

$$B = 1{\cdot}96\sigma$$

$$\therefore S \geqslant 4\sigma \text{ approximately.}$$

That is, the signal-to-noise ratio must be greater than or equal to 4.

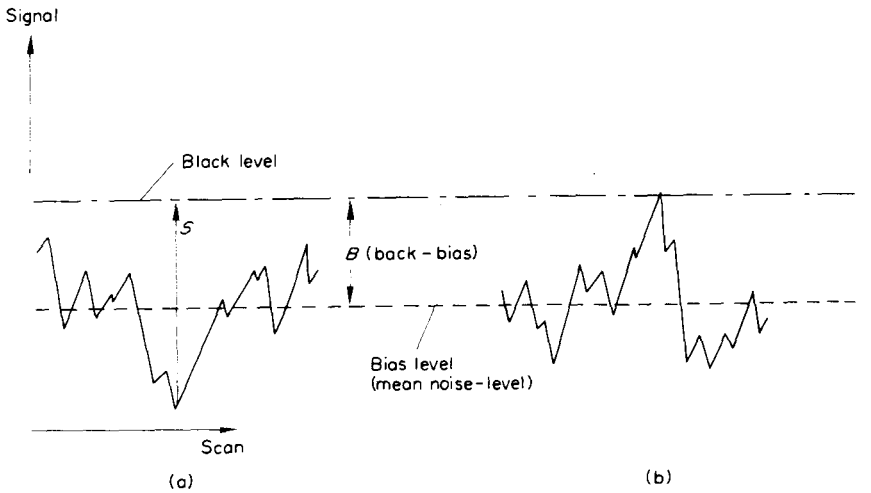

Fig. 13. Limiting cases for detecting signal in noise.

This compares with the threshold signal-to-noise ratio k found by Rose (1948) to have a value of between 3 and 6 in his experiments on perception.

For satisfactory imaging with m grey levels the signal-to-noise ratio must be m times higher; for example, for 5 grey levels the ratio must be greater than or equal to 20, i.e., better than $26dB$. Thus, the condition for imaging detail at various grey levels is more stringent than for detection of detail. This is discussed in more detail in Catto (1971).

RESOLUTION LIMITS

Consider two equal asperities with $r_a(=h_a) < L$ so that occlusion and absorp-

tion effects are small. Then, for a scanning line passing through the centre of the asperities, the signal S will be of the form

$$S = S_0\{e^{-[2(r-r_s)/r_e]^2} + {}^{-[2(r+r_s)/r_e]^2}\}$$
$$= S_0 2e^{-(2r_s/r_e)^2} e^{-(2r/r_e)^2} \cosh(8rr_s/r_e^2) \quad (26)$$

where S_0 is a constant; $2r_s$ is the separation between the asperities, with the origin taken mid-way between them (Fig. 14), and

$$r_e = [r_a^2 + r_j^2]^{1/2}. \quad (27)$$

The term $e^{-(2r_s/r_e)^2}$ in (26) is a scaling factor and the central dip in the signal C_d (Fig. 15) is approximately proportional to

$$F_d = \{1 + e^{-(4r_s/r_e)^2} - 2e^{-(2r_s/r_e)^2}\}. \quad (28)$$

When r_j is large compared with r_a and r_s, the beam Gaussian term $e^{-(2r/r_e)^2}$ in (29) swamps the separation term cosh $(8rr_s/r_e^2)$.

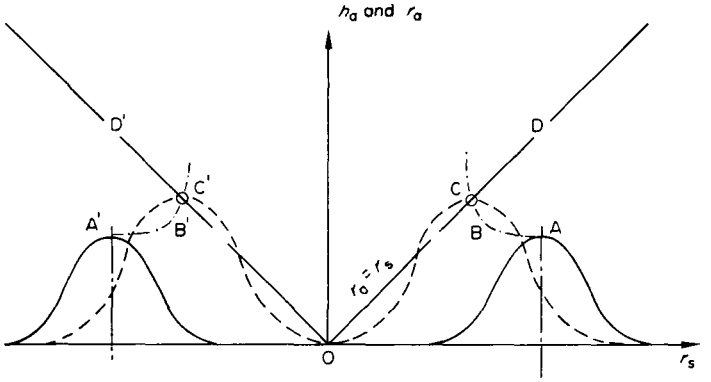

Fig. 14. Illustrating criteria for resolution of two equal asperities.

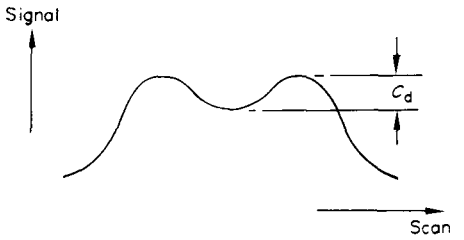

Fig. 15. Central dip signal.

If r_a is so small that one asperity on its own is at the detection limit given in Fig. 12, it should just be possible to resolve two such asperities if $r_s \geqslant r_e$. In Fig. 14, two asperities at the limit $r_s = r_e$ are shown by the full lines having peaks A and A'. If r_s is now reduced below r_e, the size of the asperities must be increased to compensate, if the asperities are still be to resolved. The chain-link curves ABC and A'B'C' trace the loci of the peaks as r_s is reduced and r_a and h_a increased. A limit is set by the lines OD, OD' (with equations $r_a = |r_s|$) when the asperities begin to coalesce (dashed lines in Fig. 14, peaks at C and C').

The loci ABC (and A'B'C') may be determined as follows: Having set $r_s = r_e$ by the values of r_a, r_j at the detection limit (point M in Fig. 16), suppose r_a is

increased until the point P is reached. The loss in central-dip signal now that $r_s < r_e$ may be counterbalanced by the excess signal-to-noise ratio represented by QP,

$$\text{i.e. } +\text{QP} = 20 \log_{10} \frac{(F_d)_M}{(F_d)_P} \tag{29}$$

where $(F_d)_M$, $(F_d)_P$ signify the values of F_d at points M and P respectively. Using equations (27) and (28), and noting that $(F_d)_M \approx 1$, $(r_s)_P$ may be determined. The resolution limit is given by the intersection of the curve ABC and the line OD at C (Fig. 14).

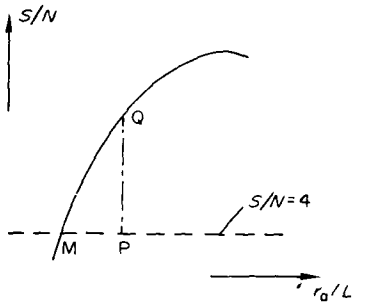

Fig. 16. Construction for finding S/N and resolution limits from Fig. 12. (Section of one curve shown in Fig. 12, intersecting line for $S/N = 4$ at point M).

The resolution limit may also be obtained by putting $r_s = r_a$ in the equation for F_d (eq. 28) and finding the intersections between the signal-to-noise ratio (S/N) curves in Fig. 12 and the locus

$$(S/N)_{\text{Resln}} = 4\left\{\frac{(F_d)_{r_s=r_e}}{(F_d)_{r_s=r_a}}\right\}$$
$$\approx 4/(1 + e^{-16/F_r^2} - 2e^{4/F_r^2}), \tag{30}$$

where $F_r^2 = 1 + (r_j/r_a)^2$.

The $(S/N)_{\text{Resln}}$ loci for beam diameters of 0·5 nm and 5 nm are marked in Fig. 12, and a summary of the detection and resolution limits derived from Fig. 12 is presented in Table 4.

Table 4. Detection and resolution limits

Primary beam energy E_p (keV)	10		20		30	
Probe diameter $2r_j$ (nm)	0·5	5	0·5	5	0·5	5
Detection limit $2r_a$ (nm)	0·32	1·6	0·38	1·9	0·44	2·1
Resolution limit $2r_a$ (nm)	0·41	2·9	0·47	3·2	0·52	3·3

DISCUSSION

The resolution limits presented in Table 4 must be interpreted with due regard to the approximate nature of the theory and experimental inaccuracies in some of the important physical quanties involved. Furthermore, the assumption that the normal secondary emission theory, valid for large surface areas and volumes of material, is also valid for small asperities approaching lattice dimensions must be questioned. Channelling and diffraction effects in a metallic

specimen may, at this dimensional level, give rise to quite different contrast mechanisms from those discussed above. The investigations of Seiler & Kuhnle (1970) on the variation of secondary emission with crystal orientation are relevant in this context. However, the theory does indicate that secondary electron diffusion effects may not be as important as hitherto supposed in limiting resolution and that, with the fine electron probes now potentially available, topographical detail at the 1 nm level might be resolvable on a metallic specimen. It should be noted that the limiting resolution does not depend upon the diameter from which the secondary electrons escape at the surface of the specimen.

A major factor limiting the available signal-to-noise ratio, and hence resolution, is the secondary electron component δ_b due to backscattered primaries. These primaries also excite secondary electrons at the lens pole-pieces and other surfaces surrounding the specimen, and the signal resulting from this source can form an appreciable fraction of the total signal current (Smith, 1956). To obtain high resolution therefore, it would appear to be advantageous to remove as much of the backscattered primary component as possible by preparing specimens in a form thin enough to allow transmission of an appreciable part of the primary beam. Secondaries produced by the transmitted beam at the back surface of the specimen and elsewhere would, of course, have to be suppressed to prevent their entering the detector.

At sharply contoured features on a specimen surface, it is possible for the penetration envelope (Fig. 1c) to break through the surface in a region remote from the point of entry of the beam. This gives rise to the characteristic bright bands which are commonly observed in scanning micrographs. Such conditions are unfavourable for the resolution of fine detail and it is to be expected that, on a rough surface, the regions in which high resolution might be obtainable would be restricted. Penetration effects are reduced at lower primary beam energies but this advantage is offset by the reduction in beam current. It will be noted that the resolution figures given in Table 4 are relatively insensitive to primary beam energy.

The difficulties of collecting the secondary electrons from the surface of a specimen immersed in a high-power magnetic objective lens has already been mentioned. In this respect some advantage may be gained from the use of a lens with an asymmetric pole-piece configuration, both the incident primary beam entering and the secondary electrons leaving on the wider-bore side of the lens.

Another important practical consideration is the problem of contamination in the SEM. Very low partial pressures of hydrocarbons and other residuals would be necessary to avoid detail at the 1 nm level from being obscured.

In this paper consideration has been given to the image formed by collecting the secondary electrons only. It has been shown by Catto (1971) that collection of the total backscattered component is less favourable for the production of high resolution images. The alternative, proposed by Wells (1972), of collecting only those electrons which have suffered very small energy losses may in principle also offer the possibility of achieving very high resolution. However, collection of the backscattered component and separation of the low-loss electrons, under the optical conditions necessary to achieve probe diameters of the order of 1 nm, would seem to be even more difficult than collection of the secondary electron component.

ACKNOWLEDGMENTS

We are grateful for the financial support received from the Wolfson Foundation, the Royal Radar Establishment and the Science Research Council.

References

Archard, G.D. (1961) Back scattering of electrons. *J. appl. Phys.* **32**, 1505.

Bishop, H.E. (1966) *Electron scattering and X-ray production.* Ph.D. Thesis, University of Cambridge.

Broers, A.N. (1969) A new high resolution reflection scanning electron microscope. *Rev. scient. Instrum.* **40**, 1040.

Bruining, H. (1954) *Physics and Applications of Secondary Electron Emission.* Pergamon, London.

Catto, C.J.D. (1971) *Production and storage display of the signal in the scanning electron microscope.* Ph.D. Thesis, University of Cambridge.

Cosslett, V.E. & Thomas, R.N. (1964) Multiple scattering of 5–30 keV electrons in evaporated metal films. II. Range-energy relations. *Br. J. appl. Phys.* **15**, 1283.

Crewe, A.V. & Wall, J. (1970) A scanning microscope with 5 Å resolution. *J. molec. Biol.* **48**, 375.

deHaan, E.F. (1960) Signal-to-noise ratio of image devices. *Adv. Electronics Electron Phys.* **12**, 291.

Everhart, T.E. (1958) *Contrast formation in the scanning electron microscope.* Ph.D. Thesis, University of Cambridge.

Everhart, T.E. (1970) Contrast and resolution in the scanning electron microscope. *Third Annual Cambridge Stereoscan Colloquim*, Chicago.

Kanter, H. (1961) Energy dissipation and secondary electron emission in solids. *Phys. Rev.* **121**, 681.

Komoda, T. & Saito, S. (1972) Experimental resolution limit in the secondary electron mode for a field emission source scanning electron microscope. *Proceedings of the Fifth Annual Symposium on Scanning Electron Microscopy* 1972, p. 129. IIT Research Institute, Chicago.

Lewis, H.W. (1950) Multiple scattering in an infinite medium. *Phys. Rev.* **78**, 526.

Lye, R.G. & Dekker, A.J. (1957) Theory of secondary emission. *Phys. Rev.* **107**, 977.

Pease, R.F.W. & Nixon, W.C. (1965) High resolution scanning electron microscopy. *J. scient. Instrum.* **42**, 81.

Reimer, L., Seidel, H. & Gilde, H. (1968) Einfluss der Elektronendiffusion auf die Bildentstehung im Raster-Elektronen-mikroskop. *Beitr. Elektronmikroskop Direktabb. Oberfl. (Munster)*, **1**, 53.

Rose, A. (1948) Television pickup tubes and the problem of vision. *Adv. Electronics*, **1**, 131.

Schade, O.H. (1955) Image analysis in photographic and television systems (definition and sharpness). *J. Soc. Motion Pict. Telev. Engrs.* **64**, 593.

Seiler, H. (1967) Einige aktuelle Probleme der Sekundarelektronenemission. *Z. angew. Phys.* **22**, 249.

Seiler, H. & Kuhnle, G. (1970) Zur Anisotropie der Elektronenausbeute in Abhangigkeit von der Energie der auslosenden Primarelektronen von 5 bis 50 keV. *Z. angew. Phys.* **29**, 254.

Smith, K.C.A. (1972) Scanning electron microscopy, the next ten years. *Proceedings of the Fifth Annual Symposium on Scanning Electron Microscopy* 1972, p. 1. IIT Research Institute, Chicago.

Wells, O.C. (1957) *The construction of a scanning electron microscope and its application to the study of fibres.* Ph. D. Thesis, University of Cambridge.

Wells, O.C. (1972) Explanation of the low-loss image in the SEM in terms of electron scattering theory. *Proceedings of the Fifth Annual Symposium on Scanning Electron Microscopy* 1972, p. 169. IIT Research Institute, Chicago.