Interpretation of Holographic Contour Maps of Reverse Biased p-n Junctions

Clara Capiluppi(1), Andrea Migliori(2) and Giulio Pozzi(1)

(1) Department of Physics and Istituto Nazionale di Fisica della Materia, University of Bologna, viale B. Pichat 6, 40127 Bologna, Italy
(2) CNR–LAMEL, via P. Gobetti 101, 40129 Bologna, Italy
(Received March 8; accepted June 12, 1995)

Abstract. — Reverse-biased p-n junctions have been observed by means of electron holography using a transmission electron microscope equipped with an electron biprism and a field emission gun. Aim of this work is to present and discuss an analytical model for the electric field associated to a periodic array of alternating p and n stripes lying in a half-plane which simulates the experimental set-up and allows the interpretation of the main features of the observed holograms and the quantitative evaluation of the effect of the fringing field on the holograms and on the reconstructed images.

1. Introduction

The problem of observing reverse-biased p-n junctions by means of transmission electron microscopy techniques is a very challenging one both from the experimental and theoretical points of view.

In particular, in addition to standard Lorentz microscopy techniques, whose image contrast is able to give some information on the region of high field [1], also holographic methods have been successfully applied and an interference map of the two-dimensional phase on the specimen plane obtained [2, 3].

Unfortunately, after the first successes, this research line was put aside, owing mainly to instrumental and specimen preparation problems. Nonetheless, it has been the analysis of the p-n junction holographic interference maps and of the troubles associated in their processing and interpretation which prompted the investigations on long range electromagnetic fields [4], whose peculiar feature is represented by the fact that the so-called reference wave is in reality perturbed by the tail of the fringing field.

Today, with the new generation of field emission electron microscopes, holographic experiments on reverse-biased p-n junctions can be carried out under less demanding and stringent conditions, so that we feel it worthwhile to describe in some detail the theoretical results we have obtained, previously reported only in the very concise form of an extended abstract [5]. It will
be shown how the analytical solution can be obtained for the electrostatic field associated to a periodic array of alternating p and n regions lying in a semi-infinite plane. From this solution the phase shift can be calculated and used for the interpretation of the main features of the holographic contour maps.

2. The Electrostatic Field Model

Let us consider the following boundary value problem [6], Figure 1, i.e. finding the electrostatic potential produced by a parallel array of stripes having pitch b, which lie in the half-plane \( z = 0; x \geq 0 \), and are biased at alternate potential, namely \(+V_0\) for \(2nb < y \leq (2n + 1)b\) and \(-V_0\) for \((2n + 1)b < y \leq (2n + 2)b\), \(n\) being an arbitrary positive or negative integer.

On the half-plane, the potential is given by

\[
V(x, y, 0) = \frac{V_0}{4} \sum_{k=0}^{\infty} \sin \left( \frac{2k + 1}{b} \pi y \right) \frac{2k + 1}{2k + 1}
\]

so that, considering separately each Fourier component, we have to find a solution of Laplace’s equation in the form

\[
G_\lambda(x, y, z) = \Phi_\lambda(x, z) \sin \lambda y
\]

where \( \lambda = \frac{2k + 1}{b} \pi \).

It follows that \( \Phi_\lambda(x, z) \) should satisfy the following differential equation

\[
\frac{\partial^2 \Phi_\lambda}{\partial x^2} + \frac{\partial^2 \Phi_\lambda}{\partial z^2} - \lambda^2 \Phi_\lambda = 0
\]
with the boundary condition

\[ \Phi_\lambda(x \geq 0, 0) = 1 \]  

As the half-plane \( z = 0; x \geq 0 \) can be considered as the limiting case of an infinite parabolic cylinder, it is convenient to solve equation (3) by the method of separation of variables in parabolic coordinates \( \xi, \eta \) related to the cartesian coordinates \( x, z \) by the relations

\[ x = \frac{(\xi^2 - \eta^2)}{2} \]  
\[ z = \xi \eta \]

where \(-\infty < \xi < \infty\), and \( 0 \leq \eta < \infty \).

The orthogonal system of surfaces consists of parabolic cylinders with foci at the origin, Figure 2, and the specimen half-plane corresponds to the value \( \eta = 0 \). In this coordinate system, equation (3) for the Fourier component of the electrostatic potential \( V \) takes the form

\[ \frac{1}{\xi^2 + \eta^2} \left( \frac{\partial^2 \Phi_\lambda}{\partial \xi^2} + \frac{\partial^2 \Phi_\lambda}{\partial \eta^2} \right) - \lambda^2 = 0 \]  

for which a solution is looked for by the method of separation of the variables [7]. Writing

\[ \Phi_\lambda = A(\xi)B(\eta) \]
the following equations result
\[
\frac{d^2A}{d\xi^2} - (\mu + \lambda^2 \xi^2)A = 0
\]
\[
\frac{d^2B}{d\eta^2} - (-\mu + \lambda^2 \eta^2)B = 0
\]
where \(\mu\) is a constant.

By introducing the new variables \(u = \sqrt{2\lambda}\xi\) and \(v = \sqrt{2\lambda}\eta\), and a new parameter \(\nu\) related to \(\mu\) by the formula
\[
\mu = -\lambda(2\nu + 1)
\]
equations (8) and (9) can be reduced to the form
\[
\frac{d^2A}{du^2} + \left(\nu + \frac{1}{2} - \frac{u^2}{4}\right)A = 0
\]
and
\[
\frac{d^2B}{dv^2} + \left(-\nu - 1 + \frac{1}{2} - \frac{u^2}{4}\right)B = 0
\]
respectively. The solutions \(D_\nu(u)\) of equation (11), called Weber's equation, (or of the otherwise identical Eq. (12) with index \(-
\nu - 1\) are called parabolic cylinder functions or Weber–Hermite functions and their properties are summarized in the book edited by Erdelyi [8], whose notation will be hereafter strictly followed. In addition to \(D_\nu(u)\) also \(D_\nu(-u)\), \(D_{\nu-1}(iu)\) and \(D_{\nu-1}(-iu)\) satisfy Weber's equation. As there are only two linearly independent solutions, the above solutions are connected.

In particular, if \(\nu = n\) is a non-negative integer, then
\[
D_n(u) = 2^{-(1/2)n} \exp\left(-\frac{u^2}{4}\right) H_n(2^{-(1/2)}u)
\]
where \(H_n(u)\) is the Hermite polynomial of degree \(n\), being \(H_0(u) = 1\), whereas if \(\nu\) is a negative integer, then \(D_\nu(u)\) can be expressed in terms of the complementary error function \(\text{Erfc}\) defined by
\[
\text{Erfc}(x) = \int_x^\infty \exp(-t^2)dt
\]
that is
\[
D_{-m-1}(u) = 2^{1/2} \frac{(-1)^m}{m!} \exp\left(-\frac{u^2}{4}\right) \frac{d^m}{du^m} \left[ \exp\left(\frac{u^2}{2}\right) \text{Erfc}(2^{-1/2}u) \right]
\]
so that
\[
D_{-1}(u) = 2^{1/2} \exp\left(\frac{u^2}{4}\right) \text{Erfc}(2^{-1/2}u)
\]

Unfortunately, the solution of equation (3) with the boundary condition (4) cannot be found by using Hermite polynomials, as outlined by Lebedev [7], since the value of the function at the boundary does not satisfy the conditions for its expansion in series of Hermite polynomials.

Therefore, we should use the more general Hermite functions: when \(\nu\) is not integer, \(D_\nu(u)\) and \(D_\nu(-u)\) are linearly independent, so that a particular solution of equations (11) and (12) can be written as
\[
A(u) = PD_\nu(u) + QD_\nu(-u)
\]
and

\[ B(v) = RD_{\nu-1}(v) + SD_{\nu-1}(-v) \]  

(18)

respectively.

Noting that the required solution of equation (3) should be symmetric with respect to the plane \((x, y)\), then \(A(u)\) should be unchanged by the substitution \(u \rightarrow -u\), which implies that \(P = Q\). It follows that the most general solution for the Fourier component of the potential can be put in the form

\[ \Phi_\lambda = \int \left[D_\nu(u) + D_\nu(-u)\right][RD_{\nu-1}(v) + SD_{\nu-1}(-v)]d\nu \]  

(19)

where the coefficients \(R\) and \(S\) are function of \(\nu\), and should be determined in such a way that \(\Phi_\lambda\) satisfies the boundary condition, (4), which in parabolic coordinates becomes

\[ \Phi_\lambda(u, 0) = 1 \]  

(20)

Putting \(S = 0\) for simplicity, this leaves us to determine only the function \(R\), and this can be done by comparing our expression with the one which can be deduced from Erdelyi’s formula (11) in Section 8.5.2 of the reference [8], specifying there \(c = -1/2, t = 1\) and \(y = 0\)

\[ \frac{(1/2)\pi^{1/2}}{2\pi i} \int_{-i\infty}^{-(1/2)+i\infty} \frac{d\nu}{\sin(-\nu\pi)} \left[D_\nu(x)D_{\nu-1}(0) + D_\nu(-x)D_{\nu-1}(0)\right] = 2^{-1/2} \]  

(21)

It results that

\[ R = \frac{i}{2\sqrt{\pi} \sin \nu \pi} \]  

(22)

so that our solution for the field in the whole plane can be written:

\[ \Phi_\lambda = \int_{-i\infty}^{-(1/2)+i\infty} \frac{i}{2\sqrt{\pi} \sin \nu \pi} \left[D_\nu(u)D_{\nu-1}(v) + D_\nu(-u)D_{\nu-1}(v)\right]d\nu \]  

(23)

Our final solution can be changed into a more manageable form, if it is compared with Cherry’s formula (8) in the same Section 8.5.2 [8], which, rewritten by putting \(\phi = \pi/2, h\zeta = u\) and \(h\eta = v\), becomes

\[ -\sqrt{2i}D_0 \left[\frac{u + v}{\sqrt{2}}\right] D_{-1} \left[\frac{u + v}{\sqrt{2}}\right] = \int_{-1/2-i\infty}^{-(1/2)+i\infty} D_\nu(u)D_{\nu-1}(v) \frac{d\nu}{\sin \nu \pi} \]  

(24)

so that it finally ensues

\[ \Phi_\lambda = \frac{1}{\sqrt{2\pi}} \left\{ D_0 \left[\frac{-u + v}{\sqrt{2}}\right] D_{-1} \left[\frac{u + v}{\sqrt{2}}\right] + D_0 \left[\frac{u + v}{\sqrt{2}}\right] D_{-1} \left[\frac{-u + v}{\sqrt{2}}\right] \right\} \]  

(25)

It should be noted that Cherry’s formula has been derived with complex arguments in relation to the solution of the Helmholtz equation to find Sommerfeld’s secondary wave, while in our case we are dealing with real arguments and the solution of Laplace’s equation. Nonetheless, it can be easily ascertained that our final solution is the right one, being the correct combination of parabolic cylinder functions and satisfying the boundary condition, but with the peculiarity that its orientation has been rotated by \(\pi/2\) with respect to the original one, Figure 2.
By using the inverse transform between parabolic and cartesian coordinates

\[ \xi = \text{sign} z \sqrt{x^2 + z^2 + x} \]  
\[ \eta = \sqrt{x^2 + z^2 - x} \]  

Equation (25) can be transformed, in the half-space \( z \geq 0 \) into

\[ \Phi_\lambda(x, z) = \frac{1}{\pi} \left[ \exp(\lambda z) \text{Erfc} \left( \sqrt{\lambda} \sqrt{x^2 + z^2 + z} \right) + \exp(-\lambda z) \text{Erfc} \left( -\sqrt{\lambda} \sqrt{x^2 + z^2 - z} \right) \right] \]  

whereas the symmetric form holds for the half-space \( z \leq 0 \). Finally, the expression for the electrostatic potential in the whole space satisfying the required boundary conditions is given by

\[ V(x, y, z) = \frac{V_0}{4} \sum_{k=0}^{\infty} \frac{\sin \left( \frac{2k + 1}{b} \pi y \right)}{2k + 1} \frac{\Phi_{2k + 1}}{b} \pi (x, z) \]  

3. Basic Principles of Hologram Recording and Processing

Let us recall some fundamental theoretical aspects concerning image codification and recording in an interferometric electron microscope \([4, 9, 10]\). By considering only elastic scattering events, the solution of the Schrödinger equation in the high energy approximation to describe the interaction between electrons and an electrostatic field gives the transmission function (i.e. the ratio between the amplitudes of the ingoing and outgoing wavefunctions) as:

\[ T(r) = \exp(i\varphi(r)) \]  

where \( r = (x, y) \) is a bidimensional vector perpendicular to the optic axis \( z \), and \( \varphi(r) \) is the phase term, given by:

\[ \varphi(r) = \frac{\pi}{\lambda E} \int_l V(r, z) dz \]  

where the integral in equation (31) is taken along a trajectory \( l \) parallel to the optic axis \( z \), and \( E \) is the accelerating voltage of the electron microscope in the non relativistic approximation. Moreover, the fact that electrons scattered by atoms at large angles are stopped by the aperture of the objective lens can be accounted for by introducing an amplitude term \( C(r) \), so that the resulting transmission function is given by:

\[ T(r) = C(r) \exp i\varphi(r) \]  

Within an interferometric or holographic electron microscope \([9, 10]\), usually equipped with a Möllenstedt-Düker electron biprism, the image wavefunction, referred back to the object plane,
is described by the following equation (33), which neglects diffraction effects due to the biprism edges and a multiplicative phase factor, and assumes the imaging to be ideal:

$$\psi(r) = T \left( r - \frac{D}{2} \right) \exp(i\pi fr) + T \left( r + \frac{D}{2} \right) \exp(-i\pi fr)$$  \hspace{1cm} (33)$$

where $f$, perpendicular to the biprism axis, is the spatial frequency wavevector corresponding to the interference fringes, referred to the object plane and $D$, parallel to $f$, is the vector that connects the points brought to interfere by the interferometry device. In absence of the object, it is found that the image displays an interference fringe system which is parallel to the biprism axis and spaced $s = 1/|f|$. The trend of this interference fringe system is modified by the object wavefunction in such a way that information on it can be obtained, either by the direct analysis of the fringes, when these are few and refer to one-dimensional objects (interference electron microscopy), or by analogic or digital processing of the interferogram when the fringe number is larger (i.e. more than 100) and the interferogram can be considered as an off-axis hologram.

Let us consider first the situation in which the reference beam $\psi_{\text{ref}}(r) = T \left( r + \frac{D}{2} \right)$ travels outside the specimen border through a field free region. It follows that $C \left( r + \frac{D}{2} \right) = 1$ and $\varphi(r + \frac{D}{2}) = \text{const} = \varphi_0$ so that the interferogram results as the superposition of the object wave $\psi_{\text{obj}}(r) = T \left( r - \frac{D}{2} \right)$ with a reference beam $\psi_{\text{ref}}(r) = T \left( r + \frac{D}{2} \right) = \exp(i\varphi_0)$.

The intensity distribution is therefore given by:

$$I(r) = 1 + C \left( r - \frac{D}{2} \right)^2 + 2C \left( r - \frac{D}{2} \right) \cos \left[ \varphi \left( r - \frac{D}{2} \right) - \varphi_0 + 2\pi fr \right]$$  \hspace{1cm} (34)$$

It should be noted that, apart from the unessential constant phase factor $\varphi_0$, the object phase $\varphi$ is stored in the hologram. When leakage fields (pure phase objects), which arise from the specimen and extend in the space around it are present, the wavefunction describing the reference wave can be written as:

$$\psi_{\text{ref}}(r) = T \left( r + \frac{D}{2} \right) = \exp \left[ i\varphi \left( r + \frac{D}{2} \right) \right]$$  \hspace{1cm} (35)$$

and the intensity distribution in the hologram becomes:

$$I(r) = 1 + C \left( r - \frac{D}{2} \right)^2 + 2C \left( r - \frac{D}{2} \right) \cos \left[ \varphi \left( r - \frac{D}{2} \right) - \varphi \left( r + \frac{D}{2} \right) + 2\pi fr \right]$$  \hspace{1cm} (36)$$

In this case the hologram stores the information due to a fictitious specimen, having amplitude and phase given by:

$$C'(r) = C \left( r - \frac{D}{2} \right)$$  \hspace{1cm} (37)$$

$$\Delta \varphi(r) = \varphi \left( r - \frac{D}{2} \right) - \varphi \left( r + \frac{D}{2} \right)$$  \hspace{1cm} (38)$$

The electron hologram, whose recorded intensity is given by (34) or (36), can be processed by optical and/or digital means. In the first case it is first developed so that its amplitude transmittance is a linear function of the intensity. The plate is then inserted in an optical bench and
illuminated by a coherent plane laser wave. A filter lens performs, in the back focal plane, the Fourier transform of the intensity distribution of the hologram.

In addition to the straight through beam, two sidebands are displayed which represent the Fourier spectrum of the image wave and its conjugate. In the off-axis scheme, if the carrier spatial frequency is three times larger than the maximum spatial frequency of the object, the sidebands do not overlap with the zero beam (autocorrelation). Therefore, one of the sideband, e.g. the object wave with spatial frequency $f_{\text{obj}}$, can be isolated by means of a spatial filter inserted in the back focal plane of the lens. Allowing a free propagation of the wave from the back focal plane downwards, the inverse Fourier transform is performed on the image plane of the lens, where an exact optical replica of the object wave function is found. To display also the phase of the image wave, a number of techniques can be applied which have no counterpart in the electron microscope. The capabilities include bright and dark field, phase contrast imaging, optical interferometry, diffraction image filtering, low angle diffraction, phase detection and its amplification.

In the simplest arrangement, an interferometric reference plane wave

$$\psi_{\text{ref}} = \exp(i(2\pi k_R r + \phi_R))$$

(39)

where $k_R$ is the spatial frequency wavevector and $\phi_R$ the longitudinal phase, is overlapped to the reconstructed wave using a Mach-Zehnder interferometer, so that the observed intensity in the image plane is given by:

$$I(r) = 1 + C''^2(r) + 2C''(r) \cos[\Delta \phi(r) + 2\pi k_0 r - \phi_R]$$

(40)

where $k_0 = f_{\text{obj}} - k_R$ is the spatial frequency wavevector. In this optical interferogram the trend of the interference fringes indicates the local phase distribution of the fictitious object.

Contrary to electron interferometry experiments, azimuth and spacing of the fringes can now be adjusted to meet the actual needs. In principle by putting $k_0 = 0$, a particular interferogram, called contour map, is obtained, whose phase can be further amplified, by optical or digital means, by a factor $n$.

Today, however, image reconstruction by numerical methods presents considerable advantages with respect to the light optical technique, also because modern instruments are equipped with electronic recording devices like CCD plates [11], so that the microscope output is in digital form from the beginning. In this case the Fourier Transform of the hologram is taken and the sideband corresponding to the object wavefunction chosen. By shifting the origin in the Fourier space, we obtain the image wavefunction, from which the lens aberrations can be eliminated in principle by multiplication with the inverse of the aberration phase factor. Finally the inverse Fourier Transform gives the amplitude and phase of the true object wave whose phase can be displayed either directly or by adding digitally a plane wave, arriving thus at the same expression (40).

4. Numerical Simulations of Holographic Contour Maps

As shown in the foregoing section, once the phase shift $\phi(r)$ is calculated inserting into equation (31) the electrostatic potential given by equation (29), the contour map can be easily obtained using equation (40). By neglecting amplitude effects, i.e. $C''^2(r) = 1$, putting equal to zero unessential phase terms, i.e. $\phi_R = \phi_o = 0$, it turns out that the ideal contour map ($k_0 = 0$) is given by:

$$I(r) = 1 + \cos[\phi(r)]$$

(41)

whereas the real contour map, taking into account the perturbed reference wave, equation (38), is given by

$$I(r) = 1 + \cos[\Delta \phi(r)]$$

(42)
The following Figure 3 shows the ideal contour map due to the external field alone of our p-n junctions array, calculated for a pitch $b = 8 \text{ } \mu\text{m}$ and a potential $V_0 = 0.5 \text{ V}$ (a), and $V_0 = 1 \text{ V}$ (b), the potential difference being the double of this value. It should be remembered that the ideal contour map is the optical interferogram which would be obtained when the reference wave in the recording step is an unperturbed plane wave, and the optical wave in the recording step is plane and parallel to the object wave.

It can be seen that the trend of the contour fringes, giving the trend of the projected potential distribution, is in good agreement with what expected on the basis of naive assumptions about the potential distribution [2]. The fringes are running parallely to the junctions far from the specimen edge, and at the edge they fan out and connect neighboring regions.

However, it should be remarked that experimental results show unexpected features, like closed loops of the equiphaselines between the junctions [3]. This effect can be accounted for by considering the real contour map, equation (42), where the phase difference is inserted instead of the phase itself. Figure 4 shows the real contour maps obtained for different values of the interference distance $D$ whose direction has been taken parallel to the junctions. Figure 4a refers to $D = 4 \text{ } \mu\text{m}$; (b) to $D = 6 \text{ } \mu\text{m}$; (c) to $D = 8 \text{ } \mu\text{m}$ and (d) to $D = 10 \text{ } \mu\text{m}$; $V_0$ has been taken equal to 1 V and hence the potential difference is 2 V.

It can be seen that, as expected, the closed contour lines become increasingly elongated along the junctions as the interference distance $D$ increases, corresponding to a lessening of the effect of fringing field on the reference wave; it is also interesting to note that the radius of curvature of the loops is larger near the edge towards the vacuum region and smaller far from the edge within the specimen, as observed also experimentally [3].

Finally, the Figures refer to a very large area, having side $2b = 16 \text{ } \mu\text{m}$, whereas the actual reconstructed differential contour maps display only a strip, parallel to the biprism wire, a few microns in width. However, an image like those reported could be obtained by reconstructing several double exposure holograms and pasting them together, as recently demonstrated for the mapping of the electrostatic field around charged microtips [12].
Fig. 4. — Calculated phase difference contour maps for the case of $\Delta V = 2 \text{ V}$, for different values of the interference distance $D$: (a) $D = 4 \mu\text{m}$; (b) $D = 6 \mu\text{m}$; (c) $D = 8 \mu\text{m}$; (d) $D = 10 \mu\text{m}$.

5. Conclusions

The electrostatic problem giving the external potential due to an array of p-n junctions lying in a half-plane with the edge perpendicular to the junctions has been solved analytically and the solution for the case of abrupt distribution of the potential obtained. This potential distribution has been used for calculating the phase shift when the specimen is illuminated by high energy electrons, and the corresponding ideal and real contour maps obtainable by means of holography methods have been calculated and displayed. Even if the model does not take into account a realistic topography of the field across the junctions, nevertheless the obtained results show that the main features of the experimental data have been recovered, and that the unexpected loops of the experimental contour lines are interpreted as due to the effect of the perturbation of the reference wave from the fringing field.
It is intended to develop these calculations by considering the case of the edge tilted with respect to the junctions, i.e. a configuration better approaching the experimental set-up, and by inserting more realistic topographies for the electric field across the junctions. In fact, the main advantage of holographic methods is the ability to recover the phase of the reconstructed wave, and hence to obtain information regarding the trend of the field across and around the junctions, which could not be obtained by standard electron microscopy methods.

Acknowledgements

Extraordinary financial support from the University of Bologna for computing facilities is gratefully acknowledged. This work has been partially supported by funds from MURST coordinated by Consorzio INFM and CNR-GNSM. Useful discussions with Prof. G. F. Missiroli, Dr. G. Matteucci, Department of Physics, and Prof. M. Vanzi, Institute of Electrotechnics, University of Cagliari, as well as the skillful technical assistance of S. Patuelli are gratefully acknowledged.

References