The system, however, is under-determined, resulting in free parameters as, e.g. for $\arccos\{[|Q_{000}|^2+|Q_{\lambda 00}|^2-(|Q_{0\mu 0}|-|Q_{00\eta}|)^2]/2|Q_{000}Q_{\lambda 00}|\} \le \arg(Q\lambda 00) \le \arccos\{[|Q_{000}|^2+|Q_{\lambda 00}|^2-(|Q_{0\mu 0}|+|Q_{00\eta}|)^2]/2|Q_{000}Q_{\lambda 00}|\}$.

Using additional terms of (7) restricts the free parameters as to provide solutions probably unique as they have components that depend on the linearly independent terms $v_x$, $v_y$ and $v_z$.

## 4. Conclusions

The retrieval of the atomic displacements from a reconstructed electron wave function at the exit surface of an object results in the algebraic equation system (5) and the particular inverse problem (7) with the difficulties of finding the roots as discussed. The procedure described has thus transformed the difficulties of solving the direct scattering problem to the mathematical problem of determining the roots of a function with an incomplete Fourier transform. Furthermore, there are certain restrictions on the existence of solutions that would enable the construction of numerical algorithms as, e.g., generic ones. From the mathematical point of view the retrieval procedure is an ill-posed inverse problem requiring additional information about the unknown reconstructed displacements in order to make the process stable and continuous, to avoid singularities, and to restrict the manifold set of solutions possible. Open questions arise, e.g., with respect to the assumptions of cyclic boundary conditions, the applicability of the completeness relation to the backward iteration and to depths, where the equations for the displacement retrieval cannot be inverted because of singular coefficients.

## References

[1] K. Scheerschmidt, R. Hillebrand: Image interpretation in HREM: Direct and indirect methods, Proc. 32nd Course Int. Centre of Electron Microscopy "High-Resolution Electron Microscopy - Fundamentals and applications", Halle 1991, p. 56
[2] H. Lichte, Ultramicroscopy **20** (1986) 293
[3] W. Coene, G. Janssen, M. Op de Beeck, D. van Dyck, Phys. Rev. Letters **69** (1992) 3743
[4] D. van Dyck, W. Coene, Scanning Microsc., Suppl. **2** (1988) 131
[5] A.K.Head, Aust. J. Phys **22** (1969) 43
[6] K. Scheerschmidt, F. Knoll, ICEM-13, Paris 1994, Vol.1, p. 333; and Phys. Stat. Sol. (1994), in print
[7] H. Lichte, Advances in optical and electron microscopy **12** (1991) 25
[8] H. Lichte, E. Völkl, K. Scheerschmidt, Ultramicroscopy **47** (1992) 231
[9] D. van Dyck, Advances in optical and electron physics **65** (1985) 295
[10] G.R. Anstis, Computer Simulation of Electron Microscope Diffraction and Images, Eds.: W Krakow, M.O' Keefe, The Minerals, Metals & Materials Society, 1989, p. 229
[11] A. Howie, Z.S. Basinski, Philos. Mag. **17** (1962) 1039
[12] P. Stadelmann, Ultramicroscopy **21** (1987) 131
[13] B.N. Zakhariev, A.A. Suzko, Direct and Inverse Problems, Springer Vlg., Bln. 1990.
We are grateful to the Volkswagenstiftung for financial support.

Electron Holography
A. Tonomura, L.F. Allard, G. Pozzi, D.C. Joy and Y.A. Ono (Editors)



# RETRIEVAL OF ATOMIC DISPLACEMENTS FROM RECONSTRUCTED ELECTRON WAVES AS AN ILL-POSED INVERSE PROBLEM

Kurt Scheerschmidt and Frank Knoll

Max Planck Institute of Microstructure Physics, Weinberg 2, D-06120 Halle, Germany

The imaging of crystal defects by using high-resolution transmission electron microscopy or the electron diffraction contrast technique is well known and routinely applied. A direct and phenomenological analysis of electron micrographs, however, is mostly not possible, thus requiring the application of image simulation and matching techniques (see, e.g. [1]). On the other hand, electron holography and other wave reconstruction techniques allow one to directly determine the scattered wave function at the exit surface of an object up to the information limit of the electron microscope (see, e.g. [2,3]). Applying such a wave reconstruction should enable an object retrieval, i.e. the determination of the object potential or the positions of the atomic scattering centres directly from the wave function reconstructed instead of using trial-and-error simulation techniques. Up to now, direct solutions have been given for very thin objects (phase grating approximation, [4]) or for the assumption that the crystal potential of the perfect structure is known and solely the atomic displacements owing to a crystal lattice defect should be determined using the dependence of the three components in the case of plane strains or stresses [5]. Based on the knowledge of the reconstructed complex electron wave and using a discretized form of the diffraction equations, an alternative method is developed [6], which, in principle, enables the direct retrieval of the atomic displacements, caused by a crystal lattice defect, relative to the atom positions of the perfect lattice. A special inverse problem of electron scattering can be deduced considering solely those atomic displacements, which are given by the zeros of a function with an incompletely known Fourier spectrum. The fundamental relations are described, with the problems of solving the ill-posed Fourier transform being discussed.

## 1. Wave Reconstruction and Electron Diffraction

The electron microscope imaging is mainly determined by two processes: First, by the electron diffraction owing to the interaction process of the electron beam with the almost

periodic potential of the matter and, second, by the interference of the plane waves leaving the specimen and being transferred by the microscope. Images are modelled by calculating both processes, they are fitted to the experiment by varying the defect model and the free parameters. This trial-and-error image matching technique is the indirect solution to the scattering problem applied to analyse the defect nature under investigation.

The principles of image formation in the electron microscope have been well established: According to the lens aberrations and the microscope instabilities the higher spatial frequencies are transferred by alternating phase shifts and increased damping, respectively. In addition, owing to the recording of solely the image intensity modulus and phase of the electron wave are always mixed. Holography with electrons offers one of the possibilities of increasing the resolution by avoiding the microscope aberrations. It also enables the complete complex object wave to be restored. Image plane off-axis holograms are recorded in a microscope which is equipped with a Möllenstedt-type electron biprism inserted between the back focal plane and the intermediate image plane of the objective lens [2,7,8]. The object is arranged so that both the reference wave and the object wave are transferred through the microscope, and owing to a positive voltage of the biprism both waves mutually overlap in the image plane creating additional interference fringes. The intensity of the latter is modulated by the modulus of the object wave, whereas their position is varied by the phase of the object wave. Thus the recorded interference pattern is an electron hologram from which both modulus and phase of the object wave can be reconstructed by optical diffraction or numerical reconstruction. A Fourier transform of the intensity distribution of the hologram generates three distinct spectral patterns if the carrier frequency is sufficiently high. In the central region of the spectrum zero peak and autocorrelation occur, representing the conventional diffractogram of the object intensity, completely identical with that obtained from a corresponding HREM micrograph. The sidebands represent the Fourier spectrum of the complete complex image wave and its conjugate, respectively, from which the object wave o(x,y) can thus be reconstructed by separating, centring, and applying the inverse Fourier transform with the complex conjugate phase filter because of the always linear transfer to the sideband.

The interaction of electrons with a crystalline object is described assuming a periodic potential with the electron structure factors as the expansion coefficients and the Bloch-wave method of solving the high-energy transmision electron diffraction. Different formulations can be given, using Bloch wave or plane wave representations of the scattered waves, applying direct or reciprocal space expansion and using direct integration or slice techniques, which, in principle, are equivalent descriptions [9]. The object wave in terms of modified plane waves with complex amplitudes $\phi_g$ yields

$$o(\mathbf{R})=\Sigma_g \phi_g e^{2\pi i((\mathbf{k}+\mathbf{g})\mathbf{R}+s_g t)} \qquad (1)$$



with reflections **g**, excitations s, wave vector **k**, and thickness t of a parallel-sided object. Amplitudes $\phi_g$ are constant with respect to z in the vacuum outside the object, which means that the plane waves are the stationary solutions to the wave equation. Within the crystal, however, the amplitudes of the modified plane waves $\phi_g$ are z-dependent according to the Ewald pendel solution as described by the Bloch waves, which are the stationary solution to the periodic potential. Using furthermore the deformable ion approximation a crystal lattice defect can be included by its elastic displacement field as a phase shift of the Fourier spectrum of the crystal potential. The evaluation of the quantum-theoretical scattering problem by the high-energy forward scattering approximation (see, e.g. [10,11]) yields a parabolic differential equation system for the complex amplitudes of the elastically scattered electron waves:

$$\partial\phi_g/\partial z = \{ik_z\nabla^2 - 2(\mathbf{k}+\mathbf{g})\nabla\}\phi_g/2k'_z + i\sigma\Sigma_h V_{g-h}\phi_h e^{i\alpha_{gh}} \qquad (2)$$

where $\sigma=2\pi me/h^2k_zk'_z$, $\nabla=(\partial/\partial x,\partial/\partial y,0)$, $k'_z=k_z+g_z+s_g$ and $\alpha_{gh}=2\pi[(s_h-s_g)z+(\mathbf{g}-\mathbf{h})\mathbf{v}(x,y,z)]$ with the elastic displacement field **v** and the potential $V=V'+iV''$ including the lattice potential V' and the absorption V'' (one electron-optical potential approximation of inelastic scattering).

Boundary and initial conditions have to be applied, too: The linearized high-energy approximation directly fits $\phi_g(\mathbf{R},t)$ at the crystal exit face to $\phi_g(\mathbf{R})$ outside, demanding $|\phi_g(\mathbf{R},0)|=\delta_{g0}$ at the entrance face, whereas the continuity of the derivatives has to be omitted in the linearized case. Instead of the boundary conditions one can assume a periodic continuation for large extended crystal slabs, i.e. $\phi_g(x,y,z)=\phi_g(x+X,y,z)$ and $\phi_g(x,y,z)=\phi_g(x,y+Y,z)$, with slab extensions X,Y tending to infinity.

Fig. 1 shows the modulus (left) and phases (right) of the exit wave function in reciprocal- (Fig. 1a) and real- (Fig.1b) space representation, simulated for a spherical inclusion with the Ashby-Brown displacement field and linear displacements within the defect. The wave function is calculated using the EMS package for multi-slice simulations [12]; the data correspond to a 400 kV microscope and [011]-silicon of high symmetric incidence having a sample thickness of t=8.4nm.

Fig. 2 shows a calculated hologram (hol) assuming a perfect microscope without aberrations (a) and its Fourier transform to demonstrate the reconstruction of the wave function in reciprocal space. Fig 2. (b) demonstrates the intensity (diffraction pattern) and phase of the sideband selected. Furthermore, the corresponding modulus (mod) and phases (pha) of the complete reconstructed sideband are shown in (c), which should be equivalent to the exit wave function. The holograms are generated assuming a reference beam with a damping of 0.2 and a carrier frequency of 13.2 $nm^{-1}$ (i.e. located approximately at 1/10(-44,43,-43) in diffraction).

Fig. 3 shows the modulus (mod) and phases (pha) of the particular reflections selected of type 000, {200}, {022} and {111}, thus demonstrating the reconstruction of the corresponding amplitudes $\phi_g$ from the holograms. The reconstruction of the {400} reflections is impossible here because of the overlap of the autocorrelation and the sideband. Thus, aperture and damping are chosen to exclude the {113} reflections, which also omits the dumbbells in the HREM reconstruction resulting in differences between the original HREM images and the corresponding reconstructions.

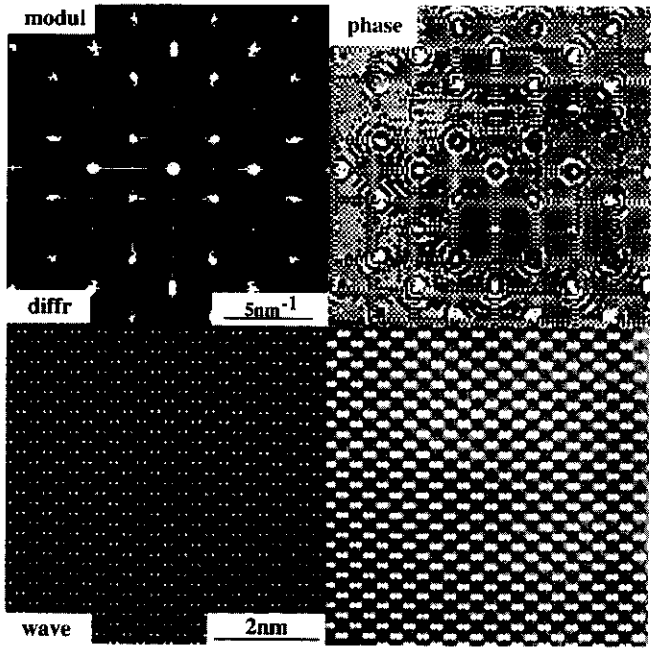


Fig. 1: Modulus (left) and phases (right) of the exit wave function simulated for a spherical inclusion: a) reciprocal-space (diffraction), b) real-space representation (U=400kV, sample thickness t=8.4nm, sphere radius $R_0$= 1.6nm (12,8√2,√2/2)-[011]-Si-supercell)

## 2. Forward-Backward Iteration

The differential equations (2) allow the diffusion-like interpretation and can be discretized using standard difference algorithms [14]. With the help of

$$V^2\phi=[\phi(x+\Delta x,y,z)-2\phi(x,y,z)+\phi(x-\Delta x,y,z)]/\Delta x^2+[\phi(x,y+\Delta y,z)-2\phi(x,y,z)+\phi(x,y-\Delta y,z)]/\Delta y^2$$
$$\partial\phi/\partial z = [\varepsilon(\phi(x,y,z+\Delta z)-\phi(x,y,z))+(1-\varepsilon)(\phi(x,y,z)-\phi(x,y,z-\Delta z))]/\Delta z \qquad (3)$$

an algebraic equation system results for the complex amplitudes and the elastic displacements at the (xyz) grid points (i,j,k), (i±1,j,k), (i,j±1,k), and (i,j,k±1). Using the abbreviations



$$A^{\pm}=\pm\sigma k_z z_0 l^2/(x_0^2K)+2\pi(k_x+g_x)(\pm\varepsilon_x-1/2),$$
$$B^{\pm}=\pm\sigma k_z z_0 l^2/(y_0^2K)+2\pi(k_y+g_y)(\pm\varepsilon_y-1/2), \qquad (4)$$
$$C=A^{-}-A^{+}+B^{-}-B^{+}+1/2-\varepsilon_z,\ \alpha(i,j,k)=2\pi[(s_h-s_g)kz_0/K+(\mathbf{g-h})\mathbf{v}(i,j,k)]$$

and denoting the maximum number of grid nodes in x,y, and z-direction by I,J,K yields

$$(1/2-\varepsilon_z)\phi_g(i,j,k-1)-(1/2+\varepsilon_z)\phi_g(i,j,k+1)=A^{+}\phi_g(i+1,j,k)-A^{-}\phi_g(i-1,j,k)$$
$$+B^{+}\phi_g(i,j+1,k)-B^{-}\phi_g(i,j-1,k)+C\phi_g(i,j,k)-i\sigma(z_0/K)\Sigma_h V_{g-h}\phi_h(i,j,k)e^{i\alpha(i,j,k)} \qquad (5)$$

which is equivalent to forward (k+1) and backward (k-1) integration with respect to the beam propagation, i.e. to $\varepsilon_z$= 1/2 or -1/2, respectively.

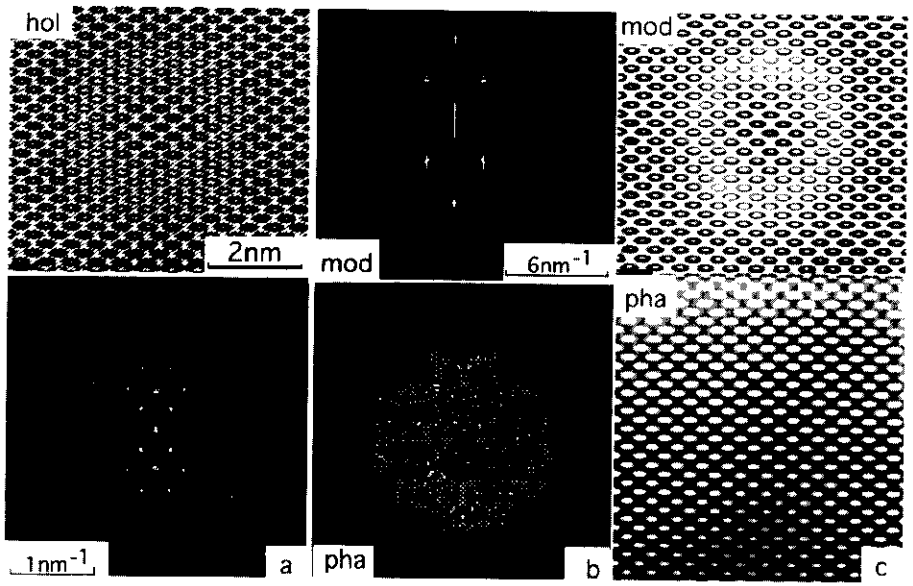


Fig.2: Calculated hologram (hol) and its Fourier transform (a; diffraction pattern with autocorrelation and sidebands) assuming a perfect microscope without aberrations for a spherical inclusion (U=400kV, $\alpha$=20nm$^{-1}$, (12,8√2,√2/2)-[011]-Si-supercell, t=8.4nm, $R_0$= 1.6nm), the selected sideband with intensity and phase (b), and its complete reconstruction (c) with modulus (mod) and phases (pha) representing the perfect exit wave function (reconstructed HREM-image). The hologram is generated assuming a reference beam with a damping of 0.2 and a carrier frequency of 13.2 nm$^{-1}$.

The periodic boundary conditions and the initial conditions may simply be written $\phi_g(i,j,k)=\phi_g(i+I,j,k)$, $\phi_g(i,j,k)=\phi_g(i,j+J,k)$, and $|\phi_g(i,j,0)|=\delta_{0g}$, $\phi_g(i,j,t)=F_g(i,j)$, respectively, with $F_g$ being known from the wave reconstruction for a certain number of reflections.

The difference equations (5) are equivalent for backward (k-1) and forward (k+1) integration, thus being insufficient for determining both the wave amplitudes $\phi(i,j,k)$ and the elastic displacement field $\mathbf{v}(i,j,k)$ at the grid points (i,j,k) considered. One of the difference equations, however, can be replaced as follows: While the optical potential in reciprocal space representation is generally non-Hermitian, the hermiticity of potential V' and of „absorption" V" yields the equation of continuity for the whole current $I=\phi_g\phi^*_g$. The continuity equation can be written as

$$\partial I/\partial z = \vartheta I - 2\Sigma_g V''_{gh}\phi_g\phi^*_h e^{i\alpha_{gh}} \qquad (6)$$

yielding the abbreviation $\vartheta I=\Sigma_g[k_z(\phi_g\nabla^2\phi^*_g-\phi^*_g\nabla^2\phi_g)+2(\mathbf{k}+\mathbf{g})\nabla(\phi_g\phi^*_g)]/k'_z$.

The equation of continuity can be discretized by analogy with the discretization of the differential equations above. The differential operator $\vartheta$, however, yields mixed terms with respect to different nodes (i,j,k) and (i±1,j±1,k). By analogy with the Gelfand-Levitan-Algorithm (see, e.g., [13]) an additional equation results, which is a kind of completeness relation, yielding

$$\Sigma_{\mathbf{g}} Q_{\mathbf{g}} e^{2\pi i \mathbf{g}\mathbf{v}} = 0. \qquad (7)$$

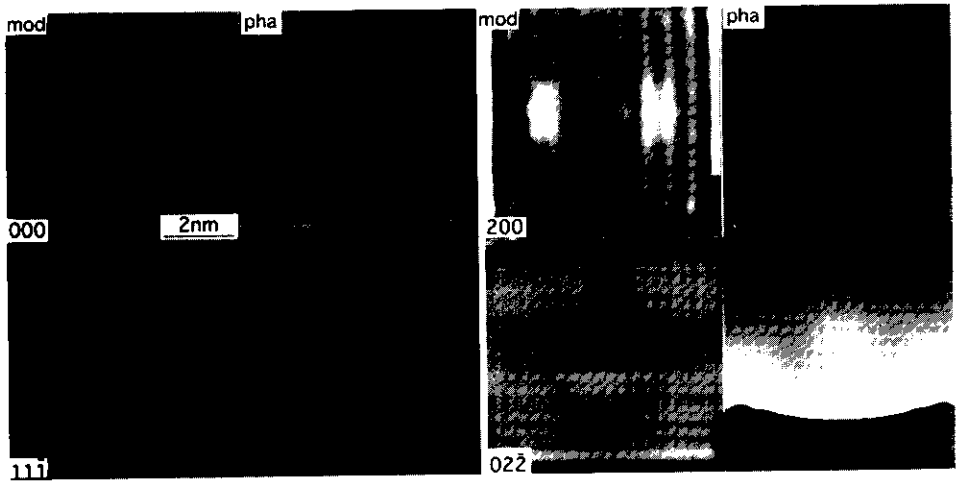


Fig. 3 Modulus (mod) and phases (pha) of the single reflections 000, {200}, {220}, and {111}, i.e. the corresponding plane wave amplitudes, reconstructed from sideband (b) of the hologram (a) of Fig. 2 by filtering and centring of the corresponding reflections.

The coefficients $Q_{\mathbf{g}}=\Sigma_h k_z/k'_z V''_{g,g-h}\phi_g\phi^*_{g-h}\exp[2\pi i(s_g-s_{g-h})z]$ for $\mathbf{g}\neq(000)$, and correspondingly $Q_{000}=-\vartheta I$ are given for the nodes (i,j,k) from eqs. (6) in forward scattering.



Eq.(7) can replace one of eqs.(5) in backward integration enabling the determination of displacements $\mathbf{v}$ at (i,j,k) by inverting the equation of continuity as an independent additional equation. Thus, in principle, the retrieval of the displacement is given by the remaining inverse problem (7), which is the same as to find the root of a function given by an incomplete Fourier transform.

At the exit surface a further equation is given applying the forward integration outside the crystal to determine $\phi(i,j,K+1)$ from $\phi(i,j,K)$ where the potential is assumed to be vanishing because of the vacuum propagation. The backward integration, however, using eq. (6) then enables the determination of $\mathbf{v}(i,j,K)$ at the exit surface.

## 3. The Remaining Inverse Problem

The inverse problem (7) is ill-posed for two reasons: Only one equation has to be solved for the vectorial root $\mathbf{v}(i,j,k)$ at node (i,j,k), which describes three unknown quantities by two conditions, and spectrum $Q_{\mathbf{g}}(i,j,k)$ is incomplete and noisy. This results in unstable numerical solutions using standard algorithms to find the roots, owing to the existence of a large number of subsidiary roots. Different algorithms are tested, viz. the Newton-Raphson algorithm itself to solve eq. (7), and of transform eq.(7) in an iterative form as a kind of quasi-regularization, e.g. using relations for the arguments yielding

$$v_x^{n+1}=1/2\pi\{\arg[Q\exp(2\pi i v_x^n)+\Sigma_{\mathbf{g}}Q_{\mathbf{g}}\exp(2\pi i\mathbf{g}\mathbf{v}^n)]-\arg[Q]\} \qquad (8)$$

and similarly for $u_y$, $u_z$. Both algorithms demand the iteration for linear independent coefficients $\mathbf{g}$, thus coplanar vectors $\mathbf{g}$ leave one component unconsidered.

Analytical solutions of eq.(7) can be performed if four terms at a maximum are considered, which are revealed most easily by interpreting equation (7) in the complex plane as the summation of rotating vectors $Q_{\mathbf{g}}$ as a function of root $\mathbf{v}(i,j,k)$ at node (i,j,k).

For non-vanishing $Q_{000}$, $Q_{\lambda 00}$, for instance, the system is over-determined demanding $|Q_{000}|=|Q_{\lambda 00}|$ and resulting in $v_x=1/2\pi\lambda\{\arg(Q_{000})-\arg(Q_{\lambda 00})+(2k+1)\pi\}$.

Otherwise, for non-vanishing $Q_{000}$, $Q_{\lambda 00}$, $Q_{0\mu 0}$, and with the other coefficients being neglected, one yields

$$v_x = 1/2\pi\lambda\{\arg(Q_{000})-\arg(Q_{\lambda 00})+\arccos[(|Q_{\lambda 00}|^2-|Q_{0\mu 0}|^2-|Q_{000}|^2)/2|Q_{000}Q_{\lambda 00}|]\} \qquad (9)$$

similarly also for $u_y$, with $u_z$ being arbitrary, however. Thus, considering solely three terms provides an exact solution to the problem if $||Q_{0\mu 0}|-|Q_{\lambda 00}||\leq|Q_{000}|\leq|Q_{0\mu 0}|+|Q_{\lambda 00}|$.

The solution based on four non-vanishing terms, in principle has the same structure, enabling the determination of the component $\mathbf{v}_z$, too, but having to fulfil further restrictions.