Effect of dimensionality on the charge-density-wave in few-layers 2H-NbSe₂

Matteo Calandra¹, I. I. Mazin², and Francesco Mauri¹

¹CNRS and Institut de Minéralogie et de Physique des Milieux condensés, case 115, 4 place Jussieu, 75252, Paris cedex 05, France and
²Naval Research Laboratory, 4555 Overlook Ave. SW, Washington, DC 20375

We investigate the charge density wave (CDW) instability in single and double layers, as well as in the bulk 2H-NbSe₂. We demonstrate that the density functional theory correctly describes the metallic CDW state in the bulk 2H-NbSe₂. We predict that both mono- and bilayer NbSe₂ undergo a CDW instability. However, while in the bulk the instability occurs at a momentum \( q_{\text{CDW}} \approx \frac{4}{3} \Gamma M \), in free-standing layers it occurs at \( q_{\text{CDW}} \approx \frac{4}{5} \Gamma M \). Furthermore, while in the bulk the CDW leads to a metallic state, in a monolayer the ground state becomes semimetallic, in agreement with recent experimental data. We elucidate the key role that an enhancement of the electron-phonon matrix element at \( q \approx q_{\text{CDW}} \) plays in forming the CDW ground state.

PACS numbers: 74.70.Ad, 74.25.Kc, 74.25.Jb, 71.15.Mb

Charge density wave (CDW) is one of the most common and most intriguing phenomena in solid state physics. The concept originated in the seminal paper by Peierls[2] pointing out a divergence in the 1D response functions at a wave vector equal to twice the Fermi vector. Later this concept was generalized onto the 2D and 3D system with “nesting”, that is, quasi-1D portions of the Fermi surface. On the other hand, the fact that the Peierls instability is logarithmic, and therefore very fragile, has led to the question of whether the actual CDW observed in quasi-2D materials are indeed manifestations of the Peierls instability[3].

In the last decades the role of nesting in various structural and magnetic instabilities has been wildly debated, particularly with respect to the most venerable CDW materials, NbSe₂ and related dichalcogenides. It has been pointed out that the bare susceptibility does not have any sharp peak at the CDW wave vector, \( q_{\text{CDW}} = (2\pi/3, 2\pi/3) \), but at best a broad and shallow peak[3, 4, 5] in the real part, while the imaginary part (which directly reflects the Fermi surface nesting) does not peak at \( q_{\text{CDW}} \) at all[3, 4, 5].

This suggests that momentum dependence of the electron-phonon interaction plays a crucial role in driving CDW instabilities[3, 4, 5]. An indirect confirmation comes from neutron[8] and X-ray scattering[6] phonon-dispersion data where a marked softening was detected in only one of the low energy modes close to \( q_{\text{CDW}} \).

A very clear test for the described hypothesis are purely 2D systems — mono- and bilayers of the same compounds. Lacking \( k_z \) dispersion, they should be much more sensitive to the Fermi surface nesting than their 3D counterparts. 2H-NbSe₂ is particularly promising as its 2D structures have been synthesized, and found to show a curious semimetallic behavior[10]. Interestingly the conductivity of a NbSe₂ layer in a field-effect transistor as a function of the gate voltage (\( V_g \)), which controls the doping, is very different from that of good metallic monolayers. While typical metals have a conductivity rather weakly dependent on doping, NbSe₂ monolayers show a strong monotonic dependence on \( V_g \). Such a large variation (more then a factor of two in the -70 to +70 V \( V_g \)-range) is characteristic of semiconductors, semimetals as graphene, or metals with a pseudogap at the Fermi level. The first principles calculations predict a good metal for a monolayer in the undistorted bulk structure[11], in disagreement with experiments, which suggest a possibility of a semimetallic band structure.

In this Letter we study the CDW instability in the bulk 2H-NbSe₂, in the bilayer and in the monolayer NbSe₂, both by calculating the phonon spectra in the high-symmetry phase and by full structural optimization[12]. We find that all three studied structures are unstable against a CDW formation with, however, different patterns, despite the extreme similarity of the Fermi surfaces. We then demonstrate that the crucial difference comes from the \( q \)-dependent electron-phonon matrix elements, rather than directly from the Fermi surface (FS). Finally we show that the CDW distorted monolayer is semimetallic and the large variation of its conductivity as a function of \( V_g \) detected in experiments is a manifestation of the CDW.

We simulate a single layer or a bilayer, as usually, by introducing a thick vacuum layer (of the order of 10 Å). We then performed the geometrical optimization in one unit cell, without a CDW. We found that the distance between the Nb and Se planes barely changes: from 1.67 Å in the bulk to 1.69 Å in the bilayer (same in the monolayer). On the contrary, the Nb-Nb interplanar distance in the bilayer is significantly enlarged compared to the bulk (6.27 Å vs. 6.78 Å).

The pseudopotential electronic structure of the NbSe₂ bilayers and monolayers is compared with that of the bulk NbSe₂ in Fig. 4 (LAPW calculations yield very similar results). For the bulk case, as known from previous calculations[5, 15], three bands cross the Fermi level...
The calculated FS for a single layer (Fig. 2) consists of a 2D, that is, seemingly more prone to the nesting effects. Four Brillouin zones are shown, Γ points are at the grid nodes. Left: Real part of the bare static electronic susceptibility (in arbitrary units) with constant matrix elements, same as the real part of the phonon self-energy $\chi''(q, \omega = 0)$. Eq. (1), calculated using the LAPW band structure of a single layer NbSe$_2$. Center: The corresponding FS. Four Brillouin zones are shown, Γ points are at the grid nodes. Right: The same, for the resulting 4×1 CDW.

(\epsilon_f) One is the antibonding Se $p_z$ bands, which is highly dispersive along the z-axis and does not cross $\epsilon_f$ in the A-L-H plane. The other two are formed by Nb d-states. These two bands are degenerate in the A-L-H plane in the scalar relativistic calculations (spin-orbit interaction removes the degeneracy) while the degeneracy is lifted for other values of $k_z$ \cite{5} by the interlayer interaction. As a result, while one of the FSs originates from Nb d-bands is two dimensional, the other acquires considerable $k_z$ dispersion, so that a model neglecting hopping between the layers does not apply \cite{16}.

In a bilayer, the Se-Se interplanar interaction is reduced, while in the monolayer it is absent. As a result, the Se band sinks below the Fermi level, and only Nb the derived Fermi surfaces are present, and they are strictly 2D, that is, seemingly more prone to the nesting effects. The calculated FS for a single layer (Fig. 2) consists of a rounded hexagon at Γ and two rounded triangles at K, just as the Nb FS pockets in the bulk 2H-NbSe$_2$ \cite{5}.

Next, we have computed the phonon dispersions for the bulk and single-layer NbSe$_2$. In Fig. 3 we plot the phonon dispersions in the monolayer (bulk) calculated using a 40 × 40 (20 × 20 × 6) k-points grid and Fermi temperature $T = 68$ meV ($\tau = 270$ meV) with Hermite-Gaussian distributions (See Ref. [17] for more details.). In the bulk, the highest energy acoustic mode is unstable at approximately $q_{CDW} = 2/3 \Gamma M$, in agreement with the experiment. These calculations implicitly include the renormalization of the bare phonon frequencies due to the electron-phonon interaction, as described by the real part of the phonon self-energy of a phonon mode $\mu$

$$\Pi'_{\tau,\mu}(q, \omega = 0) \sim \sum_{k, nm} \left| g_{k_i, k+q,j}^{\tau,\mu} \right|^2 \left( f_{k+q,j} - f_{k_i} \right)^2 \epsilon_{k+q,j} - \epsilon_{k,i}, \quad (1)$$

where $\epsilon_{k_i}$ is the one-electron energy, $g_{k_i, k+q,j}^{\tau,\mu}$ is the electron-phonon matrix element and $f_{k_i}$ is the Fermi function. If one neglects the $k$ dependence of the matrix elements in Eq. 1, this expression becomes proportional to the real part of the unrenormalized static electronic susceptibility $\chi''(q)$, as shown in Ref. [5], in NbSe$_2$ $\chi''(q)$ has a broad maximum at $q_{CDW} = 2/3 \Gamma M$, even though the imaginary part (which directly reflects the FS nesting) does not.

Thus, DFT correctly describes the CDW instability in the bulk NbSe$_2$. With this in mind, we move to a NbSe$_2$ monolayer and observe that (a) now the most unstable phonon mode along $\Gamma M$ appears near $(1/2)\Gamma M$ and (b) the instability expands over a substantially larger k-point region, as phonon are unstable also along MK. We also looked at the bare susceptibility $\chi''_0$ and found a strip-like region of enhanced values, extending from 0.4$\Gamma M$ to 0.75$\Gamma M$, with a width of approximately 0.15 of the $\Gamma M$ vector (0.15$\pi/a$). This is in agreement with the instabilities found in the linear-response phonon calculations along $\Gamma M$ and MK.

To understand which electronic states are involved in the CDW formation we have repeated the calculations with an increased electronic temperature $\tau$. We found that the phonon frequency of the highest acoustic mode strongly depends on $\tau$ for $\tau > 0.3$ eV. The instability occurs for $\tau \approx 0.3$ eV, and for smaller temperatures the (imaginary) phonon frequency of the unstable mode is essentially constant. This means that electronic states within a 0.3 eV window around $\epsilon_f$ are responsible for the CDW instability, which is not a Fermi surface effect \cite{4,18,19,20}, but involves states in the full Nb-derived d-band (which is only 1 eV wide, see Fig. 4). As discussed above, $\chi''_0$ does have a maximum at $q_{CDW}$, but this maximum is relatively weak and broad, and, most importantly, weakly dependent on $\tau$. Thus, the CDW is mainly the result of an enhancement of the electron-phonon matrix element close to $q_{CDW}$. The self-consistent screening of the electron-phonon matrix element as a function of $\tau$ is crucial to describe the CDW.

The next test is to take a superstructure suggested by
the calculated CDW vector \([(\pi/3,0)\) or \((\pi/4,0)\)] and to optimize the cell dimensions and the atomic positions. Each time we started from slightly randomized atomic positions and minimized the total energy. We found that in the bulk the 4×1 and 3×1 supercells both converge to a lower energy than the undistorted cell, consistent with the calculated phonon spectra in Fig. 3; the 3×1 being lower in energy (with the energy gain of 0.1 mRyd/Nb with respect to the undistorted structure). The other supercells converged to the undistorted structure. The distortion of the 3×1 supercell is consistent with the Σ₁ phonon pattern detected in inelastic neutron scattering [8], involving not only an in-plane deformation of the layered structure, but also out-of-plane displacements of the Se atoms and a charge transfer between the two layers.

For a monolayer, the situation is different. The lowest energy supercell is now 4×1 with a 1.19 mRyd/Nb energy gain with respect to the undistorted cell, followed by 3×1. We also checked the 2×1 and 2×2 supercells, which converged to the undistorted structure, and the 3×3 one, which converged to the same structure as 3×1. This is again consistent with the calculated phonon spectra. In the ground state Nb atoms are trimerized, qualitatively different from the bulk CDW. Finally, the bilayer behaves just like a monolayer, with a 4×1 ground state similar to that found for a monolayer.

Having obtained the distorted structure both for the bulk and for a monolayer, we investigated their electronic properties. We found (Fig. 4) that in the bulk case the density of states (DOS) at the Fermi level in the distorted 3×1 supercell is not much reduced in the CDW, contrary to what one would expect in the Peierls model, and reconfirming that the energy gain in this case is not coming from the states in the immediate vicinity of the Fermi level [5] but from the occupied Nb bands. In the monolayer 4×1 supercell the Peierls model is again violated due to the important role of the electron-phonon matrix element, however the FS is strongly gapped (see Fig. 2 right) and displays a clear semimetallic behavior in qualitative agreement with experimental data [10].

More insight on the conducting nature of the monolayer is obtained from the conductivity per layer, namely

\[
\sigma^2_{\text{D}}(E) = \frac{e^2}{3\Omega_{2D}N_k} \sum_{k'x} v_{k'x}^2 \tau(k) \delta(\epsilon_{k'} - E),
\]

where \(\Omega_{2D}\) is the area of the two dimensional unit cell of a NbSe₂ layer, \(e\) is the electron charge, \(N_k\) the number of k-points used in the calculation, \(v_{k'x}\) the electron velocity, \(E\) the energy. In general \(\sigma^2_{\text{D}}(E) = \sigma^2_{\text{D}}(E) \times L\), where \(L = 6.275\)Å is the layer thickness. Assuming, \(\tau(k) \approx \tau(\epsilon_f)\), we have that \(\sigma^2_{\text{D}}(E) = \tau(\epsilon_f) \omega_{p,ab}^2 / 4\pi\).
FIG. 5: (color online): LAPW plasma frequency and conductivity as a function of the gate-induced charge for a single-layer NbSe$_2$. The constant relaxation time approximation is adopted. Experimental data [10] are plotted assuming the charge neutrality at $\approx 30 \times 10^{12}$ cm$^{-2}$. The gate voltage for the standard configuration of a 300 nm SiO$_2$ field-effect-transistor can be obtained from the $x$ axis dividing 7.2 $\times$ 10$^{-2}$ cm$^{-2}$ V$^{-1}$. The charge is obtained from the integrated DOS.

The results of the $\omega_{p,ab}^2$ calculation using the optics code of the WIEN2k package are plotted in Fig. 5 for the monolayer NbSe$_2$ as a function of the gate-induced charge as obtained from the calculated DOS (Fig. 4). For the undistorted structures, $\omega_{p,ab}$ is essentially constant, in disagreement with the experiment. On the contrary, in the CDW state the monolayer plasma frequency shows substantial variation near $\epsilon_F$, due of its semimetallic structure. A pseudogap occurs just above $\epsilon_F$, so that the conductivity grows with the gate voltage. In the constant relaxation-time approximation, and assuming an electron self-doping of $\approx 30 \times 10^{12}$ cm$^{-2}$ we obtain reasonable agreement with experimental conductivity as function of the gate voltage [10] (see Fig. 5). For a quantitative comparison with experiments, however, one would need to know the energy and momentum dependence of the relaxation time.

To summarize, we have shown that the density functional theory accurately describes the CDW instability in the bulk 2H-NbSe$_2$, and predicts a similar instability to occur in mono and bilayer. However, while in the bulk the CDW occurs at $q_{CDW} \approx (\frac{1}{3}, 0, 0) \frac{2\pi}{a}$, in the monolayer and bilayer the CDW vector is $\approx (\frac{1}{3}, 0, 0) \frac{2\pi}{a}$. In all these systems the CDW is driven by an enhancement of the electron-phonon coupling at $q \approx q_{CDW}$. Our work solves the long standing controversy on the origin of CDW in 2H-NbSe$_2$ and in transition metal dichalcogenides.

Unlike the bulk case, when the system remains a good metal in the CDW state, in a monolayer CDW produces a semimetallic state. This leads to a large variations of the conductivity as a function of the gate voltage, contrary to the undistorted structure and in agreement with the experiment data [10]. Our calculation shows that a proper description of the CDW instability in reduced dimension is mandatory to interpret conduction data for low-dimensional transition-metal-dichalcogenides and opens the way to theoretical understanding of CDW-based field-effect-devices.

Calculations were performed at the IDRIS supercomputing center (project 081202).

[1] G. Grün, Density Waves in Solids (Addison-Wesley, Reading, PA, 1994).
[2] R. E. Peierls, Quantum theory of solids, Clarendon, Oxford (1955).
[3] M. D. Johannes and I. I. Mazin, Phys. Rev. B 77, 165135 (2008).
[4] N. J. Doran, B. Ricco, M. Schreiber, D. Titterington and G. Wexler, J. Phys. C, 11 699 (1978).
[5] M. D. Johannes, I. I. Mazin and C. A. Howells, Phys. Rev. B 73, 205102 (2006).
[6] M. H. Whangbo and E. Canadell, J. Am. Chem. Soc. 114, 9587 (1992).
[7] N. J. Doran, J. Phys. C, 11, L059 (1978).
[8] D. E. Moncton, J. D. Axe and F. J. Salvo Phys. Rev. Lett. 34, 734 (1975) and Phys. Rev. B 16, 801 (1977).
[9] B. M. Murphy et al. Phys. Rev. Lett. 95, 256104 (2005), J. Phys. Cond. Matt. 20, 224001 (2008).
[10] K. S. Novoselov et al., PNAS 102, 10451 (2005).
[11] Lébègue and O. Eriksson, Phys. Rev. B 79, 115409 (2009).
[12] Calculations were performed using the QUANTUM-ESPRESSO [15] pseudopotential code and the WIEN2k [14] all-electron LAPW package with the generalized gradient approximation. For Nb (Se) we use ultrasoft [21] (norm-conserving [22]) pseudopotentials including semicore states as valence.
[13] P. Giannozzi et al., J. Phys. Cond. Matt. 21, 395502 (2009), http://www.quantum-espresso.org.
[14] P. Blaha, K. Schwarz, G. K. H. Madsen, D. Kvasnicka and J. Luitz, WIEN2k, An Augmented Plane Wave + Local Orbitals Program for Calculating Crystal Properties, Karlheinz Schwarz, Techn. Universität Wien, Austria, 2001.
[15] L. F. Mattheiss, Phys. Rev. Lett. 30, 784 (1973), Phys. Rev B 8, 3719 (1973).
[16] D. S. Inosov et al., New J. of Phys. 10, 125027 (2008).
[17] M. Calandra and F. Mauri, Phys. Rev. Lett. 95, 237002 (2005).
[18] T. M. Rice and G. K. Scott, Phys. Rev. Lett., 35, 120 (1975).
[19] J. A. Wilson , Phys. Rev. B, 15, 5748 (1997).
[20] Th. Straub, et al, Phys. Rev. Lett. 82, 4504 (1999).
[21] D. Vanderbilt, PRB 41, 7892 (1990).
[22] N. Troullier and J. L. Martins, Phys. Rev. B 43, 1993 (1991). 32, 2217 (1971)