Fermi surface of layered compounds and bulk charge density wave systems

F Clerc¹, C Battaglia¹, H Cercellier¹, C Monney¹, H Berger², L Despont¹, M G Garnier¹ and P Aebi¹

¹ Institut de Physique, Université de Neuchâtel, CH-2000 Neuchâtel, Switzerland
² Institut de Physique de la Matière Complex, EPFL, CH-1015 Lausanne, Switzerland

E-mail: florian.clerc@unine.ch

Abstract
A review is given of recent angle-resolved photoemission (ARPES) experiments and analyses on a series of layered charge density wave materials. Important aspects of ARPES are recalled in view of its capability for bulk band, Fermi surface and spectral function mapping despite its surface sensitivity. Discussed are TaS₂, TaSe₂, NbTe₂, TiSe₂ and TiTe₂ with structures related to the so-called 1T polytype. Many of them undergo charge density wave transitions or exist with a distorted lattice structure. Attempts to explain the mechanism behind the structural reconstruction are given. Depending on the filling of the lowest occupied band a drastically different behaviour is observed. Whereas density functional calculations of the electronic energy and momentum distribution reproduce well the experimental spectral weight distribution at the Fermi energy, the ARPES energy distribution curves reveal that for some of the compounds the Fermi surface is pseudo-gapped. Two different explanations are given, the first based on density functional calculations accounting for the charge-density-wave-induced lattice distortion and the second relying on many-body physics and polaron formation. Qualitatively, both describe the observations well. However, in the future, in order to be selective, quantitative modelling will be necessary, including the photoemission matrix elements.

1. Introduction

1.1. Context

This review is written in the context of surface methods, i.e. angle-resolved photoemission (ARPES), used to map Fermi surfaces, and its aim is to illustrate and discuss the different issues involved within this approach when applying it to layered compounds and in particular to charge density wave (CDW) systems.
From the experimental point of view, it is important to distinguish between surface versus bulk properties probed by the experiments, and to study implications for the experiment if the electronic structure has a two-dimensional (2D) or a three-dimensional (3D) character. Furthermore, why are we interested in mapping the Fermi surface (FS), and are we really measuring the FS?

In fact, we would like to learn about the low energy excitations of electrons, what interactions govern the properties of the system and whether we have quasiparticles. The topology of the FS gives an indication on the possible scattering channels of quasiparticles and is therefore an important ingredient for understanding the electronic properties. Naturally, electronic and lattice degrees of freedom are not independent, an aspect particularly interesting in CDW compounds where atom positions are modified, thereby lowering the electron energy.

The interplay between lattice and electronic degrees of freedom has received renewed interest in the context of high temperature superconductivity and colossal magnetoresistance materials where electron–phonon coupling and possible polaronic effects are considered [1, 2]. A significant contribution to the discussion is given by ARPES experiments via analysis of the spectral function. Indeed, the one-electron removal spectral function is one of the ingredients of the photoemission signal, containing the many-body physics of the system.

The interpretation of the photoemission signal is a challenge, trying to extract information on the quasiparticles, their energy dispersion, self-energy and spectral function. The final goal is to be able to obtain everything ranging from the electronic band structure and the FS to the many-body physics and the excitation spectrum of electrons. The CDW materials present several interesting aspects in this context. First, electron–phonon coupling plays a role by definition and therefore these materials are a playground to study the interplay between the atomic and electronic structure, i.e. what is the strategy of the material to optimize its energy, thereby creating the CDW, and second, unconventional spectral features discovered in the most important material’s class of high critical temperature ($T_c$) superconductors can be compared to the ones obtained in this more conventional class of CDW compounds.

1.2. Transition metal dichalcogenides

In the present article we consider a series of transition metal dichalcogenides (TMDCs). We discuss TaS$_2$, TaSe$_2$, TaTe$_2$, NbTe$_2$, TiSe$_2$ and TiTe$_2$, all in a structural form related to the so-called $1T$ polytype crystal structure (figure 1). Layered TMDCs are a class of quasi-2D materials. Many exhibit CDWs and even become superconducting at low temperatures. In the ionic picture the transition metal atom (M) has a 5d$^1$6s$^2$ (Ta), 4d$^5$5s$^1$ (Nb) or 3d$^2$4s$^2$ (Ti) configuration. The chalcogen atoms (X), S, Se and Te, with high electron affinity, have all an
s^2p^6 configuration, missing two electrons for a filled shell. Therefore, the two chalcogen atoms per formula unit may attract a total of four transition metal d or s electrons and the degree of filling of the lowest lying d band will delicately influence the properties of the compound [3]. Nominally, Ta and Nb related compounds will be in a d^1 configuration, whereas the ones with Ti will remain with zero d electrons (d^0).

Of course, hybridization will modify this simplified view, but it is clear that the material’s properties will be highly sensitive to the number of d electrons. The lower the number of electrons at the Fermi level (E_F), the weaker the screening and the higher the tendency to instabilities. In this regard it seems likely that the heaviest metal (smallest electronegativity) together with the lightest of the chalcogen atoms (largest electronegativity), i.e. TaS_2, gives the most unstable combination and, on the other hand, TiTe_2 the most stable one. Indeed, TaS_2 has the most complicated phase diagram and TiTe_2 is a prototype for a 2D Fermi liquid.

Band-structure calculations (figure 2(a)) confirm that the six lowest lying bands are mainly chalcogen related sp bands followed by a partially filled transition metal d band (downwards pointing arrows) and the rest of the metal d bands. The bulk Brillouin zone (BZ) of the threefold symmetric 1T structure as well as the surface BZ (SBZ) with corresponding high symmetry points is indicated in figure 2(b). Due to the octahedral coordination of the metal atoms (figure 2(c)), the five d bands are split into a lower triplet of t_{2g} states and an upper doublet of e_{g} states (figures 2(a) and (d)). The σ bonding e_{g} orbitals have higher energies because they interact strongly with the neighbouring chalcogen atoms. The orbital degeneracy of the octahedral t_{2g} manifold is reduced in a Jahn–Teller-like fashion by a trigonal elongation of the chalcogen octahedra along the c axis. This elongation influences the overlap of the metal d band with the highest chalcogen sp bands (figure 2(a), arrows pointing upwards) and is characterized by the so-called z parameter as indicated in figure 1(a). A larger z parameter indicates a larger metal–chalcogen distance and tends to lower hybridization.

Among the 1T compounds studied, TaS_2 and TaSe_2 are similar since they both exhibit an incommensurate (IC) CDW with (√3 × √3) symmetry at high temperature. At low temperature both compounds lock into the commensurate CDW with (√3 × √3)-R 13.9°
symmetry. 1T-TaS$_2$ is more complicated, additionally exhibiting a quasi-commensurate (QC) phase at intermediate temperature. TaTe$_2$ and NbTe$_2$ both stabilize in a monoclinically distorted version of the trigonal 1T polytype and can be interpreted as a (3 × 1) reconstruction [4, 5]. Unlike TaTe$_2$, NbTe$_2$ becomes superconducting at ≈0.5 K [6]. After cooling of heat pulsed crystals to room temperature, transmission electron diffraction experiments reveal a second, CDW state with a (√19 × √19) signature [7–9], commensurate at room temperature, but readily rendered incommensurate just above. This observation is interesting since it links to the (√T3 × √T3) reconstruction of TaS$_2$ and TaSe$_2$. 

1T-TiSe$_2$ is different in the sense that the lattice is not distorted at room temperature. A (2×2×2) reconstruction appears at ≈200 K together with a maximum in the resistivity [10, 11]. 1T-TiTe$_2$ does not show any phase transition, remaining in its trigonal high symmetry phase, and is considered to be a prototype of a 2D Fermi liquid.

To get an overview on the different compounds we can classify them into two categories, the Ta (including Nb) compounds with d$^1$ configuration and the Ti related compounds with d$^0$ configuration. ARPES results within a given category are rather similar. Numerous studies of the electronic properties have been carried out for TMDCs [12–23]. For practical reasons, in this review article, results will be shown only for some of the compounds, but the discussion will involve all of them.

2. Aspects of photoemission

Even for quasi-2D materials such as the layered TMDCs, the photoemission experiment proceeds in the real 3D world and attention has to be paid to the fact that the escape depth of electrons is small. The result is that ARPES is very surface sensitive unless extremely high (of the order of a few keV) or low (5–10 eV, e.g. lasers) photon energies are used (i.e., the mean free path of electrons increases significantly at high and low electron kinetic energies). Therefore, experiments such as the ones presented here effectively probe the surface electronic structure. Depending on the material and the presence of surface relaxation and/or reconstruction of atomic positions, so-called surface states may exist besides the bulk states. While surface states are interesting by themselves, they are not relevant for the bulk physics we are interested in here. Surface states (perfectly 2D objects) may exist in a region of wavevectors parallel to the surface ($\vec{k}_1$ region) and an energy interval where there are no bulk states available for all wavevectors perpendicular to the surface ($\vec{k}_1$), i.e. where the bulk band structure is gapped. Furthermore, surface states are very sensitive to surface modification, e.g. due to adsorbates or contamination, and do not show any dispersion as a function of $\vec{k}_1$. From these criteria we did not find indications for surface states in the experiments presented here.

Another important aspect occurring due to the surface sensitivity of ARPES is the following. Within first order time dependent perturbation theory ARPES is described by the Fermi golden rule giving the transition probability between two stationary states due to the perturbation, i.e. the photon field. Since for ARPES the contribution of the surface is dominant, the stationary states are the ones respecting the boundary conditions of a semi-infinite solid. These are called inverse LEED states, named in connection with the low energy electron diffraction (LEED) experiment. The point is that, parallel to the surface, periodicity is intact (maintaining Bloch states with quantum number $\vec{k}_1$, 2D Bloch states) whereas it is broken perpendicular to the surface ($\vec{k}_1$ is not a good quantum number, i.e. it is not conserved). In quantum mechanics, a trick is to still use $\vec{k}_1$ as a quantum number but to give it a ‘lifetime’ or a ‘broadening’ ($k_1$ broadening), instead of admitting a linear combination of basis functions from a complete set (i.e. the 3D Bloch functions) and calculating the inverse LEED states. Often in
the literature this non-conservation of $\vec{k}_\perp$ is associated with an electron and hole lifetime and the finite mean free path of the electron. The lifetime and mean free path of the electron due to interactions, however, are not specific for $\vec{k}_\perp$. But the association may be understood by the fact that electron interactions are indeed responsible for the surface sensitivity and therefore for the non-conservation of $\vec{k}_\perp$.

As a consequence, a strict momentum conservation of the 3D wavevector $\vec{k}$ is not guaranteed and only $\vec{k}_\parallel$ is conserved. This may have dramatic consequences, particularly for 3D materials where for a given $\vec{k}_\parallel$ different $\vec{k}_\perp$ contribute to the ARPES signal, especially for $\vec{k}_\perp$ with a high density of states (DOS), leading to a $\vec{k}_\perp$ smearing and the measurement of the so-called 1D-DOS [24]. For 2D materials, as is our case, the situation is to our advantage in the following sense. The transition in the Fermi golden rule goes from the initial state (ground state of the $N$-electron system) to the final state, consisting (within the sudden approximation) of an $(N-1)$-electron system with a factorized independent electron in the detector. The $(N-1)$-electron system of the 2D material leads to the description of the ARPES signal in terms of the one-electron removal spectral function $A(\vec{k}_\parallel, \omega)$ depending only on $\vec{k}_\parallel$ and the electron energy $\omega$ and a 3D final state single electron reaching the detector. This 3D final state indeed suffers from $\vec{k}_\perp$ broadening, also inducing an energy broadening, with the advantage of filling possible final state energy gaps. For strictly 2D electron systems this final state energy and $\vec{k}_\perp$ broadening do not affect the line shape of the ARPES signal, it is only influenced via an energy and $\vec{k}$ dependent matrix element. However, a slight departure from the 2D character may already result in a broadening, not to be confused with spectral function effects (many-body physics) [25].

3. Experiments and calculations

Traditionally, in ARPES, relatively few energy distribution curves (EDCs) were measured for different angles in order to determine the band structure and the $\vec{k}_\parallel$ location where bands cross $E_F$. Then, Santoni et al [26] used an alternative method to obtain the same information. With a 2D display-type analyser they mapped the FS of layered graphite directly by measuring the total intensity within a narrow energy window at $E_F$. Inspired by [26], sequential angle-scanning data acquisition [27, 28], as introduced by the surface structural x-ray photoelectron diffraction (XPD) method, was used to map the intensities within a narrow energy window at the Fermi energy [29–31]. It was also realized that simple intensity mapping at $E_F$ as a function of angle or $\vec{k}_\parallel$ might not be sufficient to determine the Fermi vector ($\vec{k}_F$) and that a combination of intensity mapping and EDCs is necessary in order to verify whether a quasiparticle peak indeed crosses the Fermi level [32]. Another, most important, finding was that so-called angular distribution curves (ADCs), nowadays also called momentum distribution curves (MDCs), i.e. intensities as a function of angle or $\vec{k}_\parallel$ for a fixed binding energy, are much easier to interpret and to fit for quantitative interpretation in terms of the spectral function than the traditional EDCs because of their much simpler line shape (Lorentz-function type functionality of the spectral function) [33–36]. The analysis of ADCs instead of EDCs has been applied successfully to the cuprate high $T_c$ superconductors [37]. These ARPES modes are well established now and have proven their power on many systems [31, 32, 36, 38].

ARPES EDCs were acquired with a Scienta SES-200 hemispherical electron energy analyser with energy and angular resolution of 5 meV and 0.25°, respectively, while the Fermi surface mapping (FSM) measurements have been collected in a modified Vacuum Generator ESCALAB mark II spectrometer [39] with energy and angular resolution of 50 meV and 0.5°, respectively. The sequential motorized sample rotation has been described elsewhere [31].
Figure 3. Fermi energy maps for (a) TaS$_2$ and (b) TiSe$_2$. Raw data have been normalized by the mean azimuthal value (see text). (c) Density functional theory based simulation; inscribed are the CDW-induced reconstructed BZs. High intensity is in white. (d) Sketch of the elliptically shaped Fermi energy contours for nominally one-d-electron (large ellipse) and zero-d-electron (small ellipse) filling.

Monochromatized He I photons of energy 21.22 eV were used for all measurements reported here. The samples were prepared by vapour transport and cleaved in situ at pressures in the $10^{-10}$ mbar region at room temperature. Surface cleanliness before and after ARPES measurements was monitored by x-ray photoelectron spectroscopy (XPS), while LEED was used to check the sample orientation and the evolution of the CDW superstructures. The Fermi energy and instrumental energy resolution were calibrated by measuring a polycrystalline copper sample.

First-principle calculations were performed in the framework of density functional theory (DFT) using the full potential augmented plane wave plus local orbitals (APW + lo) method in conjunction with the generalized gradient approximation (GGA) in the parametrization of Perdew, Burke and Ernzerhof [40] as implemented in the WIEN2k software package [41] as well as the ABINIT code [42] using the local density approximation (LDA) and relativistic separable dual-space Gaussian pseudopotentials [43]. A recent extension to WIEN2k based on the OPTICS package allows the computation of the imaginary part of the static electronic susceptibility $\chi$ (q) [44]. The phonon dispersion is computed with the help of density functional perturbation theory (DFPT) capabilities of ABINIT [45, 46].

4. Results and discussion

4.1. Pseudo-gap and missing Umklapp bands

Figure 3 presents FSMs for TaS$_2$ and TiSe$_2$ collected at room temperature (figures 3(a) and (b)). A DFT based calculation for the undistorted or normal state structure, using a free electron final

3 The ABINIT code is a common project of the Université Catholique de Louvain, Corning Incorporated and other contributors (URL http://www.abinit.org).
state, is shown in figure 3(c). The measurements shown in figures 3(a) and (b) are typical for the nominally one- and zero-d-electron compounds, respectively (see above). The different filling translates into a different sized ellipse around the M point as illustrated in figure 3(d). In fact, FSMs for TaSe$_2$ [19] TaTe$_2$ and NbTe$_2$ [5] are very similar to the result for TaS$_2$, and the FSM for TiTe$_2$ [32] resembles the one of TiSe$_2$.

The calculated lines (only shown for TaS$_2$, figure 3(c)) agree well in shape with the measurements for all compounds. Differences exist mainly in the intensities with marked threefold symmetry for the experiments. This is attributed to matrix element effects [13]. Note that figure 3(a) presents normalized data. In fact, for the raw data (not shown, see [5, 12, 13, 19]) intensity is maximal near normal emission ($\Gamma$) and falls off rather quickly towards larger polar angles. This behaviour is observed for TaS$_2$, TaSe$_2$, TaTe$_2$ and NbTe$_2$ and has been attributed to the $d_{z^2}$ character of the transition metal band [13]. A normalization of the FSM by the mean intensity for each polar emission angle as shown in figure 3(a) eliminates this dependence and allows us to reveal weaker off-normal features. However, centred circular features are then suppressed, with the consequence that the closing of the ellipse near $\Gamma$ is not apparent in the normalized data.

The important point to notice is that the CDW induced lattice distortion is present at room temperature for TaS$_2$ and TaSe$_2$ as well as TaTe$_2$ and NbTe$_2$. However, the corresponding umklapp bands with the symmetry of the BZs due to the reconstruction (as inscribed in figure 3(c)) are not clearly observed. A second important observation is that, again for all but TiSe$_2$ and TiTe$_2$, the high intensity lines in the experiment do not correspond to Fermi energy crossings of intensity maxima or quasiparticle peaks [19, 5, 12]. This is verified with EDCs measured across these high intensity lines as shown in figure 4. EDCs are displayed along the $\Gamma$–$\bar{\Gamma}$ direction for NbTe$_2$, TaS$_2$ and TiSe$_2$. The dispersion of the d band is what is expected from the elliptically shaped FSM contours. What is striking, however, for NbTe$_2$ and TaS$_2$, is a characteristic back-bending of the intensity maximum when approaching $E_F$, not crossing $E_F$ but provoking considerable spectral weight at $E_F$ (as seen in the FSM in figure 3). No quasiparticle crossing has been found for any of the measured energy dispersion curves. Thus, strictly speaking, the maps for TaS$_2$, TaSe$_2$ and NbTe$_2$ are not Fermi surfaces. The observed intensity originates from bands which come close to the Fermi level, but must not be associated with quasiparticle crossings, but rather spectral weight which leaks across the

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig4.png}
\caption{He I excited energy distribution curves along the $\Gamma$–$\bar{\Gamma}$ direction for NbTe$_2$, TaS$_2$ and TiSe$_2$.}
\end{figure}
The signature of such a pseudo-gapped Fermi surface is also observed in high $T_c$ superconductors [47] and has remained a controversial topic.

Therefore, for all but TiSe$_2$, the FSs show that there is a so-called pseudo-gapped Fermi surface. As for TiSe$_2$, this is also not the case for TiTe$_2$, considered as prototype for a 2D Fermi liquid [48–50] where ARPES shows sharp quasiparticle peaks becoming increasingly narrow when approaching $E_F$ according to what is expected from their lifetime.

The point to discuss is whether the two observations, i.e. a pseudo-gapped FS and the non-observation of the Umklapp bands, are connected. From a deeper investigation of TaS$_2$ using DFT calculations it turns out that a connection can be made [19] as follows. From a band-structure calculation point of view, even the slightest atom displacement inducing a structural reconstruction with symmetry breaking (such as indicated in figure 5(a)) will immediately introduce a new BZ (figure 5(b)) with back-folded bands (figure 5(c)). However, for infinitesimal atom displacements the electron wave functions probed by ARPES will practically be unaffected. Therefore, a band-structure calculation of the reconstructed crystal structure will display a wealth of new bands (sub-bands, dotted lines in figure 5(c)), but the spectral weight distribution measured by ARPES will basically continue to follow the band structure of the unreconstructed lattice (open circles in figure 5(c)) [51]. In figure 5(d) the situation is explained for a quasi-1D atomic chain. The top and bottom panels show the unreconstructed and reconstructed chains, respectively. The influence of the reconstruction on the electronic structure (back-folded bands, FS) is sketched (arrows). The spectral weight on the new bands
Figure 6. (a) TaS$_2$ Fermi energy map measured in the commensurate phase at 20 K. Superimposed on it are the $(1 \times 1)$ BZ boundaries as well as the BZ boundaries of the commensurate $(\sqrt{3} \times \sqrt{3})$-R 13.9° superstructure in thin dashed lines. (b) Free-electron final-state calculation based on the (CDW phase) band structure (figure 5(c)). High intensity is in white.

will depend on the scattering strength of the potential with the new symmetry. With this in mind we can understand that the observed EDCs (figure 4) display an intensity distribution following the sub-bands, but are heavily weighted by the bands of the unreconstructed structure. This is in agreement with the observation of a pseudo-gap since the topmost sub-band is present all over the BZ with empty states very close to $E_F$ (figure 5(c)). Within this explanation, we always find spectral weight at $E_F$, whether it is originating from thermal occupation via the Fermi–Dirac distribution or whether states are truly below $E_F$. However, since these states are merely straddling $E_F$, we cannot observe a true crossing of a quasiparticle peak and hence experience a pseudo-gap behaviour displaying an intensity distribution of spectral weight according to the unreconstructed lattice. Figure 6 is in agreement with this view. It shows the spectral weight distribution at $E_F$ for TaS$_2$ measured at 20 K deep in the commensurate CDW phase, together with an FSM calculated for the $(\sqrt{3} \times \sqrt{3})$-R 13.9° structure (see also figure 5(c)). The agreement is nice with high intensity in the centre of the new BZs. However, TaS$_2$ when transiting to the commensurate CDW phase undergoes a Mott transition [52, 53] and is not metallic anymore. Nevertheless, the remaining experimental spectral weight distribution at $E_F$ follows the symmetry of the new BZs [54]. Therefore, a possible explanation of the pseudo-gap is given using the single-particle picture and band-structure calculations together with the fact that Umklapp bands carry only weak spectral weight. Although an identical analysis for TaSe$_2$, TaTe$_2$ and NbTe$_2$ has not been done, a similar explanation may also apply.

4.2. Fermi surface nesting

The next point, which will lead us to a deeper understanding, is the question of the origin of the lattice distortion. What is the strategy of the material to gain electronic energy necessary to compensate the elastic energy for the lattice distortion? The classical mechanism in 1D for the occurrence of a CDW is the Peierls instability [55, 56], where a metal becomes unstable with respect to a spatially modulated perturbation with wavevector $q_{CDW}$ equal to twice the Fermi vector $2k_F$ (figure 5(d)). This leads to the formation of electron–hole pairs with the same wavevector and finally to the opening of a gap which provides a gain in electronic energy in order to compensate the elastic energy paid for the lattice distortion. Figure 5(d) (top) displays the case of a strongly nested FS as indicated by arrows connecting nested (parallel) parts of the FS contour. One electron per atom (and unit cell) of chains of atoms with distance $a$ is
assumed. Therefore, the Fermi vector is \( k_F = \pi/(2a) \) since the quasi-1D band is half filled. The BZ boundary is located at \( 2k_F = \pi/a \) (figure 5d, top). If the lattice is distorted in order to double the periodicity, i.e., to introduce a CDW with wavelength \( 2a \), elastic energy has to be paid. At the same time a doubling of the periodicity in real space results in reducing the dimension of the BZ by a factor of two. The consequence is that the new BZ boundary is now located at \( k_F = \pi/(2a) \) and the band is full (two electrons per unit cell, due to the doubling of the periodicity) with the opening of a gap at the new BZ boundary at \( \pi/(2a) \) and a corresponding gain of electronic energy. However, especially for long \( \mathbf{q}_{\text{CDW}} \) (with many small new BZs) and imperfect nesting (with remaining small metallic pockets) it may be difficult to follow in detail the reconstruction and back-folding of bands unless they carry enough spectral weight (see above). The driving force for such an instability is given by the topology of the Fermi surface (FS), which has to present favourable nesting conditions. Namely, large portions of the FS have to be connected or nested by the vector \( \mathbf{q}_{\text{CDW}} \). A good indicator of the quality of the nesting is the imaginary part of the static electronic susceptibility \( \chi(\mathbf{q}) \) \[57\], which, in linear response theory, relates the response of the system to the perturbation. An estimate of the imaginary part of the static electronic susceptibility \( \chi(\mathbf{q}) \) may be defined as \( \text{Im} \chi(\mathbf{q}) = \sum_{\mathbf{k},n,\pm} \delta(E_{\mathbf{k} + \mathbf{q},n} - E_{\mathbf{k},n}) \), neglecting matrix elements \[58, 5\]. The Dirac function \( \delta \) gives a contribution of either unity or zero depending on whether \( \mathbf{q} \) is a nesting vector or not.

For \( \text{TaS}_2 \), computation of \( \text{Im} \chi(\mathbf{q}) \) \[54\] (not shown) for \( q_\perp = 0 \) gives a local maximum around \( \mathbf{q}_\parallel = 1/\sqrt{3} \times \mathbf{a}^* \), consistent with Myron’s \[59\] calculations obtained for a more limited set of \( \mathbf{q} \) vectors. The maximum is located along the \( \Gamma - M \) direction, 13.9° off with respect to the CDW wavevector of the commensurate phase. This is consistent with the CDW direction of the IC phase rotation by 13.9° while passing through the IC to QC phase transition. However, while FS nesting is a plausible explanation for the onset of the CDW, the overall quite large value of \( \chi(\mathbf{q}) \) found at least along \( \Gamma - M \) is intriguing and questions whether a relatively small and broad maximum at \( \mathbf{q}_\parallel = 1/\sqrt{3} \times \mathbf{a}^* \) is sufficient to explain the occurrence of the IC CDW. As pointed out by Johannes \textit{et al} \[57\], definitive evidence for the contribution of nesting to the CDW formation needs confirmation by calculations of the real part of the static electronic susceptibility. Currently this calculation is beyond our capabilities. Nevertheless, we can expect a local maximum in the real part. In fact, given the geometry of the FS of 1T-\( \text{TaS}_2 \), the maximum in the imaginary part of the susceptibility at the correct \( \mathbf{q}_{\text{CDW}} \) vector can be attributed to the presence of flat areas on the FS, and numerical simulation (not using DFT) for a 2D toy model (not shown), i.e., closed FS with flat areas, leads to a local maximum not only in the imaginary part of the susceptibility but also in the real part of the susceptibility.

For \( \text{NbTe}_2 \) the situation is more clear cut. Its structure is a monochlinically deformed 1T arrangement. LEED measurements suggest that the structure can be understood in terms of a \((3 \times 1)\) superstructure. For completeness, we note that the bulk structure exhibits a \((3 \times 1 \times 3)\) supercell structure, since successive layers are shifted within the plane. The results of \( \text{Im} \chi(\mathbf{q}) \) calculations are presented as linear grey scale plots in figure 7a, with white indicating a large response of the electron system. Strong nesting is present for small, but non-vanishing, \( \mathbf{q} \) vectors. These contributions are due to intraband nesting from weakly dispersing bands and can be reduced by choosing a smaller energy window.

Highly interesting is the peak at \( \mathbf{q} = \frac{1}{4} \mathbf{a}^* \) along the \( \Gamma \Omega \) and \( \Omega \Gamma \) directions in figure 7a (feature 1). We associate this peak with a nesting vector leading to the \((3 \times 1)\) superstructure observed by LEED and plotted in the calculated FS cut in figure 7b. Thus, the electronic structure of trigonal \( \text{NbTe}_2 \) appears unstable with respect to a potential with wavevector \( \mathbf{q} = \frac{1}{4} \mathbf{a}^* \). As discussed in [5], a second maximum (feature 2) at \( q = 0.19a^* \) along \( \Gamma K \) might be related to the \((\sqrt{19} \times \sqrt{19})\) CDW phase, which was observed after cooling of heat-pulsed
Figure 7. Calculations for 1T-NbTe2. (a) Results of Im $\chi(q)$ calculations within the first Brillouin zone at $q_z = 0$ ($\Gamma$MPK plane) obtained by integration of the APW+lo FS. The peak marked 1 occurs exactly at $q_z = \frac{1}{4}a^*$. Feature 2 is close to the nesting vector of the $\sqrt{19} \times \sqrt{19}$ CDW phase. (b) APW + lo calculations of cuts through the Fermi surface at $k_z = 0$ and $k_z = c^*/2$. Nesting vectors of length $q = 1/3a^*$ feature 1 and $q = 0.19a^*$ feature 2 are indicated. (c) Phonon dispersion of 1T-NbTe2 along high symmetry lines obtained by the response function capabilities of ABINIT. Frequencies below 0 meV are imaginary. All degrees of freedom within the trigonal $P3m1$ space group were relaxed using the Broyden–Fletcher–Goldfarb–Shanno minimization scheme as implemented in the ABINIT code.

crystals to a temperature just above room temperature [9]. Cuts through the FS with indicated nesting vectors corresponding to features 1 and 2 are shown in figure 7(b).

The occurrence of a maximum in the electron susceptibility alone does not explain the distortion to the monoclinic structure. The presence of a perturbation with the corresponding $\tilde{q}$ vector is necessary. In the one-dimensional Peierls scenario this potential is provided by a soft phonon mode.

The DFPT phonon band structure for the relaxed trigonal NbTe2 structure obtained by diagonalization of the dynamical matrix along high symmetry lines is shown in figure 7(c). The lowest lying acoustic branch exhibits imaginary frequencies. DFPT contains the implicit assumption that phonons are simple harmonic modes. Soft modes are by definition anharmonic and their frequency goes to zero. Zero frequency implies that the lattice structure is unstable and will transform, typically, to a lower symmetry phase. In the extreme case, electronic structure calculations may give an imaginary phonon frequency, indicating that the ideal structure is unstable [60]. The phonon frequencies are the square roots of the eigenvalues of the dynamical matrix. Imaginary frequencies correspond to negative eigenvalues of the dynamical matrix. A negative entry in the diagonalized dynamical matrix contributes a negative energy to the total Hamiltonian, indicating that the expansion was not carried out around the equilibrium configuration. Thus there exists an energetically more favourable configuration. At high temperature, the lattice has sufficient energy to overcome the energy barrier between two or more symmetry-related variants of the low temperature structure such that the average observed structure has higher symmetry. In such cases the ideal structure is stabilized by high temperature and will undergo a phase transition on cooling, to a low temperature phase whose symmetry differs by the symmetry of the imaginary mode.
Figure 8. (a) Schematic inverse photoemission (IPES, \( \omega > 0 \)) and photoemission (PES \( \omega < 0 \)) spectral functions adapted from the numerical calculations of the Holstein problem for the two limits of weak and strong coupling [65]. (b) Spectral function of a single fixed electron coupled to a bath of phonons of frequency \( \omega_0 \) [66]. The spectrum consists of a multi-peak structure and the quasiparticle ground state is shifted from the free electron energy \( \epsilon_c \) by \( \Delta \), which is larger for strong coupling. The stronger the coupling the more spectral weight is transferred to the satellites (\( \ell > 0 \)), with maximum intensity on the centroid of the spectrum.

The most unstable modes in figure 7(c) occur along \( \mathbf{q} = (1/3, 0, q_z) \alpha^* \). This strongly supports the Fermi surface nesting scenario for NbTe₂. Furthermore, from an analysis of the eigenvectors of the dynamical matrix, the distorted structure may be qualitatively constructed.

4.3. Electron–phonon coupling

Despite the explanation given above for the absence of any clear quasiparticle crossing at \( E_F \) using DFT calculations, relying on many new BZs and Umklapp bands with weak spectral weight, the broad spectral weight distribution instead of delta-function-like peaks is peculiar. Electron–phonon coupling is expected to be relatively strong for the CDW compounds [9] and consequently polaronic effects may also play a role. Polarons recently received increased attention for the interpretation of anomalously broad ARPES features [1, 50, 61]. The polaron concept was first introduced by Landau [62] and represents an electron moving in a polarizable lattice carrying the lattice deformation with it. In the Holstein model [63] one assumes an on-site coupling of the electrons with dispersionless lattice vibration modes. Thus this model is adapted to short range electron–lattice interaction and refers to a small polaron. Since for increasing strength of the coupling the polaron radius shrinks to a single lattice site, the Holstein model is more suitable to study strong interactions. In the case of the Fröhlich Hamiltonian [64] long range electron–lattice interaction is also active with lightly dressed particles polarizing the environment weakly in a large spatial extent around themselves. Thus in this case the polaron is called a large polaron. The influence of electron–phonon coupling, if strong enough, may even gap a non-nested FS as shown by recent numerical calculations of the spectral function for the spinless Holstein model [65]. An illustration of this behaviour at half filling is shown in figure 8(a). In the weak coupling regime most of the spectral weight close to \( k_F \) resides in the polaronic quasiparticle. The dispersion of this band exhibits a mass enhancement due to the weak dressing of the electron. Further away from \( k_F \) the polaron band is no longer visible since most of the spectral weight is transferred to the incoherent part which follows the bare (not renormalized) dispersion \( \epsilon(k) \). For strong coupling the spectrum consists again of the polaronic quasiparticle and an incoherent part. However, as a consequence of the predominantly local effects in the strong coupling regime, the polaron quasiparticle contains almost zero electronic spectral weight for the whole momentum range and turns into a very narrow band. The
incoherent part, where the electronic spectral weight is concentrated, broadens and does not cross the Fermi level but flattens in its proximity, forming a pseudo-gap.

Intuitively, this behaviour can be understood from the exactly solvable model for the coupling of a single electron to a bath of independent phonons of frequency $\omega_0$ [66]. The broad line shapes as well as the vanishingly small quasiparticle spectral weight comes naturally from the calculated spectral function of this model. Figure 8(b) sketches the outcome for the fixed particle with bare energy $\epsilon_c$ interacting with a set of Einstein phonons of frequency $\omega_0$. Due to the coupling the ground state energy is shifted by $\Delta = g\omega_0$, $g$ being the coupling. The spectral function is an envelope of many individual peaks spaced by $\omega_0$. The peak closest to the Fermi level is the quasiparticle peak or zero-phonon peak. It is shifted by $\Delta$ from the non-interacting energy position $\epsilon_c$. The following peaks are satellites. They indicate that the single electron is not an eigenstate of the Hamiltonian and the particle has a finite probability of occupying other states carrying $l$ phonons with it. From the photoemission point of view it means that a removal of an electron from the system occurs with a probability of shaking off a certain number of bosons. This picture also refers to what is called the Franck–Condon broadening. Depending on the coupling more or less spectral weight is concentrated on the quasiparticle peak and the broadening is more or less pronounced. A polaron scenario thus allows us to explain the apparent absence of quasiparticle crossings as well as the broadened line shape in the experimental ARPES spectra.

Figure 9 presents ARPES measurements in the IC and QC phases, taken at 350 and 344 K, respectively. The intensity plots (high intensity in black) on figure 9(a) are azimuthal cuts taken at a polar angle of 14° with respect to $\vec{F}$. The only visible band, near $E_F$, originates from Ta 5d electrons. In figure 9(b) these intensity plots are symmetrized with respect to the Fermi level and summed, in order [67, 68] to remove the perturbative effect of the Fermi–
Dirac distribution cut-off and to infer whether the spectral function peak crosses the chemical potential or not. Figure 9(c) shows the EDCs corresponding to the intensity plots shown in figure 9(a). In figure 9(d) a sketch of the FS contours is drawn, together with a superposition of the normal- and commensurate-state BZs. The curved arrow indicates the location of the intensity plots of figure 9(a). The presumed nesting vector \( q_{\text{CDW}} \) is also drawn according to the calculation of \( \chi(q) \), connecting flat parts of the elliptically shaped FS with large, possibly nested (parallel) portions which have strongly turned the discussion of the origin of the CDW towards the nesting scenario.

In the following strong similarities with figure 8 become evident, giving good reasons for a polaron scenario. Experimentally, large parts of the FS have been investigated (not shown) and we do not observe a different behaviour than that shown in figure 9 and described next.

In the IC phase (figures 9(a) and (b), left) a broad Ta 5d band flattens and narrows slightly when approaching \( E_F \), giving rise to a small maximum centred at \( E_F \) in the symmetrized plot. In a standard interpretation of the symmetrization procedure this would be attributed to a Fermi level crossing of the quasiparticle peak. However, band calculations predict a linear slope through \( E_F \) and no flattening is observed. Moreover, within the Fermi-liquid picture ARPES peaks are attributed to quasiparticle excitations whose lifetime increases when approaching \( E_F \). However, in our case the Ta 5d peak remains anomalously broad near \( E_F \). Thus we are clearly in the presence of a renormalized band which touches \( E_F \) without clearly crossing it.

An enhancement of the previous anomalies can be observed upon cooling down to the QC phase (figure 9(c)). The line shape gets abruptly broader (when lowering \( T \) by only 6 K); the centre of mass of the Ta 5d band shifts towards higher binding energy when approaching the Fermi vector and becomes clearly separated from a small maximum visible in the symmetrized plot (figure 9(b), right). This symmetrization peak originates from a finite spectral weight near \( E_F \). As described above, Bovet et al. [19] attributed this to a reconstruction induced band coming from above and merely straddling \( E_F \). In the polaron picture proposed here, this finite intensity near \( E_F \) is also compatible with a remnant weight of the coherent quasiparticle (figure 8) [65].

The changes between the photoemission spectra of the IC and QC phases are abrupt and are interpreted as a change of the electron–phonon interaction character. Indeed, in 2D or 3D systems there is a qualitative dependence of the polaron type (large or small) on the range of the electron–lattice interaction [69]. While large polarons are formed if the electron–lattice interaction due to the long range Coulombic interaction between electronic carriers and lattice ions is of predominant importance, small polarons form if the short range electron–lattice interaction such as the deformation potential interaction dominates. The formation of the commensurate domains in the QC phase introduces, as shown above, the ‘stars of David’, which can be identified with small molecules acting as potential wells and hence favouring the short range interaction. Therefore, the CDW in the IC phase leads to a static deformation of the whole lattice structure, which, in turn, leads to a more local character of the electron–lattice interaction and an enhancement of the electron–phonon coupling strength, i.e. to a lattice-distortion-enhanced electron–phonon coupling.

For TaSe\(_2\) and NbTe\(_2\) it is likely that the situation is similar, although a detailed analysis has not been made. Inspecting the spectra in figure 4 reveals a very similar behaviour of broad peaks approaching \( E_F \) without crossing it together with spectral weight at \( E_F \) responsible for the FSMs being in good agreement with DFT calculations reproducing the image of the bare dispersion. For TiSe\(_2\) and TiTe\(_2\), however, the scenario appears to be very different. Sharp quasiparticle peaks are clearly crossing \( E_F \). For TiTe\(_2\), the FL prototype, this is understandable, but TiSe\(_2\) undergoes a CDW transition and the mechanism for the system to gain electronic energy is under debate [70–72].
5. Conclusion and outlook

In summary, we have discussed a series of transition metal dichalcogenides related to the 1T structure, therefore having similar crystallographic and electronic structures. We have observed a distinctly different behaviour between the compounds with, according to the ionic picture, nominally one and zero d electrons in the lowest occupied d band.

The $d^1$ compounds all show a very similar behaviour as observed with ARPES and all present lattice distortions deviating from the 1T structure. The ARPES intensity distribution at $E_F$ is well described by the single-particle or DFT-based band structure obtained for the un-reconstructed 1T structure. However, closer examination shows that there are no quasiparticle peaks crossing $E_F$ such that the FS appears pseudo-gapped. Two explanations have been put forward.

The first is based on DFT calculations for the distorted structure taking into account the many new sub-bands (Umklapp bands) producing electronic states all over the normal state BZ with most of the ARPES spectral weight along the bands of the un-reconstructed structure. In this scenario, the unusual energy broadening of bands may be explained by the contribution of many (experimentally unresolved and energetically close) sub-bands.

The second explanation is inspired by the peak shape of dispersing features being particularly large. A spectral function based on polaron formation has all the attributes to explain the experimental observations. In fact, electron–phonon coupling dresses the quasiparticles with lattice vibration excitations, leading to a broadened and re-normalized line shape resembling strongly the experimental observation.

The $d^0$ compounds are very different. Sharp quasiparticle peaks are observed; the FS is not pseudo-gapped. However, whereas one of the compounds behaves in a Fermi-liquid-like way, the other undergoes a CDW transition. It is not clear currently what mechanism is responsible for the instability. It is certain that a delicate balance between interactions of the chalcogen sp electrons with the transition metal d electrons, their mobility and screening power together with the reaction of the lattice determines the amazing properties of these materials.

The calculated $d^1$ material’s Fermi surfaces are all similar in shape and resemble an ellipse with rather flat parts that are susceptible to FS nesting. In order to investigate FS nesting as the driving mechanism for the lattice distortion the imaginary part of the static electronic susceptibility has been calculated. The case for FS nesting is rather clear for NbTe$_2$, since phonon calculations also display a respective softening. It cannot, however, be generalized for the others. In particular, the situation is more complicated for TaS$_2$.

An important conclusion is that from the present point of view we cannot distinguish between the two explanations for the behaviour of the ‘one-electron’ compounds, i.e. single particles with many sub-bands, versus polarons. The interpretation is not unique. It is necessary to refine the analysis to find unique differences. More accurate simulations are necessary in the future, in the sense of not only modelling the physics of the $N$-electron system but including photoemission matrix elements for a quantitative comparison. This would remove ambiguities related first to strong versus weak spectral weight and second to initial state (physics, spectral function) versus final state (experiment specific, geometry) related broadening.

Acknowledgments

Skilful technical assistance was provided by the Neuchâtel mechanics and electronics workshops. Many helpful discussions with H Beck are gratefully acknowledged. This project has been supported by the Fonds National Suisse de la Recherche Scientifique through Div. II and MaNEP.
[57] Johannes M D, Mazin I I and Howells C A 2005 Preprint cond-mat/0510390
[60] Ackland G 2000 RIKEN Rev. 29 34
[62] Landau L 1933 Z. Phys. 3 664
[63] Holstein T 1959 Ann. Phys. 8 325