Singular Learning Theory 22 - Solid state physics 2

In the first talk we discussed dispersion relations, Density of States (DOS), Fermi energy and Fermi surfaces and we saw one calculation involving band gaps in semi-conductors to illustrate the process of model-making in solid state physics. This was a simple illustration of the following paradigm:

$$\text{Microstructure} \longrightarrow \text{Density of states} \longrightarrow \text{Bulk electrical properties}$$

In the context of twisted bilayer graphene we discussed how singularities in the dispersion relation cause divergences in the DOS and these divergences manifest themselves as interesting macroscopic physics. This is an eye-catching example but this pattern is common: today we will discuss such divergences in detail for carbon nanotubes following [K, Ch. 18].

1. The analogies to SLT

- The dispersion relation $E(\underline{k})$ or $\epsilon_{\underline{k}}$ is like $K(w)$, singularities in both determine the divergences in the DOS and hence many electrical properties resp. learning behaviour (in the Bayesian sense, i.e. asymptotic free energy, generalisation error).

- The power of "easy math" on top of the base theory of solid state physics (e.g. Kittel) to describe a wide range of optical & electrical properties of solids might be viewed as an encouraging sign for SLT applications (aka the field of SLT needs "physics thinking" as well as proofs).

- Singularities aren't weird: you can make twisted bilayer graphene with a pencil and scotch tape. Let go of that lamp post.

However these analogies should not be overstated, there are differences:

- Electrical properties of solids involve electrons (shock), which are fermions and are governed by the Fermi-Dirac distribution (due to the exclusion principle). SLT is about the Bayesian posterior or Boltzmann distribution, and doesn't directly involve fermions (although the role of noise in SLT makes it natural to introduce them following Parisi-Sourlas).

  We saw last talk (following [K, Ch. 8]) how in some cases $\mu - \epsilon \gg k_B T$ and we're effectively dealing with the same objects (see [K, p. 137]).

  In general, be careful about translating SSP $\longrightarrow$ SLT!

- The analogue of bands is subtle: the electrons in a crystal are arranged in energy bands, and whether a material is an insulator, conductor or semi-conductor depends on the gaps between these bands and how "filled" they are. The response of the material to electromagnetic radiation also (as we will see) depends on this band structure.

  Are there bands in SLT? Strictly speaking a band is a single branch of the $\epsilon_{\underline{k}}$ vs. $\underline{k}$ surface [K, p. 224], and in crystals the gaps between bands occur because the band electrons are not free: they are perturbed by the periodic potential of the ion cores. There is no analogue of this periodic potential in SLT.

  The first observation is that the dispersion relation $\epsilon_{\underline{k}}$ is formulated in momentum space with $\underline{k}$ meaning the plane wave solution of the Schrödinger equation $\Psi(x) = \exp(i\underline{k}\cdot\underline{x})$ (for free electrons).

In SLT we are probably right to think of $K(w)$ as being in <u>position space</u> w.
In SSP the material is infinite and periodic, but in SLT there is no similar structure in W, and consequently no decomposition of $L^2(W)$.

Nevertheless in a model with sufficiently many phases (of the Bayesian posterior) creating "dense enough" intervals of free energies, some of the same intuitions may apply. More on this some other time.

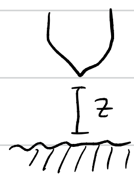
2. Spectroscopy of singularities (see [K, Ch.18])

In a scanning tunneling microscope (STM) an instrument with a sharp metal tip (~ one atom) is brought to within a nanometer of the conducting sample. A voltage bias V is applied to the sample, and a tunneling current I flowing between the tip and sample is measured. The current is proportional to P, the tunneling probability

distance

$$P \propto \exp\left(-2\sqrt{2m\phi/\hbar^2}\, z\right)$$

effective barrier height



which is exponentially sensitive to the tunneling distance ($\Delta z = 0.1\,\mathrm{nm}$, $P \to 10P$).
The STM tunneling current I as a function of bias V gives spatial and spectroscopic information about the quantum states of a nanostructure. At $T=0$

"differential conductance"

$$dI/dV \propto P \sum_j |\psi_j(\underline{r}_t)|^2 \delta(\epsilon_F + eV - \epsilon_j)$$

It is proportional to the DOS at the tunneling electron energy $\epsilon_F + eV$, weighted by the electron probability density at the STM tip $\underline{r}_t$.

Divergences in the DOS may be detected by measurements of the differential conductance as V varies (you can "see the singularities"). To give a concrete example consider a nanoscale wire. The energies and eigenstates are

$$\varepsilon = \varepsilon_{i,j} + \hbar^2 k^2/2m \qquad \psi(x,y,z) = \psi_{ij}(x,y)\, e^{ikz}$$

where i,j are quantum numbers in the x,y directions, k the wave vector in the z direction. (i.e. do particle in a box in x,y directions).

The dispersion relation consists of a series of 1D subbands, each corresponding to a different transverse energy state $\varepsilon_{i,j}$, so

recall free electron gas, $D \propto \frac{d}{d\varepsilon}(\varepsilon)^{d/2}$

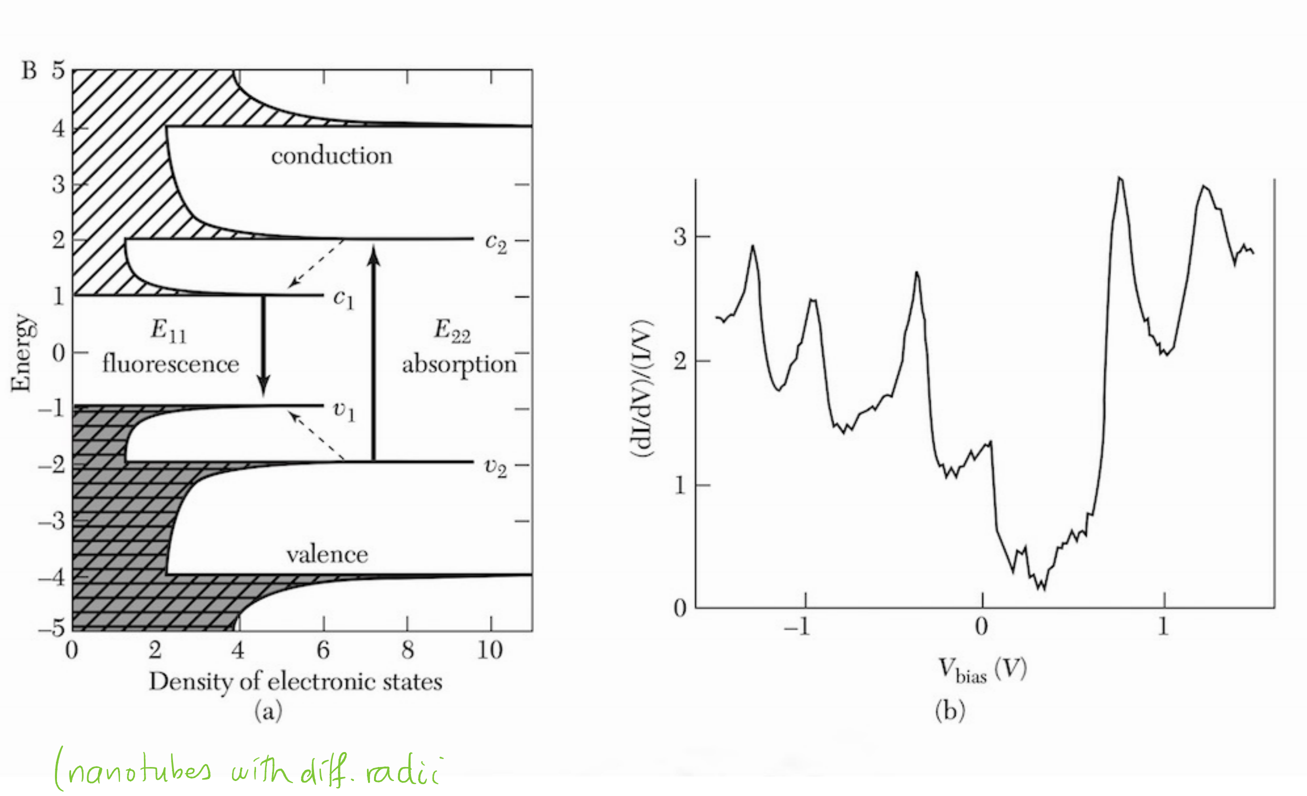
$$D(\varepsilon) = \sum_{i,j} D_{i,j}(\varepsilon)$$

$$D_{i,j}(\varepsilon) = \frac{dN_{i,j}(\varepsilon)}{d\varepsilon} = \frac{2L}{\pi}\left[\frac{m}{2\hbar^2(\varepsilon - \varepsilon_{i,j})}\right]^{1/2} \qquad \varepsilon > \varepsilon_{i,j}$$

$$= 0 \qquad \varepsilon < \varepsilon_{i,j}$$

Note that the DOS diverges as $(\varepsilon - \varepsilon_{i,j})^{-1/2}$ at each subband threshold. These divergences are examples of <u>van Hove singularities</u>.

These singularities affect the electrical and optical properties of 1D systems. To illustrate we consider a semiconducting carbon nanotube, whose band structure is shown below. VHS are seen in scanning tunneling spectroscopy, see (b). Peaks in the differential conductance are observed at bias voltages corresponding to the energies of these singularities.



(nanotubes with diff. radii & chirality)

The optical absorption and emission of semiconducting nanotubes are also dominated by these singularities, since they depend on the initial and final DOS, by Fermi's Golden Rule

$$\omega_{i\to j} = \frac{2\pi}{\hbar} |\langle j | e\underline{E}\cdot\underline{r} | i \rangle|^2 \,\delta(\epsilon_j - \epsilon_i - \hbar\omega)$$

(transition rate due to absorption)

and similarly for emission. The absorption of incident light is enhanced when the frequency (thus energy) matches the gap between two VHS, and similarly for emission

When you look at a carbon nanotube you see singularities.

VHS are a measure zero subset of the energy levels, but most of the states are there, so most of the physics is there.

3. Forward looking statements

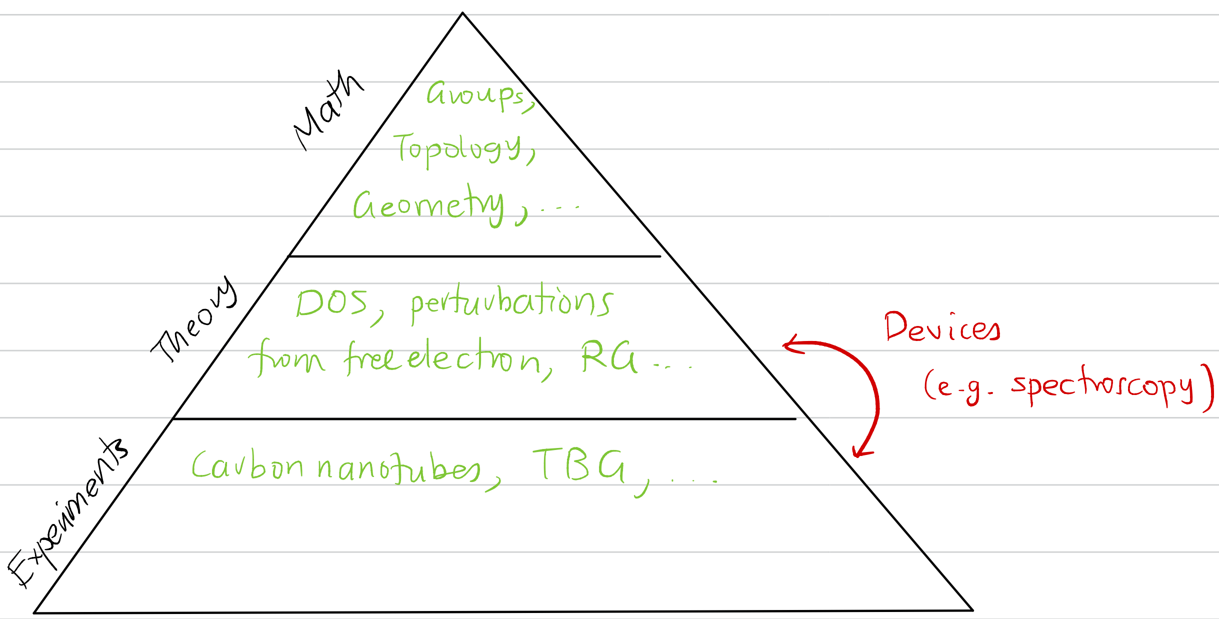
We can now understand some comments in the twisted bilayer graphene literature:

- "If the Fermi level lies in the vicinity of a Van Hove point, the singular DOS determines the physical behaviour due to the large number of available low-energy states" [C].

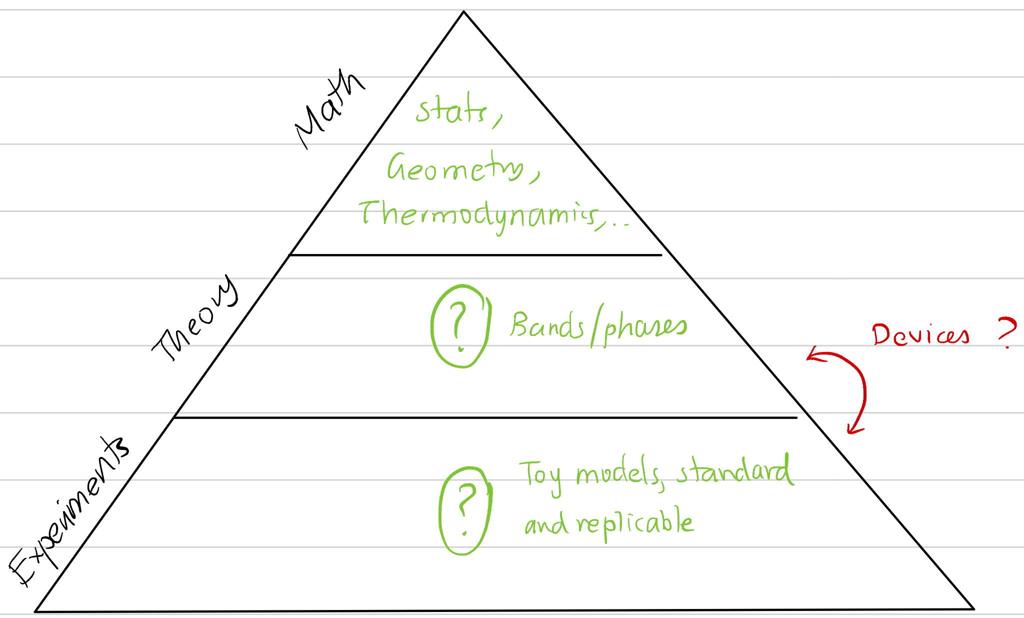
- "Due to the power-law divergence in the DOS, the high-order saddle points are more dominant than the other parts of the Fermi-surface at low energy. We thus construct the low energy theory by approximating the Fermi surface with six patches in the vicinity of these points... " [P].

In [P] they do an RG analysis to uncover the low energy effective theory (in TBG but also more generally) when multiple higher-order VHS are inside the first Brillouin zone.

Solid State Physics



SLT



We have discussed open math problems in SLT. But while of fundamental importance, proofs alone will not get us there. We need:

- Toy models Small tanh, small Transformers,... with publicly available notebooks that replicate estimations of theoretical quantities. Here we should take our cues from the mechanistic interpretability community.

- Theory of "bands" or phases that is good enough to match to measurements, with the aim to explain observed phenomena in the toy models, where the measurements come from

- Spectroscopes for learning machines There are many phenomena in learning machines sensitive to DOS, for the same reason differential capacitance is. Find the ones that are easy to engineer around, and develop a large and reliable library of these "devices".

- Develop RG methods following the solid state physicists...

References

[K] C. Kittel "Introduction to solid state physics" 8th edition

[C] L. Classen et al "Competing orders at higher-order Van Hove points" 2020.

[L] Y-P. Lin, R. M. Nandkishore "Parquet renormalisation group analysis of weak-coupling instabilities with multiple high-order Van Hove points inside the Brillouin zone" 2020.