

Phase Transitions: Metal-Insulator transition in VO_2 and Ferroelectric transition in BaTiO_3

Nov. 14, 2012 version

1. How to prepare for the measurement

This write-up contains a lot of material. You should study parts of it before coming to the lab, and read and study the entire text throughout the course of this experiment. **NOTE: a big chunk of this manual is in the Appendices, so do not treat them as secondary-importance material.** You should also supplement this write-up by looking up and reading some of the books/articles/websites from the list of references. We propose that you start in the following order:

First read sections 1 – 5 which will give you an overview of this lab. Then spend some time learning how to use the lock-in amplifier. Your **first task** will be to master preliminary practice measurements of resistance and capacitance. During the first two periods of the lab work concentrate on **section 6** of this write-up entitled “**Measurement techniques**” and study **Appendices 1 and 2**. Read the equipment manual for the lock-in amplifier and Ref. 1; study the principle of its operation. Simultaneously, study the instrument hands-on: connect a circuit and first measure a known resistor, and then a known capacitor using the lock-in.

Once you will get familiarized with using the lock-in, read about the rest of the experimental set-up, which is described in **sections 7 – 9**. Then you should start reading more of a physics theory, which is to be found mainly in a large **Appendix 5**. The Physics of Phase Transitions is not trivial, and it will take time to “digest” and understand it.

From the beginning, keep good records in your notebook. All electrical circuits that you will build, simple calculations which accompany construction of these circuits, such as estimates of the measuring current, choice of a limiting resistor, calculations of resistance and capacitance, sketches of samples or experimental apparatus, details of sample preparation and geometry, all notes concerning physics, and all your calculations of physical quantities should be recorded in the notebook. Short comments you may write should be understandable, in case you or someone else will later examine your records. The actual experimental data will be stored in the computer and need not be in the notebook.

2. Goals of the experiment

- Understand experimental techniques used in this lab, including the principle and practical operation of a lock-in amplifier when measuring resistance and capacitance, the *four-probe* method, temperature control.
- Learn to handle and prepare samples; learn how to quickly check samples before they are used in lengthy measurements.
- Study the metal-insulator transition in VO_2
- Study the ferroelectric transition in BaTiO_3
- Learn basic physics associated with phase transitions and apply it to interpreting specific experiments at hand

- We do not just do measurements for the sake of measurements. An experiment should lead to some conclusions and it should culminate and result in learning of some “new” Physics (new to us in this Lab; perhaps not new to the Physics community). Therefore it is very important to use your data to extract some meaningful physical results: namely, estimates of microscopic or theoretical quantities which enhance our quantitative understanding of the studied phenomena. You will do so for both phase transitions studied in this lab.

3. Experimental set-up

- Sample chamber with 4 electric leads.
- Temperature controller
- Lock-in amplifier to measure the samples
- Keithley four-probe resistance meter to measure temperature
- Computer (username: *gradlab*, password: *senior*)
- Decade resistance box, additional resistors, capacitors, BNC cables etc.

See the equipment manuals for details. The temperature controller has a WEB interface. Use this interface for programming the controller (see Appendix 3). The lock-in amplifier and the Keithley multimeter are connected to a computer by the IEEE 488 interface. A data taking routine, written in LabView, is provided for recording data (Appendix 4).

Samples: You will be studying samples of VO_2 and BaTiO_3 . If two students work on this lab, they should together measure at least one VO_2 sample, which will be provided, and prepare themselves and measure two BaTiO_3 samples. Sample preparation and all the measurements can be done in collaboration, but for the purposes of data analysis we will consider that one of these two samples belongs to one student, and another one to the other, so that each student will report and evaluate the data on the common VO_2 sample and on his/her own BaTiO_3 sample.

4. Principle of the measurement and an overview of data analysis

Details on these subjects will be found in the rest of this write-up; and you know what they say: *the devil is in the detail*. Here is the **brief summary**:

You will be studying two different phase transitions, both taking place as a function of temperature: a metal-insulator transition in Vanadium Dioxide (VO_2) and a ferroelectric transition in Barium Titanate (BaTiO_3). Read **Appendix 5** where the Physics of phase transitions is reviewed and detailed.

The sample is placed in a temperature-controlled chamber. The temperature is slowly changed, first rising, and then cooling down, while either resistance R (in VO_2) or capacitance C (with BaTiO_3 serving as dielectric) is measured and data is recorded. The temperature will change between room temperature and ~ 100 C in case of VO_2 , and between room temperature and ~ 150 C in case of BaTiO_3 . The measurement of R and C is performed in an AC mode, using a lock-in amplifier, which is described below in section 6 and in Appendix 1.

What will you be able to learn from these measurements? In case of resistive VO_2 data, you will be able to make conclusions concerning the nature and *order* of a phase transition, including estimates of the *metal-insulator* transition temperature, of the *percolation* transition temperature, of the *coercive* temperature.

You will also be able to analyze the data to extract the *gap* of VO₂ in a semiconducting phase and to estimate absolute resistivities of the two phases.

In case of a capacitance measurement, BaTiO₃ will be the dielectric inside a capacitor. In a parallel-plate capacitor, $C = \epsilon A/d = \epsilon_0 k A/d$, where A is the area of the capacitor, d is the distance between the plates, ϵ is the dielectric constant, ϵ_0 is a constant called *permittivity of a free space*, and k is the *relative* dielectric constant. We will be interested in finding k as a function of temperature. It will not be possible to exactly measure A and d in the type of samples you will have, but they can be roughly estimated. So, by measuring C (which is measured quite precisely), you will be able to estimate k to within about a factor of two accuracy, the uncertainty coming almost entirely from a limited ability here to evaluate capacitor geometry. At the same time, you can certainly follow how k will change with temperature with the full precision of the capacitance measurement. This T-dependence will pinpoint the temperature of a phase transition, T_C . Combined, these data will allow a calculation of microscopic parameters of a ferroelectric transition as outlined in **Appendix 5**. At the end, you will be able to estimate an interesting microscopic quantity: the very small distance by which atoms shift in the lattice of BaTiO₃ in a ferroelectric transition. The ability to estimate such a quantity is what makes this experiment especially interesting.

List of physical quantities which you will be expected to extract, calculate and estimate from your data (some of these can be measured rather precisely; others only roughly estimated):

In VO₂:

- The semiconducting gap E_g
- The resistance ratio between semiconducting and metallic phases
- Absolute resistivities in semiconducting and metallic phases
- The value of T_C
- The coercive temperature T^*
- The temperature of a percolation transition T_M

In BaTiO₃:

- T_C
- Relative dielectric constant k vs. T , from about 20 C to 150 C.
- Curie-Weiss law at $T > T_C$ and Curie-Weiss constant C_{CW}
- The value of a microscopic dipole moment p
- The effective atomic displacement x responsible for the appearance of p
- The last two quantities can be also calculated for BaTiO₃ single crystals using the literature value of C_{CW}

5. Looking ahead: what should and should not be included in your Final Report

Your performance will be judged primarily from the contents of your Final Report. In it there will be several different (but related) parts:

- A Physics introduction concerning Phase Transitions and materials which you studied and a Physics analysis of your results at the end. To get these right, you have to study Appendix 5 of this write-up and related literature.
- A description of experimental set-up (including electrical circuit diagrams) and experimental techniques that you used (lock-in and how it was used; 4-probe method), including significant details of how you obtained your data. For example, if you are measuring some voltage, but your “final product” is capacitance, you will need to explain in some detail how you proceeded from that voltage to knowing that capacitance. This may include phasor diagrams and some formula derivations given in section 6 and Appendix 1.
- Initially you will be measuring known resistances and capacitances just to learn how to connect the circuits and how to use the lock-in amplifier. It is very important to perform such *calibrations* or *baseline* measurements in order to be sure that later you will be measuring your samples correctly. While a detailed description of such calibrations may be perfectly appropriate for your mid-lab report, it should not take center-stage in the Final Report, except as a short mention of the fact that these calibrations were done. You may want to report the uncertainty (i.e. % of relative deviation) in the measurement of a known R and C (of course, assuming that the values of these R and C were known very accurately in the first place). But detailed data on such baseline measurements would be out of place; the focus of the Final Report should be on measuring real samples which exhibit phase transitions.
- You will present results of sample measurements, your *data*, in the form of graphs, possibly some tables (but do not list all of the data points in long tables – this is neither necessary nor desirable in the Final Report). If you are measuring resistance R and capacitance C as a function of temperature T, your data may consist of graphs of R(T) and C(T).

Let us discuss this in a little more detail. There exists a notion of *raw data*, the data you directly obtain in your measurement, directly read off the instruments. And then there is data which was to some degree analyzed and organized. For example, you may be measuring separate values of voltage V and current I, and later you may be combining the two when obtaining resistance $R = V/I$. But your interest – and the interest of the reader of your report -- is in that resistance R, and not in V or I. Thus one should use common sense when reporting data: probably in this example, separate values of V and I should not be reported (they can remain in your electronic data storage, to be examined if necessary), and R should be reported. Likewise, if you are measuring a voltage signal which is being converted to a capacitance, do not report those voltages; report capacitance directly. You may however mention the range of the voltage signals that you measured in order to tell your reader something meaningful about the quality of your data.

- The data should be presented with *uncertainties*, which not only should be included, but accompanied with a brief explanation of how they were obtained or estimated. If a calculation is involved, the error must be propagated from the raw data to the calculated quantity. If curve fitting is involved, the uncertainty will come out of the fitting routine. You need also to point out which uncertainties are more significant, and which can be ignored; in this lab this last comment will be particularly relevant, as some of the uncertainties will far exceed others.

- There will be data analysis: corrections which you introduced to your data, extraction of physical parameters of interest. And at the end there will be some conclusions based on your results. You will need to study this write-up and references in order to “connect the dots”, to make your experiments meaningful. When applicable, you will compare your results to the published results. Finding relevant publications by doing literature searches, reading them, and extracting from these publications important parameters, is part of a skill of an experimental Physicist.

6. Measurement techniques

The main instrument you will be using in your measurements in this lab is called *lock-in amplifier*⁽¹⁾. This wonderful instrument was invented in the 1950's by a physicist Robert Dicke; it is used extensively for a variety of measurements in Physics experiments and elsewhere.



Robert Dicke (1916 – 1997), a famous American physicist who made important contributions to the fields of astrophysics, atomic physics, cosmology and gravity. He also invented lock-in amplifier in the 1950s, with the first commercial model developed in 1962.

The advantage a lock-in gives to an experimentalist is in that it rejects a good deal of noise, increasing signal-to-noise ratio and allowing using small currents which will not damage even the most delicate of samples. Additionally, by virtue of always measuring in an AC mode, it automatically averages the readout voltage, getting rid of the so-called *thermal emf* offsets and DC level drifts which are endemic with DC measurements, even when we are using the best and most accurate DC voltmeters.

While the lock-in amplifier was invented to measure weak signals in the presence of large noise, in our measurements we will not be pushing the limits of its capacity in that regard. Instead, we will be using its ability to measure signals of different specified *phases*, as explained below and in Appendix 1. This lab gives you an opportunity to learn about this important instrument.

The name “lock-in” refers to locking to a *frequency* and to a *phase* of an AC signal, and measuring only that part of a signal which is in a specified phase relationship to the reference. This is a version of the *phase-sensitive* detection technique. The phase can be continually varied and also changed in 90° steps.

You should read about lock-in operation in our lock-in manual (available on line through Blackboard) and elsewhere⁽¹⁾. Here and in Appendix 1 we give a brief general introduction to its operating principle, and discuss in more detail the two types of measurements we will be doing in this lab.

If you have not been dealing with AC circuits lately, this would be a good time to thoroughly review them: read about AC currents in general, and the *phasor method* in particular in your introductory

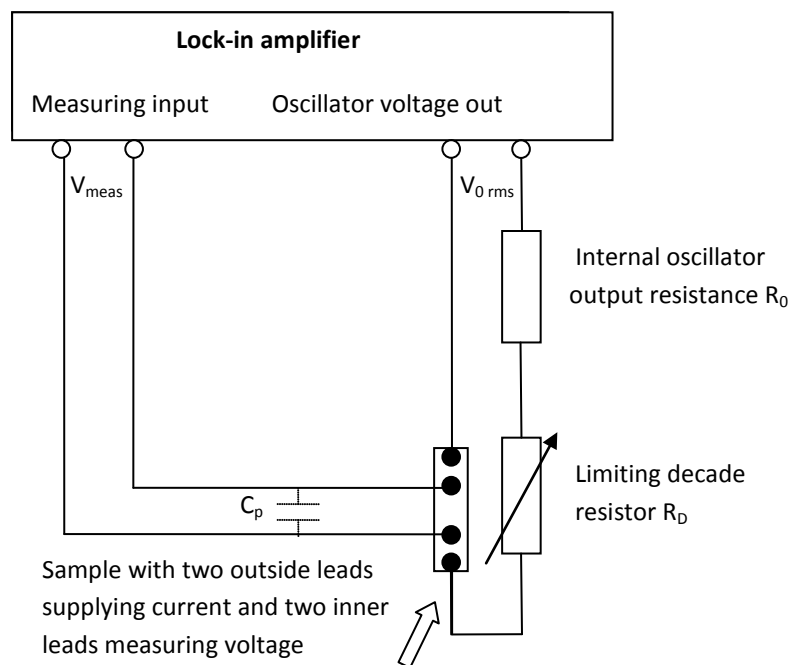
Physics course textbook; some of this is also discussed in Ref. 2; study this section and Appendix 1; it is essential that you understand this topic in order to do this lab.

The principle of lock-in measurement is as follows: lock-in contains its own AC voltage source called *oscillator output*, in which you can adjust the amplitude (but only up to 2.0 V), and frequency (adjustable in a wide range from a fraction of 1 Hz to 100 kHz). The frequency, once you will set it, will be kept with great precision and stability by a lock-in. The readout is tuned to read only voltage at this frequency (actually, at a double frequency, but this is immaterial for our general discussion), all the other frequencies being filtered out. Moreover, it will *look* into a specified *phase* of this signal. The exact meaning of this last statement will be explained below and in Appendix 1.

The *measuring input* (actually, there are two of them, A and B, and you can also hook the instrument up to measure the difference $A - B$) is an AC voltmeter, which will measure *rms* values, and where the sensitivity of a readout can be changed according to experimental needs, from several volts to a few mV full scale. It can be connected across different parts of a circuit: for example, it can be connected across R or across C. Let us now discuss specific measuring arrangements: one for measuring a resistance, another for measuring a capacitance.

How to use lock-in to measure resistance?

Suppose we want to measure a resistance, which may be changing in the course of a measurement, for example, as a function of temperature. This will be the case in our VO_2 measurement in this lab. We may connect the lock-in as shown here:



Looking at the front panel of the lock-in reveals that the oscillator circuit has internal resistance $R_0 = 600 \, \Omega$, which is really a part of the lock-in, but, for clarity, is shown on the diagram above outside of a lock-in

block. It is connected in series with *limiting* decade resistance R_D and the sample. The *measuring input* of a lock-in is connected across the inner contacts of a sample.

Let us first discuss **the current**; it is alternating, AC current; we will be always talking about *rms* values of all voltages and currents in this discussion. The *limiting* resistor R_D serves to limit the current, and to make this current approximately constant. To do so, R_D should be much greater than the sample resistance R_{Stot} , which is the total resistance of a sample between the outside current contacts shown in the figure above, and also includes the resistance of these current contacts (a contact to a sample may have appreciable resistance, see Appendix 2); it is the total resistance encountered by the current as it flows through the sample. The value of R_{Stot} will be *changing* in the course of a measurement, as we change the sample temperature. Let us say its maximum value is $R_{Stot} = 500 \text{ k}\Omega$, and the minimum value can be neglected compared to $500 \text{ k}\Omega$, which is indeed the case with VO_2 (assuming contact resistance is not very large). Suppose we decided that we will be satisfied with keeping current constant to 5%, i.e. we are not worried about up to 5% possible overall distortion in our measurement. The maximum and minimum *rms* currents then will be $I_{\max} = V_0/(R_0 + R_D)$ when R_{Stot} is negligible, and $I_{\min} = V_0/(R_0 + R_D + R_{Stot})$ when R_{Stot} is maximum; here V_0 is the *rms* voltage at the oscillator output. We want to have $\Delta I/I_{\min} = (I_{\max} - I_{\min})/I_{\min} = 0.05$, or, as it is easy to see, simply $R_{Stot}/(R_0 + R_D) = 0.05$, or $R_0 + R_D = R_{Stot}/0.05 = 10 \text{ M}\Omega$. Compared to this value, $R_0 = 600 \text{ }\Omega$ can be neglected, and essentially our limiting resistor R_D should be set at $10 \text{ M}\Omega$ (which, incidentally, is the maximum resistance which can be set in our decade box). If we use this large limiting resistor, the current will be equal to $I = 2 \text{ V} / 10 \text{ M}\Omega = 0.20 \text{ }\mu\text{A}$.

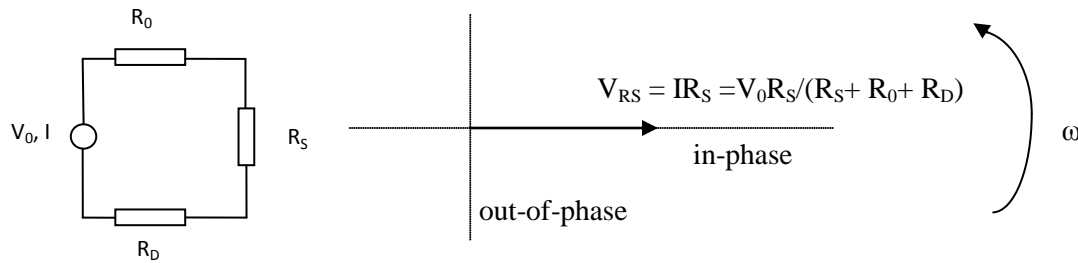
Let us now discuss **the voltage signal**. Small current I will result in a fairly small voltage IR_S , especially when R_S also becomes small. Lock-in can measure small voltages, but below about 1 mV level, the precision will suffer. If we want to increase the current, in the present scheme we will have to sacrifice some additional accuracy in our overall measurement of R_S over the whole range of its variation.

The sample has four contacts: two outside contacts for the current and two inner ones for the voltage. **You should read about four-probe measurement in Appendix 2, and brief description of this principle should be given in your Final Report.** The voltage from the inner leads V_{meas} (see the figure above) is supplied to the lock-in's *measuring input*, where it will be amplified (thus the name lock-in *amplifier*), and provided for the readout on the front panel and, through the digitizing interface, to the computer. The resistance between the inner contacts will be found as $R_S = V_{\text{meas}}/I$. Above, when we were discussing the current, we considered R_{Stot} between the outside current contacts, but now we are measuring resistance R_S between the inner voltage contacts. Lock-in will display the *rms* value as a DC-level signal. The maximum voltage V_{meas} in our example (measuring VO_2) is going to be found at room temperature; it will correspond to the maximum value of R_S . Suppose maximum R_S is equal to $300 \text{ k}\Omega$ (note that we are using here $R_S < R_{Stot}$). Then maximum $V_{\text{meas}} = IR_S = 2 \times 10^{-7} \times 3 \times 10^5 = 60 \text{ mV}$. Such a signal would be quite sufficient, but the trouble is that it will be a lot lower once R_S decreases at high temperatures, where, above the semiconductor-to-metal transition (see below) we may get a signal as small as about 0.1 mV . The lock-in is going to give us poor resolution at that range. Recall also that, with our proposed numbers, we will have a distortion in resistance values of about 5% over the whole range, as our current will increase by 5% when R_S drops by such a large factor.

An alternative way of dealing with changing R_S is to measure the current instead of stabilizing it; it can be best measured by monitoring voltage across the known limiting resistor. Then we do not need such a large limiting resistor; instead of being $10 \text{ M}\Omega$ it can be, for example, $10 \text{ k}\Omega$, and the current will be 1000 times higher, i.e. 0.200 mA instead of $0.200 \text{ }\mu\text{A}$. But we will have to monitor the current and to use the measured current I_{meas} in a calculation of a sample resistance at every point along the way. This can be done with the use of an additional computer-interfaced voltmeter, and with modification of a Lab View program.

Note that the current should not be increased indefinitely; excessive currents can produce self-heating and even damage the thin-film sample. Do not use currents greater than 1 mA with our VO₂ samples.

So, how is the lock-in measuring resistance? The oscillator output, as we said above, provides sinusoidal voltage of *rms* value V_0 at a fixed stable frequency which we can set at will. The readout is tuned to read only voltage at this frequency (actually, at a double frequency), all the other frequencies being filtered out. Additionally, it will read only a specified *phase* of this signal. For our resistive measurement we expect that the specified phase should be zero, as resistor does not produce a phase shift. So, if we are measuring pure resistance, we should be interested in the *in-phase* signal (see Appendix 1). The circuit and the phasor diagram of a pure resistor R_S are shown here:



As always, the phasor should be imagined rotating counterclockwise with ω ; the *in-phase* direction rotates with it; the projection onto the horizontal axis is the real signal. The current is determined by the source voltage V_0 applied across all of the resistors. We see that in any other direction the signal will be smaller than in this *in-phase* direction. Thus, if we will adjust the phase till we see the maximum voltage at the measuring input; this maximum voltage will correspond to the true value of resistance according to $R_S = V_{\text{meas}}/I$. Simultaneously, the 90°-shifted, *out-of-phase* signal should be zero. It is easier to tune the *out-of-phase* signal to zero than to find the true maximum *in-phase*, because when measuring the near-zero signal we can increase the sensitivity. So the usual procedure is to change the phase till one brings *out-of-phase* signal as close to zero as possible and then to shift the phase by 90° displaying the *in-phase* signal. Further, all of the above should be independent of frequency $f = \omega/2\pi$. This is what one expects; however, there are complications.

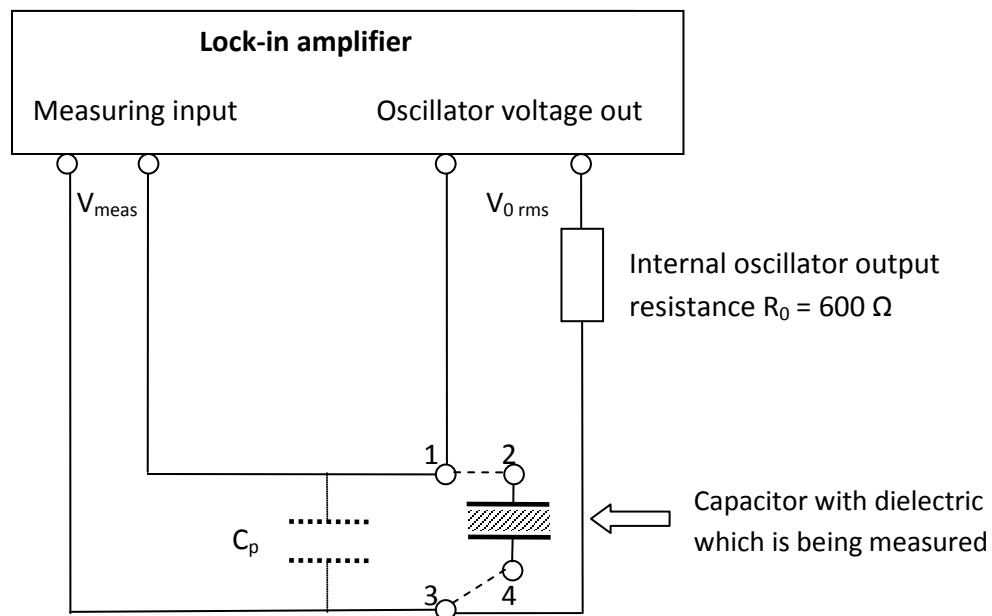
Complications arise because, in addition to the resistance, the circuit contains the so-called *parasitic* or *stray*, capacitance C_p . This is the capacitance of shielded cables, etc. In our circuit it is typically of the order of 10 – 100 pF. This parasitic capacitance can be thought of as being plugged in parallel to the sample resistance R_S , as shown in the circuit diagram above where it is drawn in dotted lines. It will effectively *shunt* (short out) the resistance, leading to incorrect decreased values of V_{meas} .

Solution: Recall that lock-in allows you to adjust the oscillator frequency f , which of course implies that *reactance* $X_C = 1/(2\pi f C_p)$ can be controlled. The solution is then clear: in order to minimize the detrimental effect of C_p we need to decrease f as much as possible: at low f the reactance $1/(2\pi f C_p)$ will be large, and the shunting effect of the parasitic capacitance will be minimized. On the other hand, using extremely low frequency $f \sim 1\text{ Hz}$ may be difficult, as the lock-in tends to lose its ability to lock the phase at extremely low frequencies. Additionally, the so-called flicker or $1/f$ noise increases at low frequencies. But we do not need to reduce f that much: using $f \approx 50\text{ Hz}$ is comfortable in terms of this phase-locking and achieves what we need. Indeed, at $f = 50\text{ Hz}$, for $C_p = 100\text{ pF}$, we find $X_C = 1/(2\pi f C_p) \approx 30\text{ M}\Omega$, which is more than sufficient to prevent a decrease in V_{meas} due to shunting.

Practical advice: insert a known resistor in place of your sample (preferably, with the value similar to R_s) and empirically find the frequency range at which your measurement agrees with that known resistor value. This is an example of a *baseline* measurement which is done to establish the validity of the measurement technique.

How to use the lock-in amplifier to measure the capacitance?

The sample is now a capacitor; the material we want to study (BaTiO_3) is a dielectric inside the capacitor. In this measurement we do not need to worry about additionally limiting the current; the current we are dealing with here is sufficiently limited by $R_0 = 600\ \Omega$. Here we will be using lock-in as a device to measure signals that are *phase shifted* from the phase of V_0 due to the presence of resistors and capacitors in the circuit. The basic circuit diagram for capacitance measurement is as follows:



As in a resistance measurement described above, the signal is generated by the lock-in amplifier's internal reference (oscillator) output; in this measurement we will be setting it at $V_0 = 2.0\ \text{V}$ (actually, at 1.999V), $f = 25\ \text{kHz}$. As before, the built-in internal resistance of the signal generator $R_0 = 600\ \Omega$, and there exists parasitic capacitance C_p shown in dotted lines; this capacitance can be thought of as connected in parallel to sample capacitance.

Surprising as it may sound, we will first do the measurement of an *empty* circuit, without the sample. Later we will insert the sample by connecting sample leads 2 and 4 between points 1 and 3, as shown in the diagram above. The purpose of the measurement without the sample is in tuning the lock-in, so as to make it later correctly measure the sample's capacitance.

First, leaving the sample unconnected, we will tune the lock-in so that it is *in-phase* with the oscillator; it will measure $V_0 = 2.0\ \text{V}$. We then change the phase by 90° and look at V_{meas} while gradually increasing the sensitivity of a measuring readout. On a sensitive-enough scale we will see some non-zero signal due to parasitic capacitance of our "empty" circuit. We will adjust the phase watching the signal on increasingly more and more sensitive scales till this *out-of-phase* signal will be made as close as possible

to zero. So, in other words, we will follow the same procedure as in measuring resistance, as was described above. See Appendix 1 for the details of the phasor diagrams in this case. Once we zeroed the out-of-phase signal, we will insert the sample, so that points 2 and 4 are connected to 1 and 3; once we have done that, the *out-of-phase* readout from the sample V_{meas} will appear at the measuring input and the magnitude of that *out-of-phase* signal is going to be proportional to the capacitance of the sample. Indeed, the capacitance of our sample is usually small, of the order of 100 pF – 1 nF, and so, with $R_0 = 600 \Omega$ and at our chosen frequency, we have $\omega R_0 C \ll 1$. In this limit the *in-phase* component of the signal will be very close to the voltage provided by the signal generator, V_0 , and the *out-of-phase* component will be close to $V_0 \omega R_0 C$ (see Appendix 1 for the details). By measuring the *out-of-phase* component V_{meas} , knowing V_0 , ω and R_0 , we can calculate the capacitance of the sample as

$$C = \delta / (\omega R_0)$$

where $\delta = |V_{\text{meas}}| / V_0$.

Lock-in amplifier settings

The standard settings of the EG&G model 5209 lock-in amplifier are:

input: A, sensitivity: 3V, filter: flat, track, osc. lvl: 1.999V, osc f: 25kHz, ref ph: tuned (see later), reference: int, osc out: BNC cable connected, time constant 300ms, dyn res: high res, output display: signal



Before measuring the sample, we do a practice run by measuring the capacitance of a piece of BNC cable. See Appendix 1 for relevant phasor diagrams.

First we check the signal generated by the internal source of the lock-in. Connect one end of a BNC cable to the “osc out” connector, and the other end to input A. In the “tuning” section scroll until the display

shows the reference phase, “ref ph”, and set it to 0.0. The analog meter and the output display on the right should show a voltage slightly less than 2V.

We are going to be interested in the out-of-phase component, and we can get that by pressing the 90 DEG button once. Ideally, the out-of-phase signal should be zero. However, with all the electronics inside the lock-in, keeping the phase of the output voltage under control is not an easy job and there will be a small but finite out-of-phase signal.

The engineers designing the lock-in provided a way to correct any phase error: Run an *autotune* cycle. First press the red “auto” button, and then press the button slightly up and left. Wait a few seconds, and the lock-in will find the best phase so that the in-phase signal is maximum, and the out-of-phase signal is zero. To check the out-of-phase, press the 90DEG button again and you should see zero.

Upon closer inspection you may discover that the autotuning is not perfect. Change the sensitivity from 3V to 300mV, and you will see that there is a finite signal. You should do a manual fine tuning of the phase, so that this signal is as close to zero as possible. You may increase the sensitivity to 100mV and continue tuning the phase. Do not increase the sensitivity further, because the input of the lock-in will be overloaded. **If you see the red overload sign flashing, change to a less sensitive setting. Never set the lock-in to sensitivity of 10mV or higher.**

Now that you have the out-of-phase voltage at zero, select the sensitivity of 300 mV again. Insert a BNC “T” connector to the line and connect another BNC cable. Suddenly, there will be a finite *out-of-phase* signal V_{meas} . Calculate C , using $C = (V_{\text{meas}}/V_0)/(\omega R_0)$ with $V_0 = 2.00\text{V}$, $\omega = 1.57 \times 10^5 \text{ s}^{-1}$ and $R_0 = 600\Omega$. Measure the length of the cable. Test a longer cable.

Q: In round numbers, what is the capacitance of one foot of BNC cable?

7. The rest of the experimental apparatus in detail

7.1 Sample chamber

The photograph of the sample chamber is shown below. There are four wires reaching the inside of the chamber, “red” and “black” pairs. The sample will be either a film of VO_2 with four resistive leads, or a capacitor with BaTiO_3 requiring only two leads attached to its plates. If two students do the measurement, it is recommended that each one of them mounts his/her own sample and evaluates the data independently. However, VO_2 samples are fragile, and a staff member should be asked to mount them in the chamber, at least initially.

A heater is wrapped around the aluminum block of the sample chamber. The platinum resistor for temperature control is visible in the upper right. There is another platinum resistor (barely visible in the photo) in the center of the sample

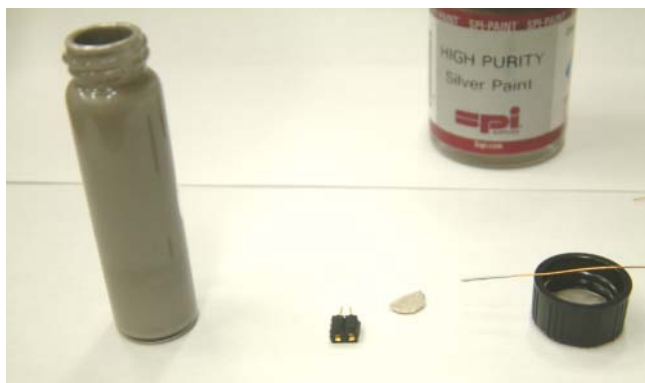


chamber. That resistor is used to measure the temperature of the samples.

In normal use, the sample chamber is covered by an aluminum cup (see the upper right corner of the photo). The whole assembly is covered by a double-walled stainless steel cover for thermal insulation. A double-walled evacuated vessel is called a *dewar*. If you wish to cool the sample faster, remove the stainless steel *dewar*. You may also remove the aluminum cup, but in that case stop taking data, since the temperature readout will not be reliable.

7.2 Samples and sample preparation

Samples of VO_2 in a form of thin films (about 100 nm thick) have been deposited on Si/SiO_2 substrates in another Lab on Campus. Contacts to them will be attached by staff members. These samples are rather fragile, especially those contacts, and should be handled with great care. Ask a staff member to plug the sample into the holder, at least initially.

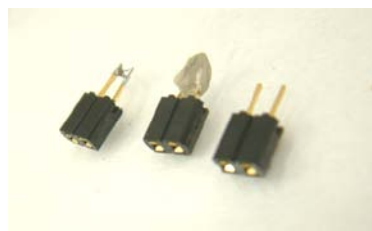


The BaTiO_3 sample is from a commercial distributor. It comes in ceramic form with many different, randomly oriented microcrystals. Feel free to split off your sample from larger chunks of the ceramic blocks. Pick a flat and thin piece of ceramic BaTiO_3 , and shape it to a thin plate by lightly sanding on the sides. Keep in mind that capacitance $C = \epsilon A/d$: the thinner is the better and the larger area is the better. Also remember that you will need to estimate A and d numerically, and so the more regular is the sample shape, the easier and more accurate will

be this estimate. Use silver paint to create the capacitor plates and mount the sample (with the silver paint) onto the two-terminal connector shown in the middle of the Figure. The best way to do this is by placing the sample between the two pins and using the silver paint to create the capacitor plates. The silver paint should be applied by a piece of thin wire (visible in the upper photo). Be careful not to paint over the edges because that creates a short circuit and ruins the sample. Remember that you are making a capacitor.

The silver paint used for the sample preparation dries very quickly. Do not leave the container open for an extended period of time; open it for as short an interval of time as possible, just to get the paint onto a piece of wire.

The photo on the right shows an empty two-pin connector (right) a mounted sample (center) and a 100pF test capacitor soldered to the connector (left).



A stereo microscope is provided for mounting the sample, although with good eyes and eye-hand coordination it can be done without the microscope.

The sample of BaTiO_3 (but not of VO_2) you make for the studies should be returned with you final lab report. Simply attach the sample with Scotch tape to the first page of your report.

Connect the experimental apparatus: one BNC cable from the signal source to input A of the lock-in, and another cable from the lock-in to the BNC leading to the sample. Check that there is no sample in the sample chamber. (For later use, when you put in the sample, you need to terminate the other BNC with a short circuit to ground, as shown in the photo).



The *out-of-phase* signal you see now is due to all of the capacitances leading to the sample, including the capacitance between the teflon-insulated black wire and the ground. Do a fine tuning of the phase so that this signal is zero. Switch to sensitivity of 300 mV and insert the sample. The signal you see now is entirely due to the sample, and you can use $C = (V_{\text{meas}}/V_0)/(\omega R_0)$ again, to calculate the capacitance of the sample. (If the signal is small, you may need to change the sensitivity to 100mV.)

For all of this to work $V_{\text{meas}}/V_0 = \omega R_0 C$ should be much less than 1. In Appendix 1 we derived the exact formula for the calculation of the capacitance, and we noted that the difference between the exact and the approximate formulae is about 1% when $\delta = 0.1$. Therefore, for $V_0 = 2\text{V}$ the systematic error caused by using the approximate formula will be less than 1% if V_{meas} is smaller than 200 mV. If you made a very thin sample and the capacitance is large, you may go outside of this limit. When that happens you should either use the exact formula or reduce the frequency to 10 kHz or 5 kHz and repeat the tuning.

At this point you are ready to measure the temperature dependence of the capacitance. Keep in mind that we expect the capacitance to increase as we warm the sample, and select a range where the current signal is less than one third of the full range. This way you do not have to change sensitivity when the most interesting part of the measurement is happening.

7.3 Temperature control

The “guts” of the temperature controller⁽³⁾ are shown in the Figure. The “brain” is an Omega CNi8DH44 unit, connected to a solid state relay (on the left) and other circuitry, like fuses and connection pads. The solid state relay converts the low-power output of the converter into high power that drives the heater wrapped around the sample chamber (see the photo of the sample chamber above). The input for the temperature controller is a platinum resistor that is placed in the aluminum block of the sample chamber. The resistance of platinum (and all other metals) depends on the temperature, and this resistor has a resistance of 100Ohm at temperature 0C. For platinum, a 1C increment of temperature the resistance changes by 0.38Ohm.



The simplest way to warm the sample to a pre-set temperature and keep it there is the “on/off” control mechanism. You specify the *set point* (or *setpoint* in one word), which is just the highest temperature needed, and the controller starts to heat the sample chamber with full power. If the *ramp* is specified, the controller will slow down this heating rate by lowering the power to keep the specified ramp. This is important in order to obtain good data. You should experiment with the ramp rate and settle at the acceptable rate which you will judge by the quality of your data. Once the temperature at the control thermometer reaches the set point, the heating power is switched off. It is most likely, however, that at this moment there is an amount of excess heat in the system, and temperature keeps climbing (this is called an “overshot”). After a while the readout of the control thermometer starts decreasing, and crosses the *setpoint* from above. The “on/off” controller switches “on”, but by his time the rest of the sample chamber is too cold, and the temperature keeps dropping for a while. Then the temperature *overshoots* again (in the opposite direction) and the process goes into a cycle. The desired temperature will be never reached; instead there will be an oscillation around it. To fix this problem the temperature controller is programmed to use “PID” control, with the P, I and D parameters optimized for this particular sample chamber. **Read about the PID control in Ref. 3.**

Q: Be ready to answer this question: What is the role is the “I” in PID?

Note: Most of the capacitive measurement is done at temperatures over 100 C. Be careful, the sample chamber will be hot. Do not exceed 160C, because the sample chamber will be damaged.

The sample chamber can be seriously overheated if left alone with the heater “on”. Always switch off the temperature controller when the measurement is done. If you are using the “ramp” mode remember to make sure that “soak” is also on for a short time to ensure that the heater shuts off once the desired temperature is reached; more on this in Appendix 3.

7.4 Temperature measurement

In measurements of temperature dependent properties it is a good practice to have two thermometers: one to control the temperature (described above) and another one to measure the actual temperature of the sample. This second thermometer is often directly attached to the sample. In our case, for the sake of easy sample changes, the thermometer is simply inserted to the center of the sample chamber. This thermometer is also a platinum resistor.

We use a Keithley 199 multimeter in four probe resistivity mode to read out the resistance of the temperature probe. Learn about the four probe method in **Appendix 2.**

Q: Be ready to answer this question: What is the advantage of the four-terminal method over the two terminal method, and how is this advantage achieved?

The multimeter is set to “Ohm” mode and the four leads should be connected to the four terminals. (Note the two red leads go to the same side of the resistor, and the two black leads go to the other side. For the extension wires the reds are connected to browns and the blacks are connected to whites.)

When the computer reads out the resistance values, it will convert them to temperature according to the (approximate) $T = 100C + (R - 100\Omega) / 0.385$.

The two data files contain: 1) the time dependence of the temperature and 2) the temperature dependence of the lock-in output voltage. It is expected that you transfer the relevant data files to a thumb-drive to carry out the evaluation of the data.

Before measuring the sample, **study the temperature control process**. Start at room temperature or at a temperature below 50C. Select a set point of 70C. Record the time dependence of the temperature. Put the temperature controller to “on/off” mode and keep recording. Temperature oscillations will develop. Put the temperature controller to “PID” mode. Discuss these observations in your lab report.

8. Suggested protocol for a measurement

- 1.) Tune the lock-in, and set up the sample (VO_2 or BaTiO_3) in the sample chamber. Cover the sample chamber with the stainless steel cover (the *dewar*).
- 2.) Start the Ferro.vi LabView routine
- 3.) Start saving data in LabView
- 4.) Program the “Set point 1” of the temperature controller to 90.0 C for a VO_2 measurement and 150.0 C for the capacitance BaTiO_3 measurement.
- 5.) Program the ramp rate. Change the ramp rate and examine the data: you want to make sure that your ramp rate is slow enough for the data to be reliable. Discuss with the laboratory Staff what this means.
- 6.) Watch the signal as the sample crosses the phase transition. In VO_2 this will take place over a fairly wide temperature range; in BaTiO_3 , in a narrow temperature range.
- 7.) Once the setpoint is reached you will take more data while the sample is cooling.
- 8.) If you are not in ramp/soak mode, program the lower set point of the temperature controller to 10.0 C
- 9.) The sample will slowly cool while the data is recorded. In the BaTiO_3 measurement, nothing interesting happens below 80C, so when the temperature reaches 80C you can take off the *dewar* to speed up the cooling. When measuring a VO_2 sample, you want to take the data in a slow-enough ramp mode in the whole temperature range.

9. Advice for evaluating the capacitance data

From the graph of the measured capacitance as a function of temperature you will find the T_c for this material.

Estimate the area of the capacitor and the distance between the plates and calculate the relative dielectric constant k at room temperature. To find k , you will need to calculate the capacitance C_0 of an empty (i.e. filled with vacuum or air) capacitor with the *same geometry* as your sample, and then take your measured C and divide it by that calculated C_0 : $k = C/C_0$.

Give an estimate of the error in k due to the necessarily inaccurate measurement of the area and the distance between painted plates. This error is expected to be large. Compare the value of k to the known dielectric constant of this ceramic material.

Our main interest, however, is to use the data from our ceramic sample to access some of the intrinsic, fundamental properties of BaTiO_3 . This means that your values of k will need an important **correction**, which we discuss here. If the dielectric constant is very large, the capacitance of the sample should be very large too. However, in this measurement, the capacitance will not increase near the transition that much. A possible explanation is that in a ceramic sample the dielectric polarization is interrupted at the

grain boundaries. The capacitance between the grains provides coupling, but that inter-grain capacitance is not changing with temperature, and certainly does not diverge at T_c . You can model this behavior by taking the sample as two capacitors in series: C_b (capacitance of the grain boundaries) and C_{BTO} (capacitance of the bulk grains of $BaTiO_3$). For the effective (measured) capacitance in this case we have $C = [C_b^{-1} + C_{BTO}^{-1}]^{-1}$. Here C_b is independent of the temperature and C_{BTO} corresponds to an ideal $BaTiO_3$ uniformly-filled capacitor. In this model you can assume that at the phase transition point, when C_{BTO} is very large, C_{BTO}^{-1} term can be neglected, and all you measure is C_b , $C \approx C_b$. Thus, from your measurement of the peak capacitance at T_c , you will know C_b . For all the other temperatures you can calculate C_{BTO} as

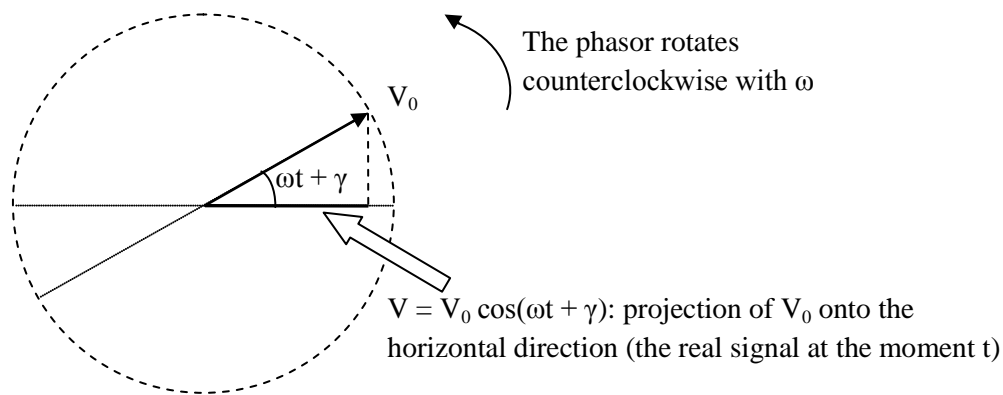
$$C_{BTO} = C_b C / (C_b - C) = C / (1 - C/C_b)$$

where C is the measured capacitance and C_b is the maximum (peak) capacitance at T_c . Therefore, **you should use this corrected C_{BTO} and not the measured capacitance C directly in finding k values: $k = C_{BTO}/C_0$.**

Once you obtain corrected k vs. T data, refer to Appendix 5 (which you should be reading continually while doing this experiment; it contains most of the Physics), analyze Curie-Weiss dependence of k in the vicinity of the transition, and find Curie-Weiss constant C_{CW} . From it, calculate the microscopic dipole moment and the displacement of Ti ion in the unit cell based on your data. Compare it to the displacement you will find in the literature (you will have to find this value, which is known from X-ray measurements of $BaTiO_3$). Also calculate the same quantities using the *intrinsic* (single-crystalline) value of C_{CW} from the literature.

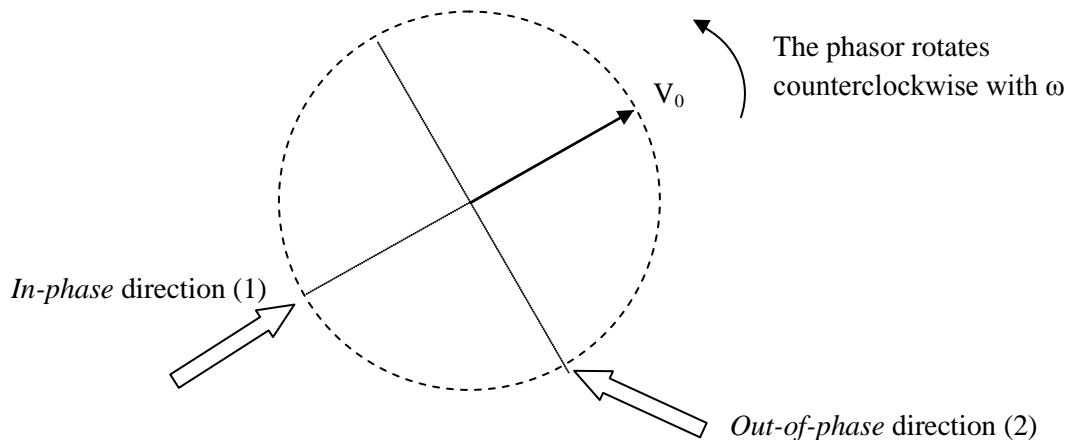
Appendix 1: Elements of AC circuit theory relevant to our measurements using lock-in amplifier

We are dealing with sinusoidal AC voltages and currents, such as $V = V_0 \cos(\omega t + \gamma)$, where V_0 is the amplitude, ω is an angular frequency, and $V(t=0) = V_0 \cos(\gamma)$ is the initial condition, i.e. a signal at $t=0$; γ is called the *phase angle*, or the *phase*. You will recall that on a *phasor diagram* which you studied in your first Physics course, all signal amplitudes are represented as rotating vectors; ω is the angular velocity of a rotating phasor related to regular frequency by $\omega = 2\pi f$; projections of these vectors onto a chosen axis are real instantaneous signals. A *cosine* function used above corresponds to a projection onto a horizontal axis; if we used a *sine* function, it would represent a projection onto a vertical axis. In the figure below we show V_0 at the moment of time t , when it makes an angle $\omega t + \gamma$ with the horizontal; we see V_0 's projection onto the horizontal direction $V_0 \cos(\omega t + \gamma)$, which is the real signal at the moment t :



We will recognize that a *phase angle* also specifies a *direction* on a phasor diagram; there is an infinite number of possible radial directions in a circle.

It is convenient to measure all phases with respect to a phase of a source, V_0 . A zero phase angle then corresponds to the direction of V_0 ; we will call it in what follows the *in-phase* direction. A direction perpendicular to the *in-phase* is called *out-of-phase*. It is clear that projections of V_0 onto the *in-phase* and *out-of-phase* directions are V_0 and zero respectively.

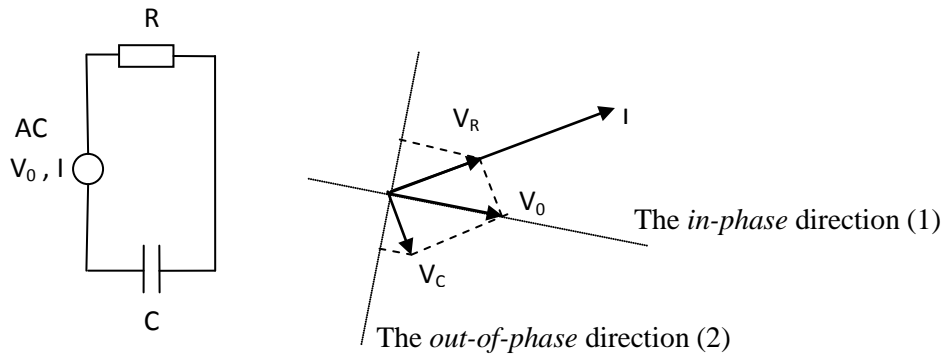


Deciding to use these two phases (or these two phase directions), we effectively setting our initial γ to zero. But we will be eventually measuring average rms values which do not depend on the initial phase; at the same time, all *relative* phases will be preserved in this new system of phase referencing.

Measuring a specified phase of a signal (or “*looking*” at a phase, as we say in electronics) is equivalent to *looking* at a signal projection onto a chosen direction in a phasor diagram. While we chose two such directions, *in-phase* and *out-of-phase* with respect to the source voltage V_0 , lock-in is capable of *looking* at any phase direction. This is one of its advantages over a normal voltmeter.

With this, let us now consider a phasor of an RC circuit. Below we show a phasor of a series connection of R and C, with AC source supplying voltage V_0 ; it is a vector sum of voltages V_R and V_C . For completeness we also show the current I, which is established in the circuit; it is shown pointing in the direction of V_R . The current is shown only once because it is the same for all elements in a series connection.

Note that current is said to be *in-phase* with the voltage across the resistor, V_R . We should be careful not to mix our *in-phase* direction along V_0 as it was defined above with the *in-phase* direction for the current; these are two different phase directions, as can be seen on the diagram below.



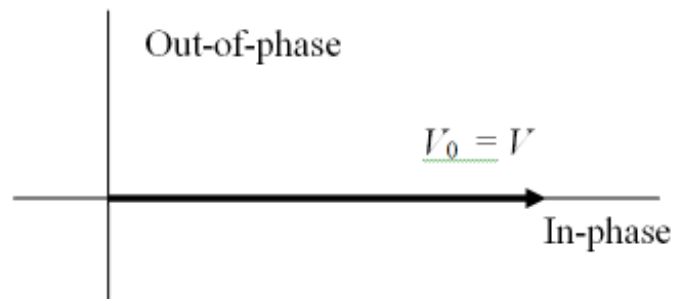
It is clear that projections of V_R and V_C onto both *in-phase* and *out-of-phase* directions are non-zero, while the algebraic sum of projections of vectors V_R and V_C onto the *in-phase* direction is V_0 and to *out-of-phase* direction is zero. Projections of V_R and V_C constitute signals measured with the lock-in amplifier across R or C respectively, either *in-phase* or *out-of-phase*.

Measuring signals in a particular phase as described above is called “phase-sensitive” detection. We are not discussing here how it is practically implemented in the inside electronics of a lock-in amplifier, but only its meaning in terms of a phasor diagram.

Measuring resistance is discussed in the main text in section 6; we need to watch for the detrimental effect of a parasitic capacitance, and the practical solution is to measure at low-enough frequency.

Measuring capacitance: Here is an illustration of the lock-in tuning process in terms of phasor diagrams:

Initial tuning: The source voltage V_0 is directly measured $V = V_0$. The phase is



adjusted so that there is no *out-of-phase* component.

When the capacitor is added to the circuit, the voltage drop on the resistor $V_1 = V_R$ and the voltage drop on the capacitor $V_2 = V_C$, via vector addition, add up to the source voltage V_0 . There is a 90° phase difference between V_R and V_C . The phasor diagram looks like this:

The phase angle α is given by

$$\tan(\alpha) = V_R/V_C = \omega RC.$$

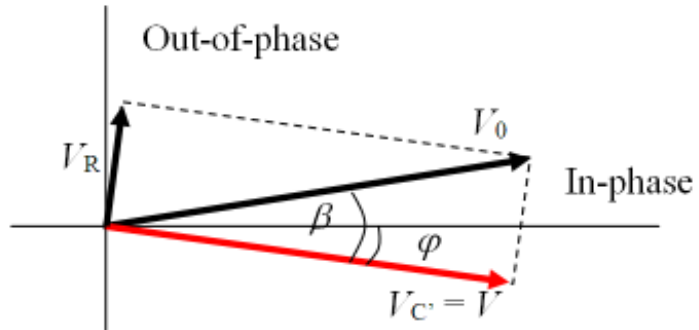
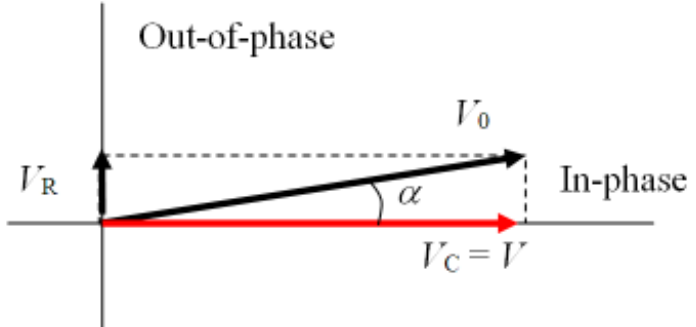
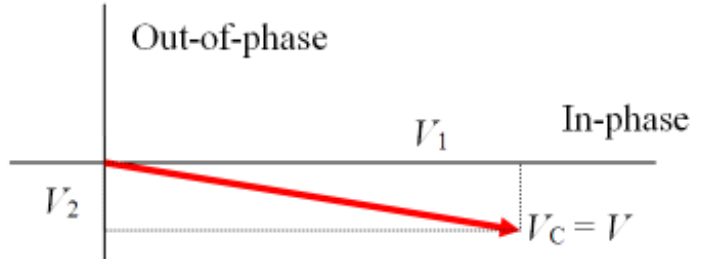
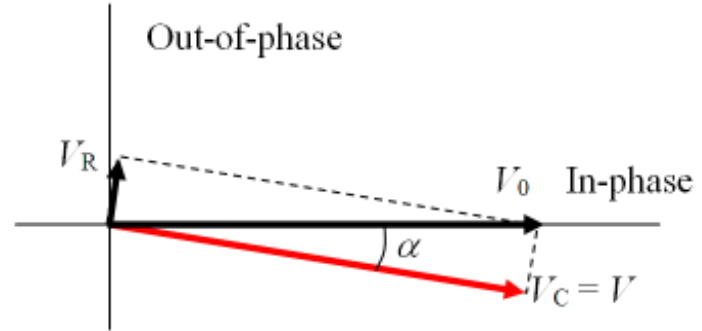
For $\omega = 2\pi f = 2\pi \times 25\text{kHz}$, $R_0 = 600\Omega$ and $C = 100\text{pF}$ we get $\tan(\alpha) = 0.0094$. For such small angles (in radians!) we can safely take $\tan(\alpha) = \alpha = \omega RC$.

For $V_0 = 2.0\text{V}$ the measured out-of-phase voltage is $V_2 = V_C \sin(\alpha)$. For small α , $V_C \approx V_0$ and $V_2 = -V_0 \omega R_0 C = -19\text{ mV}$.

Connecting the measurement circuitry is like adding a capacitor to the circuit. Here $C = C_p$ represents the sum of all parasitic capacitances. Now we tune the lock-in so that the *out-of-phase* component is zero again. The angle α is still small, so $\tan(\alpha) = \alpha = \omega R_0 C_p$

Finally we add the sample. The capacitance gets larger: $C' = C_p + C$, where C is the capacitance of the sample. Since the larger capacitance results in smaller impedance, the current in the circuit increases, and the voltage drop on the resistor increases. The phasor diagram looks like this, where $\tan(\beta) = \omega R_0 C'$.

The measured out-of-phase signal is $V_2 = V_C \sin \varphi$, where $\varphi = \alpha - \beta$. As long as all of these angles are small, we get $V_2 = V_0 \omega R_0 C$, where C is the capacitance of the sample. Therefore



$$C = \frac{\delta}{\omega R}$$

where $\delta = -V_2/V_0$ (δ is a positive quantity, and, of course, so is the capacitance). Here $V_2 = V_{meas}$.

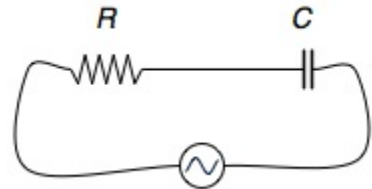
The capacitance of the typical sample is in the 100pF range. For 100pF, at 25 kHz and $R_0 = 600\Omega$ the quantity $\omega R_0 C$ is much less than 1. Therefore in this limit

$$V_1/V_0 \sim 1$$

$$V_2/V_0 \sim -\omega R_0 C$$

$$C = -(V_2/V_0)/(\omega R_0) = \delta/(\omega R_0)$$

Alternative derivation: The same result can be obtained using *complex numbers* instead of phasors. Complex numbers and symbols for treatment of AC circuits were introduced in the “Primer to Linear Circuits and Transmission Lines”, posted with the Course Documents on Blackboard. The impedance of the capacitor is $Z = 1/j\omega C$. The current in the RC circuit shown on the right is $I = V_0/[R + 1/(j\omega C)]$, and the voltage on the capacitor is:



$$V = V_0 \frac{1/j\omega C}{R + 1/j\omega C} = V_0 \frac{1}{1 + j\omega RC} = V_0 \frac{1}{1 + (\omega RC)^2} - jV_0 \frac{\omega RC}{1 + (\omega RC)^2}$$

The real and imaginary parts are

$$V_1 = V_0 \frac{1}{1 + (\omega RC)^2} \quad \text{and} \quad V_2 = -V_0 \frac{\omega RC}{1 + (\omega RC)^2}$$

With the proper setting of the lock-in amplifier, described in the main text, we can measure V_2 . Since we know V_0 , ω and R , we can calculate the capacitance of the sample.

$$C = \frac{1}{\omega R} \frac{1 - \sqrt{1 - 4\delta^2}}{2\delta}$$

where $\delta = -V_2/V_0$ (since the measured voltage is negative, δ is a positive number).

So, to summarize, when $\omega R_0 C$ is small ($\ll 1$) we can use formula $C = \frac{\delta}{\omega R}$, but for a very large and thin sample, capacitance will not be small, and one has to use the complete expression derived above. A simple calculation shows that the difference between the exact and the approximate formula is about 1% when $\delta = 0.1$ and about 5% when $\delta = 0.2$.

Appendix 2: The four-probe method: advantages and limitations

A simple resistance measurement is performed either by measuring both the current through the resistor and the voltage across the resistor (I and V measurement) (Fig.1),

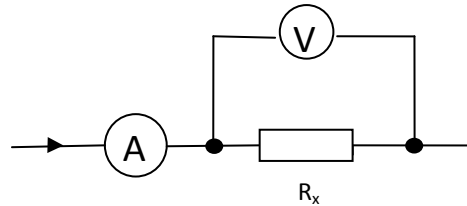


Fig. 1 Two-terminal resistance measurement

or by measuring a current, supplied by a battery, which flows through an adjustable resistor and the unknown resistor connected in series, with the scale being set as to read the unknown resistance (an *ohmmeter* measurement, which we do not show here). Either way, it is a *two-terminal* measurement: the unknown resistor is connected to the rest of the circuit at two points, or two terminals, shown in Fig. 1 as black dots. In this measurement it is implicitly assumed that *contact resistance* and resistance of the wires is negligible compared to the resistance R_x . In many cases, especially when dealing with commercially-made resistors, this is an excellent assumption.

There are however many experimental situations in which this measurement is complicated by the presence of finite, non-negligible *contact resistance*, and sometimes also by the presence of *lead resistance* (resistance of electrical connections, wires). Consider for example a measurement performed on a sample shaped as a uniform-width strip (for example, it may be a thin film sample). When performing a two-terminal measurement, we place two metal contacts at the ends of a film strip and attach to them current and voltage leads, as shown in Fig. 2:

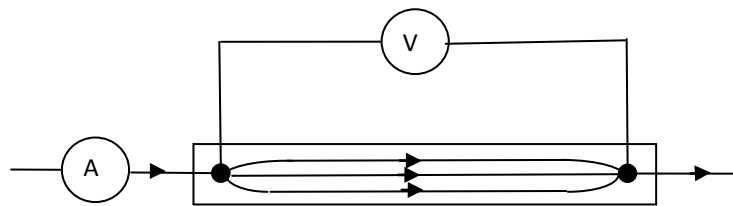


Fig. 2 Two-terminal measurement on a sample shaped as a strip, with two contacts

There are at least two problems with this arrangement:

- 1). If there is appreciable contact resistance between the film and the contacts (black dots), the equivalent electrical circuit describing this measurement will look like the one depicted in Fig. 3, where resistors R_{cont} indicate the presence of said contact resistance:

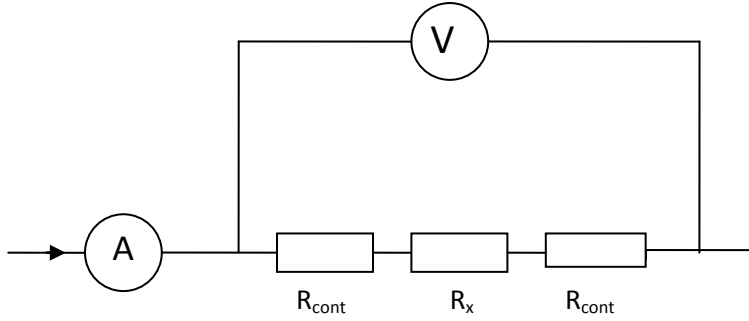


Fig. 3. Equivalent circuit showing presence of contact resistance in a two-terminal measurement

If R_{cont} is not negligibly small compared to R_x , voltage drops across them, $2IR_{\text{cont}}$ will be included in the voltmeter measurement. Note that here the full sample current I flows through these R_{cont} .

2). Current flow lines in Fig. 3 may be non-uniform around the contacts. This makes the length of a sample ill-determined, and thus will introduce an error in the calculation of *resistivity*, even if contact resistances were negligible.

We note that the same un-avoidable problems exist in an *ohmmeter* two-terminal measurement.

These problems with the two-terminal connection can be overcome by making a *four-terminal measurement* described below. This measurement is known under different names such as *4-probe method*, a *4-terminal sensing*, a *4-probe geometry*, or as *Kelvin sensing* (the latter name is rarely used today; it acknowledges that a similar method was first described by Lord Kelvin a very long time ago). It is universally used in Solid State Physics to perform accurate measurements of true material parameters, primarily of resistivity ρ , as well as in other situations when contact resistance needs to be excluded from the measurement.

The idea is simple: make current and voltage contacts separately. The current is measured, and so the results will not depend of current contact resistance. The two inner voltage leads detect *potential difference* between them. The current practically does not flow into these inner leads, as the voltmeter resistance is very high. So again, their contact resistances do not enter into the picture, except for the case when these resistances are becoming very large, namely, comparable with the input resistance (impedance) of a voltmeter.

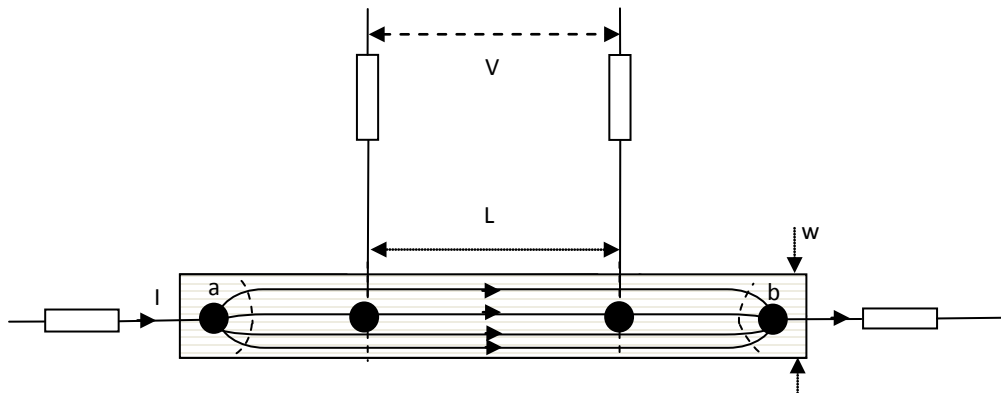
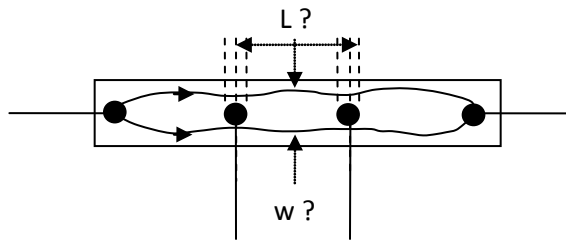


Fig. 4 The basic 4-probe method

If current is I and the potential difference between the inner leads is V , then the resistance of a sample which lies between the voltage leads is simply $R_S = V/I$. Now we can find resistivity ρ from the formula $V/I = R_S = \rho L/(w\delta)$ or $\rho = R_S (w\delta/L)$, where w is the sample width, δ is its thickness, and L is the distance between the voltage leads. Here we assume that we know or can measure sample geometry (uniform width w , uniform thickness δ , distance between voltage leads L).

As we said, this method was designed to exclude from the measurement contact and wire resistances, and it always achieves this goal once we separate the two sets of contacts. However, the question of how accurately one can find the absolute value of ρ depends somewhat on the placement and shape of the middle voltage contacts. If we have four contacts which are essentially large dots placed along the sample (as is shown in the figure above, and is the case in this lab when we measure VO_2), we will not be able to find ρ exactly. The reason is that our Fig. 4 is somewhat idealized in that current may not flow along the sample in such a perfect uniform fashion, being somewhat disrupted by the middle voltage contacts; additionally, we may have a difficulty measuring L precisely, as it is essentially known to the uncertainty in the contact dot size.

This is illustrated below in Fig. 5:

**Fig. 5** Imperfections associated with placing contacts on the sample surface

We will ask you to roughly estimate resistivity of VO_2 sample, despite these imperfections. The thickness δ will be given to you. You may not have perfectly uniform w , but you can estimate average w between voltage leads; and you can roughly measure L between them as well.

We note that one can further improve the 4-probe method in order to accurately measure ρ by moving inner voltage leads out of the way of the current and having well defined sample width w . We will not address these improvements here.

Finally, how much of a voltage contact resistance can be tolerated before we run into a problem with this measurement? The answer is that resistances in the voltage contacts will become important if they are comparable or exceeding the impedance of our voltmeter (in our case, the lock-in amplifier). The input impedance of a voltmeter can be anywhere from 10 Mohm to a Gohm (giga-ohm); one should consult with the instruction manual of a particular instrument one is using.

To understand why this is so, recall that a real voltmeter does need a tiny current for its operation. The last statement applies to both analog and digital voltmeters: whether the current is needed to move the *d'Arsonval galvanometer* in an old analog meter, or to charge a capacitor in a digital one, it is

fundamentally needed nonetheless. (The only voltage-measuring instrument which does not need a current is a *potentiometer*, which is not practical in the type of measurements discussed here). That is why all voltmeters have finite – not infinite – impedance R_{imp} (internal resistance): they must admit some current. So, a voltmeter is essentially a current-sensing instrument converted to measure a voltage. It fundamentally comprises a current-sensing element (such as a galvanometer G) in series with a very large resistor R_{imp} , as depicted in Fig. 6:

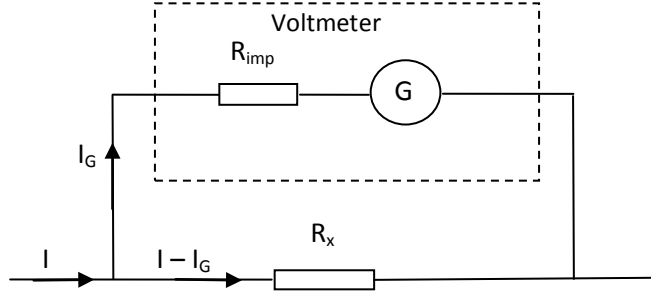


Fig. 6. Recalling how a voltmeter works

We can see that $(I - I_G)R_x = I_G R_{\text{imp}}$, and thus $I_G = (I - I_G)R_x / R_{\text{imp}} \approx IR_x / R_{\text{imp}} = V_x / R_{\text{imp}}$, where we neglected $I_G \ll I$. The instrument, while measuring I_G , is calibrated to read voltage across the sample V_x , which is proportional to I_G .

Now, suppose that voltage contact resistances are comparable to R_{imp} . The corresponding electrical circuit showing these unknown contact resistances R_{cont} is shown in Fig. 7:

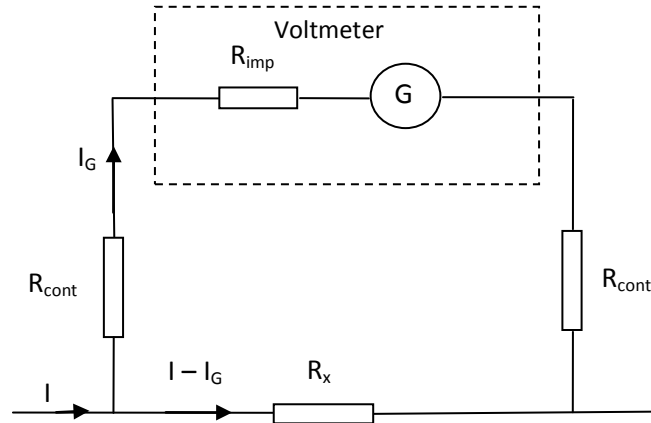
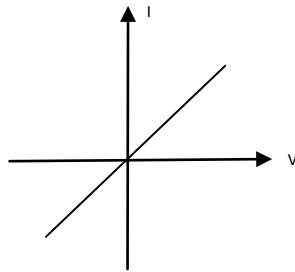


Fig. 7. A voltmeter measuring a sample with large contact resistances

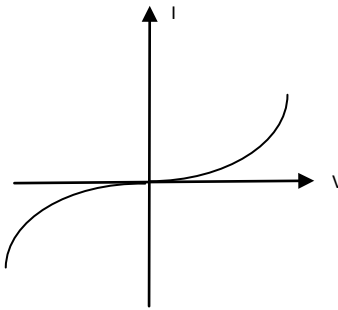
Here we have $(I - I_G)R_x = I_G (R_{\text{imp}} + 2R_{\text{cont}})$, or, again neglecting $I_G \ll I$, $IR_x \approx I_G (R_{\text{imp}} + 2R_{\text{cont}})$, or $I_G = V_x / (R_{\text{imp}} + 2R_{\text{cont}})$. Since R_{cont} is not known, the voltmeter will measure incorrect I_G and thus it will display an incorrect V_x .

Note the difference between this circuit and the one depicted in Fig.3: here a very small current I_G flows through contact resistances; there, in a 2-terminal measurement, the full current I was flowing through these resistances. This is the main advantage of a 4-terminal over 2-terminal measurement.

Generally it is rather difficult to make good contacts to high-resistivity semiconductors. Contacts to semiconductors and avoidance of contact problems constitute an active field of study, with many publications devoted just to this subject. One characteristic of a good contact is that it is *ohmic*, that is, it has a linear I-V curve:



One often finds that high-resistance “bad” contacts are usually also *non-ohmic*, having a non-linear I-V curve, for example like this one:



As one can see, here not only R_{cont} is unknown, but it also is not well-defined, depending on the current I_G that is being passed through it. If in doubt about a contact, measure its I-V characteristic, and stay away from the ones with a non-linear I-V.

Practical advice: when preparing to measure the VO_2 sample, first check its contacts with a simple 2-probe ohmmeter. If any of the pairs from the four contacts turn out to have high resistance (say, of the order of several Mohms), report this to the TA or to the Instructor, for you will probably not be able to measure such a sample with the four-probe method.

Appendix 3: Programming the temperature controller

The temperature controller can be programmed from the front panel (see the manual) or by using its WEB interface. Front panel programming is cumbersome and not recommended. There are two ways to access the WEB interface:

If you want to change the setpoint only:

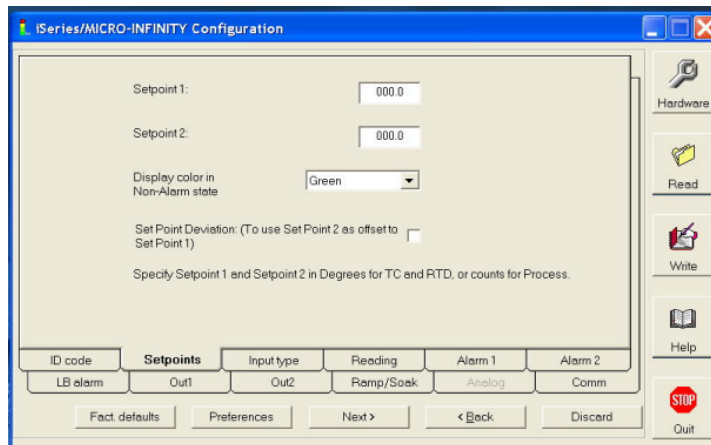
Open “Shortcut to internet” on the Desktop (Internet Explorer). The home page is served by the temperature controller (the IP address is 192.168.1.200). An “I Server” page will appear, with Device Type Selection. No need to change anything, just click on “Update”. On the next page (“Server Home Page”) click on “Read Devices”. The current temperature readout of the controller should appear. Click on “1.” Right before “Omega” on the screen. A login screen will appear, enter 12345678 for password. You reach the “Device Setpoints” screen.

Change “Setpoint #1” and press update. The temperature controller should show the new setpoint in the lower left side of the display.

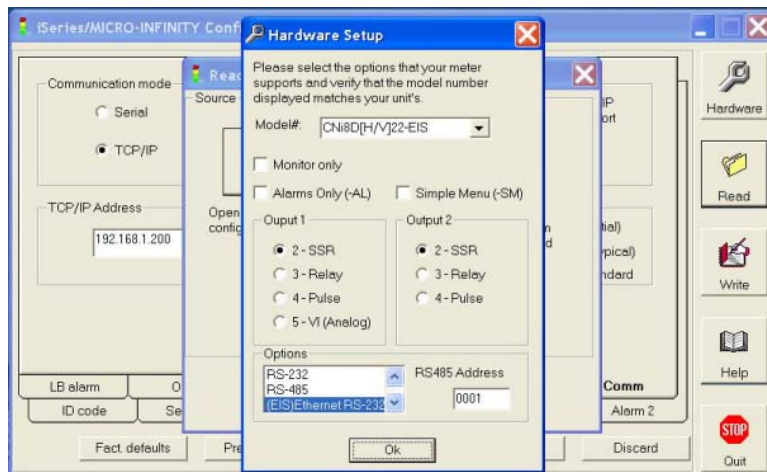
As long as you keep the window open you can change setpoints many times without going through the login process.

If you want to change other settings:

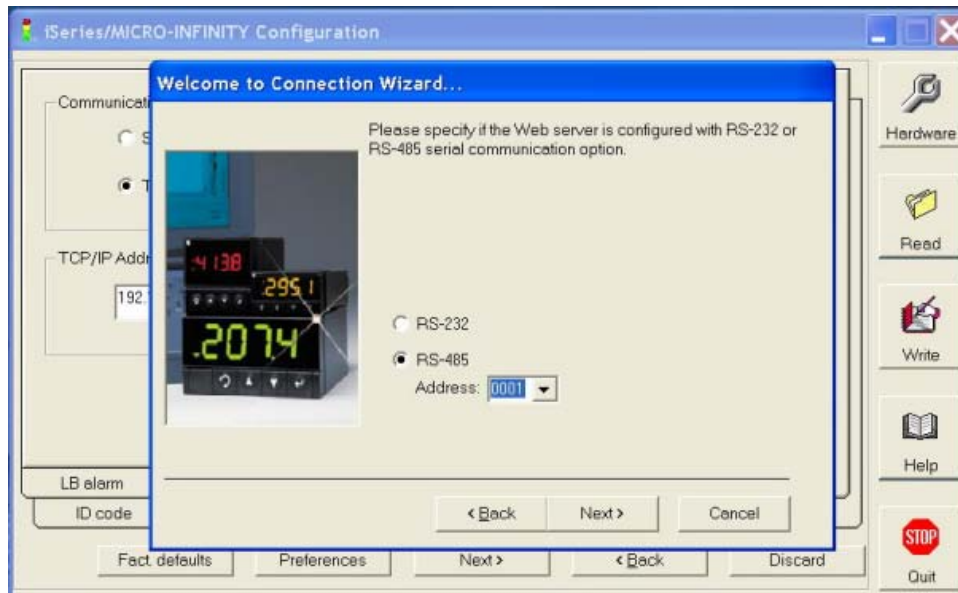
Click on the “iSeries and MICRO INFINITY Config” icon on the desktop. The opening screen will look like this (next page):



Click on “Read” and on the next screen click on “File”. The default settings of our temperature controller are stored in the “temperature_controller.cfg” file (in the “My Documents” folder). Click on the file and the next screen will be like this:



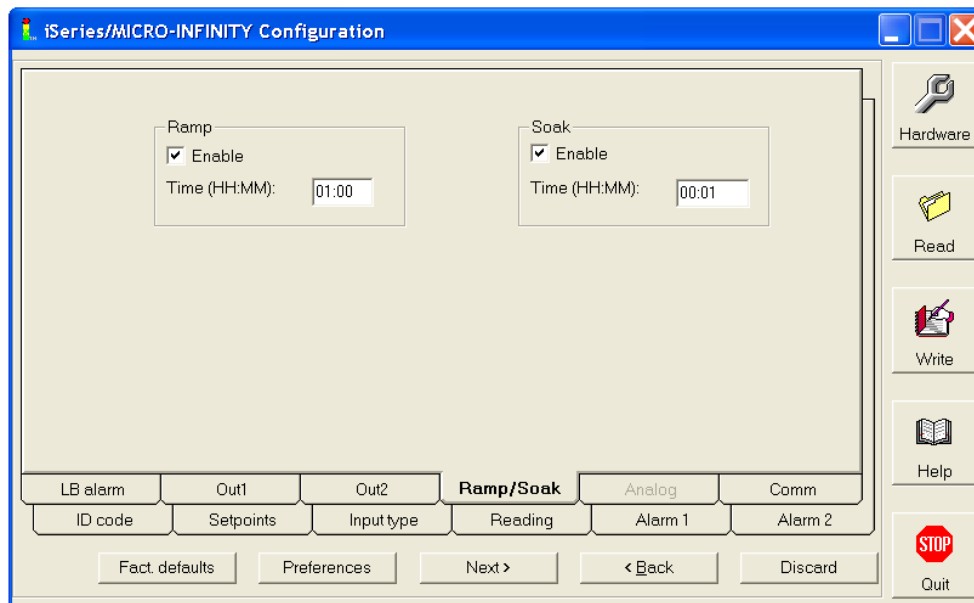
If the Model# and the options come up correctly, click OK. Next the software needs to establish connection with the temperature controller. Click on the “Comm.” Tab and click on the “Connect” button under Connection Wizard. The IP address should be 192.168.1.200. Next:



In this window the RS-232 is the default value, and you need to change it to RS-485. In a few seconds the connection should be established, and you need to push “Finish”.

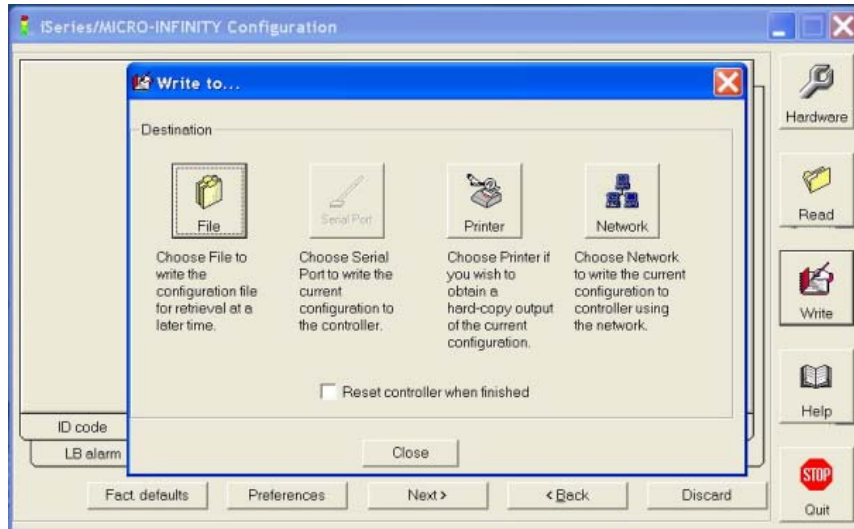
To select a new setpoint, click on the “Setpoints” tab and enter the new value to “Setpoint 1”.

To activate the ramp mode click the “Ramp/Soak” tab on the main window. The window below should appear:



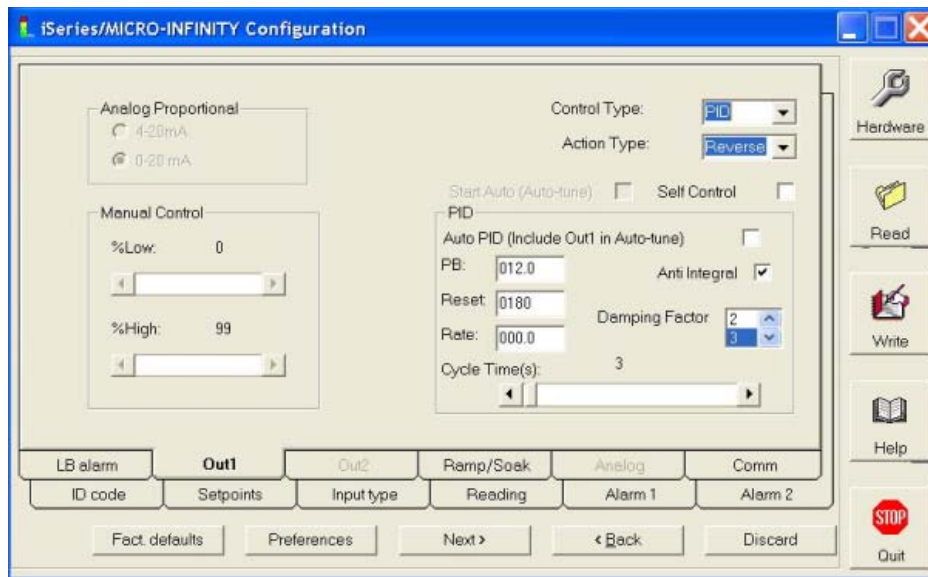
Now you can click “enable” for the ramp option. The Ramp option allows you to control the rate of heating in the sample chamber. For example if you set the ramp to 1 hour then it will take 1 hour to get to your setpoint 1, etc. This is important since you want to keep your sample thermometer consistent with your cavity thermometer. **If you decide to use the ramp option you MUST enable the soak option as well.** The soak option will guarantee that once your temperature is reached and your soak time is complete that the heater will turn off.

After that you upload the change to the controller by pressing “Write” and selecting “Network” in the screen:



During the upload process hit OK several time when the software stops for confirmation.

To access the PID control parameters click on the “Output 1” tab. The next screen shows the “Control Type”. Here is the place where you can change to “On/Off Control”:



Appendix 4: Data-taking with LabView

- 1.) Click on the “Ferro” shortcut on the Desktop.
- 2.) Press the right arrow in the upper left corner to start the program. The communication to the voltmeter and the lock in should start, but the data is not recorded yet.
- 3.) Write the file names where the data should be stored into the fields on the top part of the window. Use this format `c:\StudentData\2012Fall\Group1\yourfile.txt`

The directories Group1m, Group2, Group3 correspond to the first, second and third labs in any given semester.

- 4.) If you wish to record the data, press “Save”. Two files can be stored: one is the temperature dependence of the capacitance, and to other one is the time dependence of the temperature. It is crucial that you write the file name in your lab notes, where you also record the lock-in settings and any other information you may need to reconstruct the circumstances of the measurement.
- 5.) Stop data taking by pressing “Save” again. Write in a new file name, so that the next time you save data it is recorded separately. (If you do not change the file name, the new data will be appended to the existing data.
- 6.) To clear the screen, press stop. When you restart the measurement (right arrow in the upper left corner) a blank screen will appear.

Appendix 5. Physics of Phase Transitions

1. Introduction to Phase Transitions

An interesting, important and useful class of phenomena goes under the name *Phase Transition*^(4,5). Phase transitions are widespread in nature, and a number of them find useful applications in technology. In this lab we will primarily concentrate on phase transitions which manifest themselves in electrical properties of solids.

Microscopic theories of different phase transitions involve diverse phenomena specific to these transitions. Yet there are some similarities which allow their classification and a (somewhat) unified description. The *phenomenological* theory (i.e. non-microscopic theory, or a theory which mathematically describes given phenomena without asking detailed questions about their mechanism) of phase transitions was developed largely by a famous Soviet Physicist Lev Landau around 1937⁽⁴⁻⁷⁾. It is based on thermodynamic ideas and on symmetry. While the detailed description of this theory exceeds the level at which we can present the subject here, some of the Landau ideas are actually quite simple and transparent, and we will present them here, specifically in sec. 1.2.2.

We will start with a general principle of phase transition classification. Then, the following sections will detail two phase transitions: the *metal-insulator* transition⁽⁸⁻¹⁰⁾, and the *ferroelectric* transition⁽¹¹⁻¹³⁾ that you will study in this lab.

A phase transition is usually defined as an abrupt change in physical properties upon a continuous change in external parameters. Examples of external parameters are temperature, pressure, magnetic and electric fields. Of these temperature is the most common cause of a phase transition, and in this lab you will be studying such temperature-induced transitions. The value of the temperature, pressure, or other physical quantity at which a phase transition occurs is called a *transition point*. If it is temperature, it is usually denoted as T_C .

1.1 Classification of Phase Transitions

A distinction is made between two general classes, or, as they are commonly called, *orders* of phase transitions. This distinction follows pioneering classification of phase transitions proposed in the 1920-th by Paul Ehrenfest. Why do we speak of a transition *order*? The reason is that Ehrenfest classified phase transitions based on the behavior of the *thermodynamic free energy*^(4,5) (see below) as a function of other thermodynamic variables. Under this scheme, phase transitions were labeled by the lowest derivative of the free energy that is discontinuous at the transition. Thus, according to this scheme, *first-order phase transitions* are defined as transitions exhibiting a discontinuity in the first derivative of the free energy with respect to some thermodynamic variable, and *second-order phase transitions* are continuous in the first derivative of the free energy with respect to some thermodynamic variable, but exhibit discontinuity in a second derivative. Free energy itself is always continuous at the transition point, which is required by conservation of energy.

A modern classification is only slightly different from Ehrenfest's. The meaning of transition's *order* in terms of discontinuous free energy derivatives remains, while the distinction between *first* and *second-order* transitions is formulated as follows: in a *first-order* phase transition, specified amount of heat, called the *heat of transformation*, or *latent heat*, is released or absorbed per unit mass. In a *second-order* phase transition, there is no latent heat -- heat is neither released nor absorbed -- but instead some physical quantity, called an *order parameter* appears; *order parameter* is equal to zero on one side of a *transition point* (in other words, it does not exist in one of the phases), and it gradually increases from zero on the other side. It was Lev Landau in his above-mentioned theory of *second-order* phase transitions⁽⁴⁻⁷⁾ who first raised the *order parameter* to prominence. We will say more on this subject later.

Examples of *first-order* phase transitions include many of the most familiar ones, such as water-ice transition, or water-vapor transition; more generally -- all the solid/liquid/gas transitions are of first order. An interesting and practically important *first-order* phase transition manifests itself in the electrical and optical properties of a solid: it is a *metal-insulator* transition⁽⁸⁻¹⁰⁾, or, as it is sometimes called, *semiconductor-to-metal* transition. It can be found in a fairly large number of solid compounds (but never in pure elements), including Vanadium Dioxide (VO_2), the material you will be studying in this lab.

Second-order phase transitions also include many of the transitions of great interest in Physics and its applications. A familiar and important example is a *ferromagnetic* transition^(4,5,12,14) at a *Curie point*: above the Curie temperature (in Iron it is $T_C = 770 \text{ C}$) the material is in a paramagnetic state, with randomly oriented magnetic moments; at the Curie point the material forms domains with magnetic moments all aligned in the same direction within a domain. This phenomenon is called *spontaneous magnetization*. An *order parameter* we mentioned above in this case is the net magnetization inside a domain; it is zero at T_C and increases from zero below T_C . We note also that zero value of order parameter corresponds to a completely disordered state (random orientation of magnetic moments), while it's finite value corresponds to an appearance of order (common direction for all the magnetic moments in a domain). At the same time, the material in paramagnetic state is more *symmetrical* than in the state of spontaneous magnetization: rotational symmetry of the macroscopic collection of randomly-oriented microscopic magnetic moments in a paramagnetic state is *broken* when material transitions into the

oriented state below T_C . Thus such a phase transition is intimately related to symmetry, the fact which was fully explored by Landau in his theory. We note here that *ferroelectric* transition in Barium Titanate (BaTiO_3) which you will study in this lab is similar to *ferromagnetic* one, with *electric dipole moments* and *electrical polarization* taking place of their magnetic counterparts (we will discuss the differences in sec.2.7). We also note that ideas of spontaneous symmetry breaking migrated from Solid State Physics into cosmology and elementary particle theory, and are presently of singular importance in these fields, illuminating a large number of fundamental puzzles; for example, the celebrated Higgs mechanism responsible for the appearance of mass was first proposed by a famous Solid State theorist P. W. Anderson in 1962, who at the time worked on the aspects of theory of Superconductivity.

Another example of a second-order phase transition is a *superconducting* transition⁽¹⁴⁾, which takes place below a transition temperature (also called *critical* temperature) T_C : at $T > T_C$ one observes a normal, resistive metal; at $T = T_C$ electrical resistance abruptly (in pure metals) disappears and a normal metal turns into a superconductor with zero resistance; at lower temperatures $T < T_C$ resistance is identically zero. The *order parameter* in this case is the superconducting gap: an energy gap which is equal to zero at and above T_C and opens up at lower temperatures. It is the gap in the electronic spectrum, of the order of a meV, separating *paired* electronic state (so-called Cooper pairs) from the normal electron spectrum. Note that superconducting gap is quite different from the gap in a semiconductor or insulator electronic spectrum, of which we will speak in Sec. 2. The entropy of a superconducting state is lower than that of a normal metal state, indicating that electron pairing provides for a more ordered state.

1.2 Thermodynamic approach to Phase Transitions

In discussing phase transitions, it is instructive to take a thermodynamic approach, considering *Helmholtz free energy*

$$F = U - TS \quad (1)$$

of a material at the two sides of a phase transition. Here U is *internal energy*, T is temperature, and S is entropy. Sometimes instead of Helmholtz free energy F , Gibbs free energy G is being considered; it is equal to $F + pV$, where p is pressure and V is volume. In a phase transition there are generally two phases, one being stable on one side of the transition, another on the other side. Thermodynamically, a phase transition occurs at the external parameter value (temperature, or other) at which the free energies of two the two phases become equal (which guarantees, as we said above, that the free energy is continuous across the transition). One can make some conclusions about the two general types of phase transitions by making some general observations concerning formula (1).

1.2.1 First-order transitions from thermodynamic point of view; mixed-phase states

We recall the First Law of Thermodynamics which states that the heat Q absorbed by a system is equal to the change in internal energy of a system ΔU plus the work done by a system W : $Q = \Delta U + W$. If the system does not produce any work (i.e. if it does not expand or contract against some outside pressure), then $Q = \Delta U$. This is generally the case in solid-state transitions of interest to us here. This tells us that *first-order* phase transition, which, as we said above, is accompanied by the appearance of latent heat Q , is possible only if internal energy of the system changes. This energy change must be electronic or ionic in nature: either the electron energy spectrum, or the electrostatic energy of the ionic lattice, or both, must undergo a change at the transition point. For example, there may be an energy gap E_g opening up on one side of a phase transition, as is the case in VO_2 . Such a rearrangement in the energy spectrum of a solid will be abrupt: in our example, the semiconductor energy gap in VO_2 exists on one side of a transition and does not exist on the other; in other words, it does not gradually opens up starting from zero at the

transition, as the superconducting gap. The latter, as we said, serves as an order parameter in that *second-order* transition. Semiconducting gap, in contrast, could not be considered an order parameter.

Electronic energy spectra in VO_2 on the two sides of a transition are shown schematically in Fig. 1:

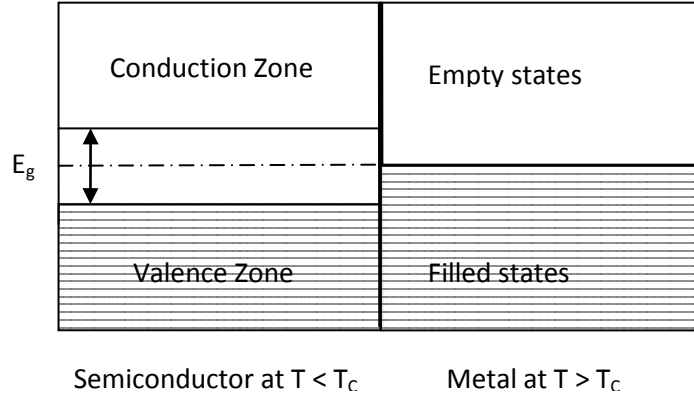


Fig. 1. A diagram schematically representing electron energy spectrum in a *first-order* phase transition in which energy gap is opening below T_c ; such transition is observed in VO_2 .

Further, we see that in formula (1) there are two terms: U and TS . We already decided that in the *first-order* transition U must change. But we also know that free energies of the two phases must be equal at T_c . Therefore the second term, TS , must also change at the transition in order to compensate for ΔU . Indeed, S , the entropy, must change if there is latent heat Q , being related to it by $\Delta S = Q/T$. Change of entropy signifies appearance or disappearance of *order* (here the word *order* is used in a context of order-disorder and not in a context of derivative's order). For example, in VO_2 one finds different lattice symmetry in the two phases: one phase is tetragonal, another monoclinic (these terms refer to types of crystalline atomic arrangements). And indeed one finds a change in entropy which comes predominantly from this lattice rearrangement (see below a section on VO_2). The lattice constant of a unit cell and material's density are also changing in the *first-order* transition. Note that this is different for the *second-order* transition, where *order* does not change at T_c , but changes gradually away from T_c .

Let us continue with general considerations concerning a *first-order* transition. We are considering two phases, phase 1 and phase 2, on the two sides of a transition. Suppose we are crossing the transition temperature T_c going from phase 1 to 2. Free energy, as we said above, is a continuous function across the transition, so that $F_1(T_c) = F_2(T_c)$. However, $F(T)$ is not constant; each phase has lower F on its "rightful" side of a transition. This is schematically shown in Fig. 2, where phase 1 is stable at $T < T_c$ and phase 2 is stable at $T > T_c$:

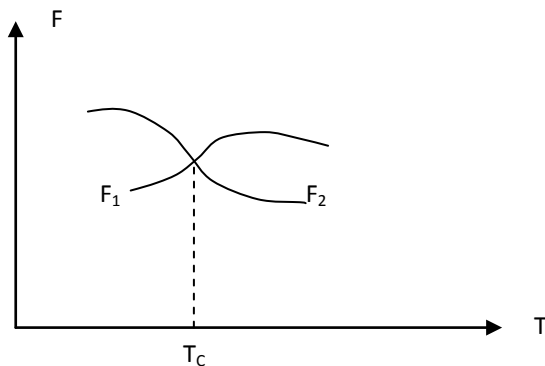


Fig. 2. Schematic behavior of free energy of the two phases across the phase transition

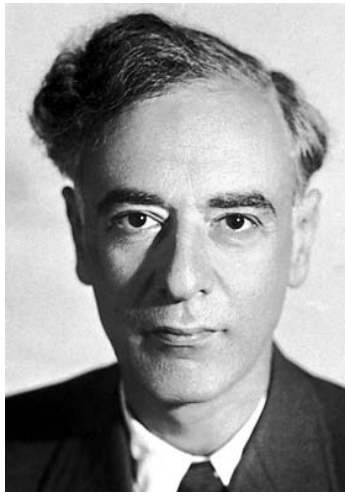
If a given phase – for example, phase 1, is energetically favored (having lower F), one might expect it to appear right away at crossing the T_c . But consider in more detail the mechanism of its appearance: what has to happen for it to appear? Some atomic rearrangement and/or lattice symmetry change will be required; electronic and ionic energy spectra will have to shift; latent heat will have to be released or absorbed. Very importantly, unless the whole volume of the material uniformly transitions all at once, *domains* of phase 1 may form, surrounded by the phase 2, and, as a result, there appears the *boundary* between the two phases. This boundary may have energy, either positive or negative. In case boundary energy is positive, the boundary is said to “cost” energy. Appearance of a boundary thus is changing somewhat the energy balance dictated by the lower free energy of phase 1. There may appear a *local energy minimum*, which depends on local configuration of the phase domains; this local minimum may preclude the system from falling into the *global minimum* dictated by the lower F_1 value. Thus the phase transition may not happen exactly at T_c , and it may develop in a complex, non-uniform fashion. During this process, the temperature of the system will stay constant as heat is added: the system is in a *mixed-phase regime* in which some parts of the system have completed the transition and others have not.

Such phenomena are well known in *first-order* phase transitions between solids and liquids and liquids and vapor: a liquid may be heated to a temperature above the boiling point or *supercooled* to below the freezing point. Something similar happens in a *first-order* phase transition in a solid. *First-order* transitions are often characterized by the formation of *metastable* states. Instead of taking place exactly at T_c , the transition going up in temperature takes place at $T_c + T^*$, and in going down in temperature at $T_c - T^*$. Here T^* is called *coercive* temperature. This leads to a phenomenon of *hysteresis*. If we fix a temperature T somewhere in the interval $T_c - T^* < T < T_c + T^*$, a *metastable* state, often in a form of domains, may survive for long (sometimes indefinitely long) periods of time on the “wrong” side of a transition. In sec. 2.4 you will find a qualitative discussion of the reason for the appearance of coercive temperature and hysteresis.

In this lab you will study one such *first-order* phase transition, namely, the semiconductor-to-metal transition in Vanadium Dioxide (VO_2), which also exhibits a mixed-phase regime. Additionally, the semiconductor phase of VO_2 has high resistivity, while its metallic phase has low resistivity. Thus, as we explore the mixed-state with electrical resistance measurement, we are encountering another interesting type of a transition here, called *percolation*⁽¹⁵⁾ transition. We will briefly discuss these phenomena when talking about VO_2 in Sec.2 In this lab, when performing measurements on VO_2 , you will be able to see how, with increasing temperature, resistivity gradually changes from high to low, dramatically changing by several orders of magnitude in the process. The material goes through the metal-insulator transition, a two-phase mixture occurs, and, as a result, it also goes through the percolation transition.

1.2.2 *Second-order* transitions from thermodynamic point of view; elements of Landau theory

In a second-order transition there is no latent heat, that is, $Q = 0$ and thus, from the First Law of Thermodynamics, $\Delta U = 0$. Thus an energy spectrum is not expected to modify abruptly at the transition point in a *second-order* transition. At the same time, the general condition $F_1 = F_2$ at $T = T_C$ must be satisfied. Looking at formula (1) for F , this tells us immediately that entropy (order) cannot change at T_C , but has to change gradually as the temperature deviates from T_C . In other words, the *order parameter* must be zero at T_C , and will increase away from T_C , as we said above. This, among other considerations, allowed Landau to construct his theory of the second-order phase transitions. Namely, the fact that order parameter is small in the vicinity of a transition point allowed him to expand free energy F as a Taylor series in order parameter, which we will call m , and then, making general physical arguments, to extract significant physical results from this expansion.



L. D. Landau (1908 – 1968); Nobel Prize of 1962

Let us follow his lines of thought, at least in part, here. The Taylor expansion of a continuous function $F(T, m)$ in powers of a small parameter m around point $m = 0$, written to the 4-th order of m , is

$$F(T, m) = F_0(T, 0) + [F'(T, 0)/1!]m + [F''(T, 0)/2!]m^2 + [F'''(T, 0)/3!]m^3 + [F''''(T, 0)/4!]m^4 + \dots$$

Here primes indicate derivatives with respect to m ; for example $F'' = d^2F/dm^2$. So far this is just a mathematical formula based on continuity of F and smallness of m . Landau then argued that F should be invariant with respect to a change in sign in m : changing m to $(-m)$ should keep the energy unchanged. Indeed, in the absence of external field, if m is magnetization M or polarization P , it seems obvious. This is a symmetry argument we mentioned before. This implies that our Taylor expansion must not contain odd powers of m . Omitting odd terms with m and m^3 , and calling even-term coefficients $[F''(T, 0)/2!] = \alpha(T)$ and $F''''(T, 0)/4! = (1/2) \beta(T)$ ($1/2$ introduced here following a convention found in most of the literature), we will obtain

$$F(T, m) = F_0(T) + \alpha(T) m^2 + (1/2) \beta(T) m^4 + \dots \quad (2)$$

We want to relate behavior of $F(T, m)$ to the observed physical phenomena in the vicinity of a phase transition. A minimum of F with respect to an order parameter m physically implies stable (observable) configuration of our system: the value of magnetization or polarization observed in a real system will

correspond to the value of m at which F has a minimum. We will find that these minima are different for $T > T_C$ and for $T < T_C$, on the two sides of a phase transition. After finding the minima with respect to m , we may be able to see how F depends on m in the vicinity of these minima. Once we decide on the form of $F(T, m)$ for fixed T , we will be in the position to discuss T -dependences. Further, by taking derivatives of $F(T, m)$, one can construct other quantities, such as entropy, specific heat, etc. Such is our “program” of how to explore physical consequences of expansion (2).

To find minima of F with respect to the order parameter m , we require that $dF/dm = 0$ and $d^2F/dm^2 > 0$: the usual conditions we always use when finding a minimum of a function. The first one, $dF/dm = 0$, guarantees that we have zero slope, i.e. that we are at the *extrema* of a function; the second condition, $d^2F/dm^2 > 0$, guarantees that this extremum is actually a minimum and not a maximum or the inflection point.

Differentiating (2) we find $dF/dm = 2 \alpha(T) m + 2 \beta(T) m^3 = 2m [\alpha(T) + \beta(T) m^2] = 0$, which gives us three possible solutions for m :

$$\begin{aligned} m_1 &= 0 \text{ and} \\ m_{2,3} &= \pm [-\alpha(T)/\beta(T)]^{1/2} \end{aligned} \quad (3)$$

The second condition for a minimum is $d^2F/dm^2 = 2\alpha(T) + 6\beta(T) m^2 > 0$. The simplest way to satisfy this last inequality is to require that at $T > T_C$, $\alpha(T) = a (T - T_C)$ with coefficient $a > 0$, and $\beta(T)$ is a positive constant $b > 0$. We note that if we adopt this form for $\alpha(T)$, then $\alpha(T_C) = 0$.

So, now the expansion looks like this:

$$F = F_0(T) + a (T - T_C) m^2 + (1/2) b m^4 + \dots \quad (4)$$

And F has minima at

$$\begin{aligned} m_1 &= 0 \text{ and} \\ m_{2,3} &= \pm [(a/b)(T_C - T)]^{1/2} \end{aligned} \quad (5)$$

Clearly the expression for $m_{2,3}$, by virtue of containing a square root, works only for $T < T_C$ (temperatures below the transition); thus $m_1 = 0$ corresponds to $T > T_C$ (temperatures above the transition). As can be seen from (5), at the transition temperature, $m = 0$ in both cases.

Now we will try to explore the physical consequences. Recalling the physical meaning of m , we see that equilibrium magnetization (or polarization) is zero at $T > T_C$: this is clearly the description of a paramagnetic (paraelectric) state, in which magnetization from randomly-oriented magnetic moments equals to zero. According to the expansion, at $m = 0$, $F = F_0(T)$. At $m \neq 0$, and with positive $T - T_C$ (above the transition), expansion (4) predicts that F will rise from the lowest energy level F_0 as $F - F_0 = Am^2 + Bm^4$, where A and B are positive coefficients. This is schematically shown in Fig. 3(a), where we took F_0 to be zero.

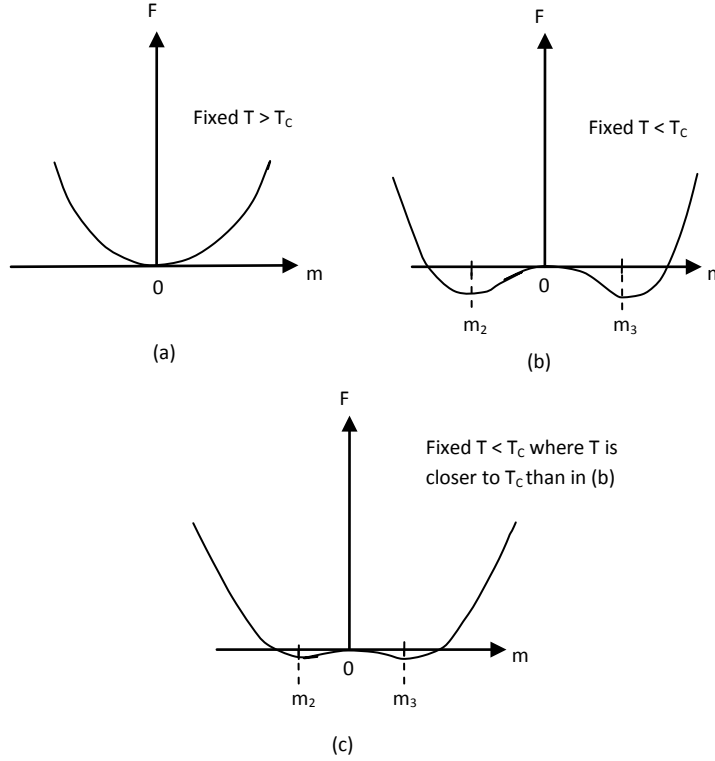


Fig. 3 Schematic representation of free energy F dependence on order parameter m at three fixed temperatures

The minima corresponding to $m_{2,3}$ describe behavior of F below the phase transition: our theory predicts that there will be two such minima, symmetrical in m with respect to $m = 0$. Between these two minima, at $m = 0$, in virtue of (4), we have $F = F_0$ (or zero, since we have chosen zero at F_0). This is depicted in Fig.3(b). Physically this means that our system will have two possible stable states with non-zero magnetization or polarization. Thus in a ferromagnetic (or ferroelectric) state magnetization (polarization) will have a finite value. This finite value may be positive or negative; it can reverse, still being consistent with our theory.

The values of F at these minima can be calculated by substituting m_1 and $m_{2,3}$ into (4). They will depend on temperature T , which is now less than T_C . We see from (4) that the closer is T to T_C , the smaller will be the negative term $a(T - T_C)m^2$ in the expansion, and thus the shallower will be the two minima below the transition. At the same time, with decreasing $T - T_C$, the two minima will move in towards the $m = 0$ value. In the limit $T \rightarrow T_C$ the two minima will disappear, transforming the double-minima curve into a single minimum curve (a). At $T = T_C$, because of the m^4 term, the curve will have a significant flat region in the vicinity of $m = 0$. Physically this must imply that, near the transition point, magnetization (polarization) in a ferromagnetic (ferroelectric) state can be very easily reversed. In other words, the material at a temperature just below the transition point must be very easily polarizable. This suggests that *dielectric constant* ϵ will be high. We will see that indeed this interesting conclusion is experimentally verified, and you will be able to see this very clearly in your measurements of BaTiO_3 .

So far our conclusions based on the analysis of expansion (4) have been very satisfying: we were able to explain some of the most prominent features of a second-order phase transition. We could continue and explore physical consequences for the derivatives of F , which will give us entropy and specific heat. We will find^(4,5) that entropy S , while being continuous at the transition, decreases below the transition, where

there is more order. We will also find that specific heat experiences a jump at the transition. Such a discontinuity in specific heat at the *second-order* phase transition is well documented experimentally.

One can add external field (magnetic field H , or electric field E) to this analysis. In doing so one finds^(4,5) that *susceptibility* χ , which is defined as $\chi_M = dM/dH$ for the magnetic case and $\chi_E = dP/dE$ for the electric case^(4,5,14), obeys Curie-Weiss Law (here written for the ferroelectric case):

$$\chi_E = [2a(T_C - T)]^{-1} \text{ for } T < T_C \text{ and } \chi_E = [4a(T - T_C)]^{-1} \text{ for } T > T_C \quad (6)$$

This shows that Landau theory predicts that *susceptibility* becomes infinite at $T = T_C$. We also see that it approaches $T = T_C$ in an asymmetric way (constants being different by a factor of 2). In an experiment you will perform in this lab you will see that this is indeed so, except that theoretical infinity is substituted by a sharp peak.

We recall Curie-Weiss law^(4,5,14) which states that $\chi = C_{CW} / (T - T_C)$ for the paramagnetic (paraelectric) phase, i.e. for $T > T_C$. Here C_{CW} is a Curie-Weiss constant (often called Curie constant). We see that Landau theory identifies coefficient $1/(4a)$ with C_{CW} .

2. Description of phase transitions and materials which you will study

2-1. Metals, semiconductors and insulators and their conductivity

With respect to an ability to transmit electrical charge, materials can be divided into conductors and insulators; this was recognized from the 18-th century. Ohm's law, *resistance* and *resistivity* were defined and clarified in 19-th and the beginning of 20-th century. Today, with respect to electrical resistivity ρ (or its inverse $1/\rho = \sigma$, electrical *conductivity*) we distinguish not two but three classes of materials: *metals* (good conductors), *dielectrics* or *insulators* (very bad conductors), and *semiconductors*, which are in some sense intermediate between the first two classes. We also distinguish them by different ways in which conductivity depends on temperature (below). True understanding of what is at the core of these different classes came only in the 1930s, when physicists understood electronic spectra (band structure) of these materials^(12, 14). This required creation of Quantum Mechanics, formulation of Pauli's Exclusion Principle, and replacement of classical statistics with Fermi-Dirac statistics. In Fig. 4 we show a schematic representation of the electronic *bands* of a metal, semiconductor and an insulator:

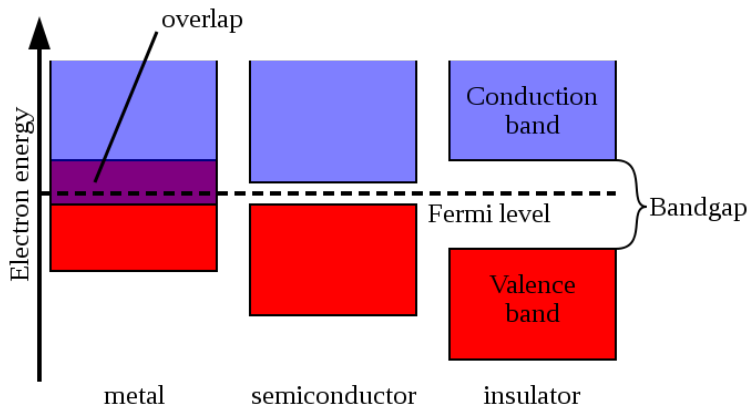


Fig. 4. Electronic spectra (band structure) of metals, semiconductors and insulators

The main distinction between them, which influences electrical conductivity and its temperature dependence, lies in the presence (in semiconductors and insulators) or absence (in metals) of an *energy gap* in the electronic spectrum. The value of the energy gap is called E_g . The highest energy band which contains states which are completely filled with electrons is called a *valence* band; an empty band, separated from it by an energy gap, is called a *conduction* band. Electronic states of the type that may conduct current (the so-called *extended* or *Bloch* states) do not exist in the energy interval inside the gap. The difference between semiconductors and insulators is quantitative rather than qualitative: they both have the gap in their spectra, but the gap in a semiconductor is smaller than that in an insulator. The rough borderline between these two classes lies around $E_g = 2 - 3$ eV. One can say that semiconductor materials are small band-gap insulators. It should be said that in addition to the presence of a relatively small gap, another defining property of a semiconductor material is that it can be *doped* with impurities that alter its electronic properties in a controllable way. Here we will be concerned mainly with the so-called *intrinsic* semiconductors, which do not have impurity levels inside the gap. The most famous semiconductors today are Si and GaAs (and their extended families); but there are hundreds of known semiconductors, and many of them find useful applications. VO_2 you will study in this lab is one of them.

Let us understand why conduction will be so very different depending on the presence or absence of a gap in the electronic energy spectrum. You will need to understand this, as we will ask you to find the value of the gap in VO_2 from your data.

In a conductor, application of electric field E leads to acceleration of charge carriers (such as electrons) with mass m and charge e : $a = eE/m$. This acceleration acts for some average time between collisions τ , leading to the appearance of drift velocity $v_d = a \tau = (eE/m) \tau$. Current density is then

$$j = env_d = (e^2 n \tau / m) E = \sigma E \text{ (Ohm's Law)}$$

Here n is the number of carriers per unit volume (carrier density), and σ is electrical conductivity which characterizes the material. Its inverse is resistivity ρ . It is clear from these familiar considerations (constituting part of what is known as *Drude model*^(12,14) of electrical conduction) that an electron must be able to change its energy in order to participate in electrical current; indeed, it acquires extra kinetic energy $mv_d^2/2$ from the field; this extra energy will be eventually lost in inelastic collisions of an electron with phonons (lattice vibrations), other electrons, etc. We know that eventually this extra energy is deposited into the conducting material, raising its temperature: the familiar I^2R power dissipated in a conductor comes from the field (the battery) and ends up as heat in the conducting material. Note that energy increase $mv_d^2/2$ is, generally, quite small compared to the total kinetic energy of an electron at the highest filled energy level.

So, to repeat: electron needs to change its energy in order to accelerate under the influence of electric field; this is the pre-requisite for electrical conduction. But, to change its energy, it has to find un-occupied energy levels at that higher energy. This is possible in a metal, where un-occupied states lie directly above the occupied ones (at zero temperature we call the highest energy of occupied states *Fermi level*; it is shown as a dotted line in Fig. 4; there are empty states right above it). In a semiconductor or insulator, in view of the existence of a forbidden energy interval (energy gap) this is impossible: there are no available states above the top edge of a conduction band (see Fig. 4). At further reflection, however, we see that this last statement is strictly true only at $T = 0$. Indeed, at a finite temperature T electrons may be excited across the gap. Relative probabilities of such processes can be found from Fermi-Dirac distribution functions; approximate (but very good) estimates can be done, however, using classical Boltzmann factor $\exp(-E_g/kT)$. Such an excitation process creates an electron near the bottom of a conduction band above the gap, and a *hole* (which is equivalent to a positively charged electron⁽⁵⁾) near the top of the valence band. These charge carriers (electron and hole) are called *mobile carriers*; they can

now participate in conducting the current, because they can find un-occupied energy levels very close to their respective energies.

The number of mobile electrons and holes depends on the values of T and E_g , and on how many electrons where available for excitation. In good insulators, i.e. materials with $E_g > 2$ eV, at practical temperatures (limited, generally, by the melting point; often we are interested in conductivity at $T = 300$ K), their numbers are negligibly small. In other words, resistivity of an insulator is many orders of magnitude higher than in a semiconductor or in a metal. For example, resistivity of a very good insulator Amber (the one that ancient Greeks were rubbing to produce static electricity) at 300 K is $5 \times 10^{14} \Omega\text{m}$, while resistivity of pure Si is $2300 \Omega\text{m}$, a difference of 10^{11} orders of magnitude! Metals are still orders of magnitude more conductive than semiconductors: resistivity of Copper at 300 K is only $1.7 \times 10^{-8} \Omega\text{m}$. Thus metals and good insulators differ by staggering 22 orders of magnitude (!) in resistivity (or conductivity).

In addition to very different values of ρ , its temperature dependencies are very different as well: in good metals resistivity increases with T ; in Copper and many other metals this increase is linear at sufficiently high T . In contrast, resistivity in a semiconductor depends on temperature as

$$\rho = \rho_0 \exp(E_g/2kT) \quad (7)$$

Here the factor of 2 comes from the fact that activation energy is counted from the Fermi level, or, as it is called in semiconductors, from *chemical potential*, which, in intrinsic semiconductors, lies in the middle of the gap^(12,14); an electron has to receive energy equal to $E_g/2$ rather than E_g to cross the gap. Thus the study of the temperature dependence of electrical resistance allows us to find E_g . The most convenient way of doing this is to plot $\log(\rho)$ vs. $1/T$.

2.2 Metal-Insulator Transition

One type of a phase transition is that between metal and insulator or metal and semiconductor; it is most commonly called *Metal-Insulator Transition*⁽⁸⁻¹⁰⁾ or MIT, even when the poorly-conducting phase is a semiconductor.

If we will take a piece of metal, such as Copper, and heat it up to its melting point, the melt will still be a metal. If we boil it (in vacuum) and evaporate the molten Copper, the Copper atoms constituting the vapor will be of course, strictly speaking, forming some sort of an insulator in a sense that they will not be conducting a current if placed in the electric field. This is not however what we call a metal-insulator transition. We are interested in a solid material which retains its general structure and composition, but, under the influence of an external parameter, for example, temperature undergoes an abrupt transition from a poorly-conducting semiconductor or insulator state to a well-conducting, metallic state or visa-versa. There are some 50 or more compounds (but no pure elements) which exhibit MIT. For example, MIT was found in Fe_3O_4 , NiS , NbO_2 , V_2O_3 , etc. One such compound you will have a chance to study in this lab: it is vanadium Dioxide, or VO_2 . All MIT's are *first-order* transitions, although this fact can be obscured by the presence of a disorder⁽⁸⁾.

2.3 Essential Physics of one type of MIT: Mott transition

The theory of MITs is rather complex. What is more, several different MIT mechanisms may be at work in different compounds. Here we can only give a "taste" of a relevant theory by qualitatively describing the transition as envisioned by N. F. Mott^(9,10). Other MIT models exist⁽⁸⁾.

Sir Nevill Mott, a famous British physicist, in 1949 discussed^(9,10) the following model: consider N hydrogen atoms are in a periodic lattice. If the distances between these atoms are small, N electrons which they contribute will fill half of the band which consists of $2N$ states (the reason there will be $2N$ states in such a band is that there are two states for each electron due to electron spin $\frac{1}{2}$: one up and one down^(12,14)). Thus such a hydrogen crystal will be a metal. Now Mott imagined increasing the distance between atoms (increasing lattice constant a). With increasing a , the band's energy span will decrease; it will become more and more narrow (consult Solid State Physics texts^(12,14) to see why), but, within the confines of the band theory, such "stretched out" hydrogen crystal will still be a metal with half-filled band for any, arbitrarily large, separation between atoms. This is of course nonsensical, because we know that eventually, at large enough atomic separations, the crystal must become an insulator.



Sir Nevill Francis Mott (1905 – 1996); Nobel Price of 1977

This leads one to believe that, with increasing a , something new has to happen at a certain distance between atoms. Mott assumed that, starting from some critical value of a , electronic states no longer can be described as extended Bloch states. Electronic wavefunctions will become *localized*. Such a crystal will be an insulator at $T = 0$: each electron will be localized at its respective atom. Such an insulator is called *Mott insulator*.

At sufficiently high T , however, such a material can conduct current, turning from an insulator into a metal. Indeed, consider two neighboring atomic sites, 1 and 2; each has an electron on it. If these two electrons have equally-oriented spins, the conduction is impossible in view of Pauli principle: an electron 1 cannot jump to site 2 which is occupied by electron 2 with the same quantum numbers. But suppose these electrons have opposite spins (what is called *antiferromagnetic* order). An electron from an atomic site 1 can in principle jump to the neighboring site 2, which has an empty energy state for electron #1's direction of spin (the other state being occupied by the electron with opposite spin which is sitting on site 2). If electron #1 jumps, it will leave behind a positively-charged *hole* on site #1. If such jumps are frequent, with an application of an electric field the current will flow, this current being carried by both electrons and above-mentioned holes they leave behind (holes moving in the direction of the field, and electrons against this direction, as usual⁽¹⁴⁾). To make this jump, however, an electron has to overcome *Coulomb repulsion* which exists between the two electrons residing on the same site – it has to have energy at least of the order of this Coulomb repulsion energy. In a Mott insulator, this Coulomb repulsion barrier energy serves as an analog of a forbidden gap E_g in a band picture of a dielectric. With raising temperature more electrons will be excited across this Coulomb gap, and thus the Fermi level (kinetic energy) will grow. Once the kinetic energy of electrons will become comparable to Coulomb repulsion energy, the current can flow. Thus, at a certain sufficiently-high temperature, Mott insulator will undergo

a transition to a metallic state. Similar transitions can be envisioned as a function of doping, which likewise can increase electron concentration and thus kinetic Fermi energy, or pressure, which can reduce lattice constant a and, as a result, broaden the band and raise Fermi energy again.

Therefore, MIT in this picture takes place essentially as a function of free-electron concentration n , at a certain critical value n_C . External parameters which can drive n to this value of n_C can be temperature, pressure, doping, etc. Lattice transformations can also play a role too, leading to changes in n .

The nature of a MIT in our material, VO_2 , has been debated for many years. The latest opinion shared by the majority of workers in the field is that it is essentially a Mott transition as described above, and not a result of a lattice transformation. Instead, it is believed that lattice transformation itself is the secondary effect from an electronic Mott transition.

2.3 MIT in VO_2

The MIT in VO_2 was discovered in 1959, and extensively studied ever since. As a function of temperature it takes place at $T_C = 340 \text{ K} = 67 \text{ C}$ (transition is sharp in single crystals⁽¹⁶⁾ but broad in polycrystalline samples⁽¹⁷⁾; you will measure it in thin films of VO_2 which exhibit a fairly broad transition). It is a *first-order* transition characterized by latent heat, density change, and lattice symmetry change, as well as by a dramatic change in resistivity. We have no way of detecting aspects of a transition other than resistivity in this lab. You will measure resistance as a function of temperature. In Fig. 5 we show an example of such a measurement taken from Ref. 17, performed on a 95 nm film of VO_2 . Your measurement will be different, as different films exhibit a somewhat different behavior; for example, the ratio of resistance just below the transition to resistance above the transition may be quite different in different samples. The overall features, however, are similar.

The transition, as is typical of all *first-order* transitions (see Sec. 1), involves formation of a mixed-state, in which there are domains of metallic phase and domains of semiconducting phase. This is discussed in greater detail in the next section, 2.4. The transition develops differently depending on the temperature increasing or decreasing – there is substantial hysteresis. As the temperature increases from room temperature (25 C), resistivity will first decrease because our material is a semiconductor. But at some temperature, which is going to be close to the $T_C = 67 \text{ C}$, metallic-phase domains will appear in an otherwise semiconducting material, and resistivity will start to decrease faster. Examination of Fig. 5, which is plotted on a logarithmic scale in resistance, suggests that the first deviation from semiconducting behavior indeed takes place at about 67 C. The fraction of metallic domains grows with increasing T , and at some point a *global metallic cluster* connects the voltage leads on a sample; as a result all remaining semiconducting domains will be *shorted out*, and resistivity will drop to its metallic value. This is the result of the percolation transition⁽¹⁵⁾, which, of course, is made possible by the appearance and growth of those metallic domains.

Q. If – for simplicity – we would imagine that resistivity of a semiconducting phase is infinitely greater than that of a metallic phase -- where, at what temperature on the curve in Fig. 5, would be the point at which a global metallic cluster is forming? This would be the point of a *percolation* transition. Read at least some of the Ref. 15 and briefly discuss percolation physics in your report. At what fraction of metal phase is percolation transition expected to take place in a 2D (thin film) system?

If the temperature will be now lowered from about 90 C back to 25 C, the process will reverse, but with substantial hysteresis. All of this is shown in Fig. 5. The transition, as can be seen, spans about 2.5 orders of magnitude in resistance. This measurement was performed on a VO_2 film; in single crystals resistance

can change by 5 (!) orders of magnitude, and transition is quite sharp, with hysteresis of only about 1 -2 degrees⁽¹⁶⁾.

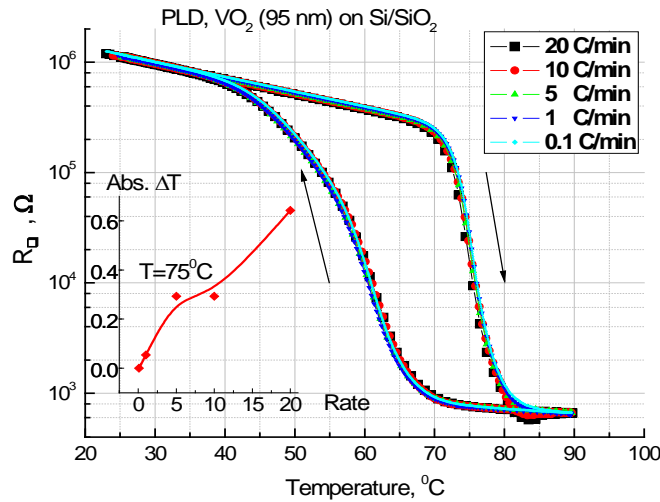


Fig. 5 An example of a hysteresis loop in $R(T)$ measured on a 95 nm film of VO_2 which was deposited by a Pulsed Laser Deposition method (PLD)(from Ref. 17). The figure shows several loops taken at different rates of temperature increase and decrease (temperature ramp rates); the upper inset gives a symbol key for different ramps. The lower inset shows temperature shifts ΔT for different ramps. These are systematic errors in measuring temperature. Although on a logarithmic scale of the picture all these loops are essentially overlapping, there is a temperature shift of up to 0.6 C for the fastest ramp compared to the slowest. **Note that you will have to worry about such ramp rates and temperature shifts in your experiment.**

2.4 Development of a mixed state and hysteresis in VO_2

The text below is taken from Ref. 17 which has been slightly edited here. In this text you will find a qualitative explanation for the appearance of coercive temperature T^ and hysteresis which was advanced by the authors of Ref. 17, primarily by Professor Serge Luryi (Chair, ECE Department, Stony Brook). Other explanations of the same phenomena were given by different authors, and they can be found in the literature⁽¹⁸⁾.*

*The major issue in dispute here is whether each domain exhibits hysteresis on its own, or hysteresis is a collective property of the whole system, arising, as argued here, as a result of the boundary energy “cost”. Phases S and M refer to semiconducting and metallic phases. At the end there is a discussion of a phenomenon called **non-hysteretic branches** which you can observe if you do appropriate measurements on your sample in this lab. Talk to the Professor in charge of this experiment if you want to explore this phenomenon.*

At a given temperature T inside the major hysteretic loop, some parts of the film have $T_C(x, y) < T$ and others $T_C(x, y) > T$. Here x, y are coordinates of a given location in a film. In the first approximation, the boundary wall between the S and M phases, at any given temperature T , is determined by the condition $T_C(x, y) = T$. In this approximation, the wall is highly irregular and its ruggedness corresponds to the scale at which one can define the local $T_C(x, y)$, i.e. to the characteristic length scale of the nanoscopic phase domains. On closer inspection, however, we need a refinement that takes into account the *boundary*

energy, associated with the phase domain wall itself. The boundary energy is positive and to minimize its contribution to the free energy the domain walls are relatively smooth. Let us examine the process of boundary motion. For concreteness, let us consider the heating branch. Below the percolation transition, M phase resembles lakes in the S phase mainland. With raising temperature the area of the M phase increases, lakes grow in size. When a boundary of a given lake is far from the other lakes, infinitesimal $\pm\delta T$ changes boundary length by infinitesimal amount, and the lake area by $\pm\delta A_M$. In other words, when the lakes are sufficiently separated, we envision a continuous, reversible, hysteresis-free process of $M \leftrightarrow S$ area re-distribution, with neighboring configurations differing *microscopically*. Let us now look at the formation of a link between two neighboring regions, which is the elementary step in the topological evolution of a global percolation picture. Let us focus on two metallic lakes that are about to merge. Since the boundary is smooth, at some temperature the distance between the lakes becomes smaller than the radius of curvature of either lake at the point they will eventually touch. Therefore, at some T the following two configurations will have equal energies: one comprising two disconnected M phase lakes that are near touching, but not quite, and the other with a finite link formed, Figs. 6(a) and 6(b) respectively.

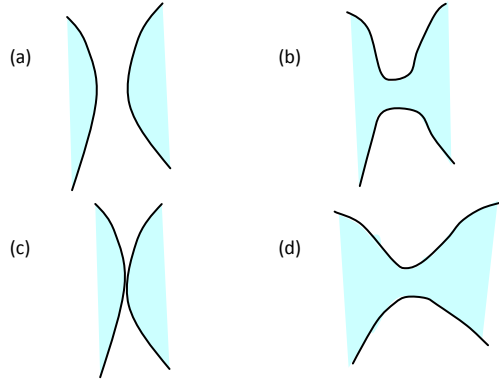


Fig. 6. Semiconductor-metal boundary; metallic phase is shown shaded. Top row (a), (b) corresponds to temperature T and the bottom row (c), (d) to a higher temperature $T_0 = T + T^*$.

Both configurations are characterized by equal boundary lengths and therefore have equal free energy. The actual transition forming a local link, however, does not occur at that temperature because of an immense *kinetic* barrier [*meaning: it takes significant atomic rearrangement*] between these two *macroscopically* different configurations. The transition occurs at a higher $T_0 = T + T^*$ when it is actually forced, i.e. when the two phases touch at a point. Here T^* is the *coercive* temperature. In this picture coercive temperature arises as a result of having a boundary between different phases; it does not pre-exist intrinsically within each domain. We associate the steep slopes of the major loop with the quasi-continuous formation of such links, i.e. with local topological changes, specifically with the merger of metallic lakes on the heating branch (HB) and semiconductor lakes on the cooling branch (CB). Near the major loop ends T_S and T_M the global map consists of widely-separated M and S lakes, respectively. In these regions we expect to see non-hysteretic behavior, as was indeed observed.

Non-hysteretic branches (NHB): Consider now a small *backward* excursion from T_0 on the HB. As the temperature decreases, some of the M-phase recedes and the S-phase grows, changing the geometry of the global two-phase map. However, topologically, the last formed M-link does not disappear immediately for the same kinetic reason. One has two S regions that need to touch in order to wipe out the M-link. It takes a backward excursion of amplitude ΔT^* to establish an S-link and thus disconnect the last M-link.

So long as we are within ΔT^* , i.e. stay on the same NHB, the area of S and M domains changes continuously, but the topology is stable and no new links are formed. Within the range of that stable or *frozen topology*, $\Delta T^* = \Delta T_{\text{NHB}}$, the resistivity of NHB will be single-valued and its T-dependence will be controlled by the percolating semiconductor phase. This explains the appearance of *non-hysteretic branches* and why they have semiconducting slopes.

2.5. BaTiO₃ and ferroelectricity

Ferroelectric materials^(4,5,11-13) exhibit spontaneous electric polarization, the separation of the centre of positive and negative electric charges. Recall Landau theory conclusion discussed at the end of section 1, that free energy will have a double-minimum below T_C : in a ferroelectric state electric polarization, which is the order parameter, is non-zero. This creates an excess positive charge on one side of the material, and negative charge on the other side. The polarization can be reversed by the application of an appropriate electric field: order parameter can be flipped from one minimum to the other with E-field. As we said above, ferroelectricity is named in analogy with ferromagnetism, which occurs in such materials as iron. The prefix *ferro*, meaning iron, was used to describe the property despite the fact that most ferroelectric materials do not contain iron.

In most ferroelectric materials the ferroelectricity disappears at high enough temperature, and the material turns into a regular dielectric. This state is called the *paraelectric*, in analogy to paramagnets. Most of the known transitions from the ferroelectric to the paraelectric state are *second-order*; the transition temperature is called the Curie temperature, in full analogy with ferroelectric transition. (Note: Some ferroelectric materials melt before reaching this temperature.) In the paraelectric state these materials have very large dielectric constant: the temperature dependent dielectric constant $\varepsilon(T)$ increases as the temperature approaches the Curie temperature T_C , and shows a divergent behavior at $T = T_C$. We also note that near T_C dielectric constant is essentially equal to *susceptibility*, a quantity you will find in many references on ferroelectricity. The relation between these two quantities is given in the next section. Again, recall our (or rather, Landau's theory⁽⁴⁻⁷⁾) prediction made in sec. 1: we predicted very large polarizability near T_C , on both sides of a transition. Large polarizability means that induced charges in a dielectric will appear easily, compensating free charges on a capacitor plate – in other words, that relative dielectric constant k will be high.

Ferroelectricity is often accompanied by significant *piezoelectricity* (developing electric polarization under mechanical stress), and *pyroelectricity* (developing electric polarization when heated). Referring to our favorite (by now) Landau's theory, we see that this is perfectly understandable: the system wants to fall into an energy minimum, and just about any outside stimulus will induce it to do so.

Ferroelectric materials have many applications, such as multilayer capacitors, gate dielectrics and waveguide modulators. Pyroelectricity is used in infrared detectors and other sensors. Piezoelectricity is widely applied in microphones speakers, cigarette lighters, and in scientific applications such as the Scanning Tunneling Microscope.

In this lab you will be studying a well-known ferroelectric material, an oxide of Ba and Ti, BaTiO₃^(11,12,19). We note that oxides of this type are important and well-studied in Solid State Physics and Materials Science, and not only because of their ferroelectric properties. These oxides all have chemical formula of the type ABO₃, where A is a metal with valency 1 or 2, either A¹⁺ or A²⁺; B is a metal with valency 4 or 5, either B⁴⁺ or B⁵⁺, and O²⁻ is oxygen. The crystal structure of these materials is called *perovskite* structure, after the naturally-occurring mineral CaTiO₃, which itself is called *perovskite*. Other examples of *perovskite* ferroelectrics include KNbO₃, PbTiO₃, KTaO₃ and PbZr_xTi_{1-x}O₃ (the latter called also PZT).

When talking about crystal structure, we often show one chemical-formula unit of this structure, called the *unit cell*. The whole macroscopic (*bulk*, as it is often called) crystal of a material can be built by stacking unit cells in all 3 spatial dimensions.

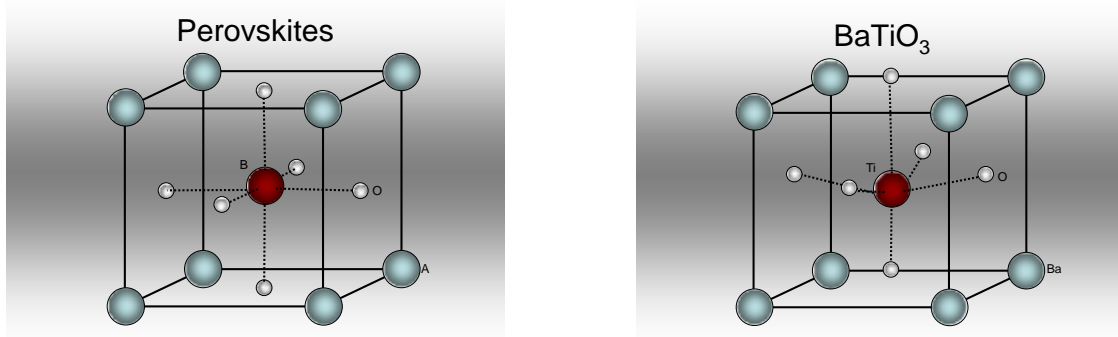


Fig. 7 (a) Unit cell of a *perovskite* structure

(b) Ferroelectric distortion in BaTiO₃

Below a certain temperature T_C , some of the *perovskites* are characterized by an appearance of a *lattice distortion* which changes their structure from *cubic*, as in the Figure 7a above, to a less symmetrical one (the structure which appears below T_C in our material is called *tetragonal*), as in Fig. 7b. Although all of the atoms constituting the unit cell move at least somewhat in this distortion, a rough picture of what happens can be obtained if you will visualize the central Ti^{4+} ion shifting within the unit cell by a small distance shown to be downwards in Fig. 7b; simultaneously, the four oxygen O^{2-} ions surrounding Ti^{4+} also shift, but in the opposite direction, i.e. upwards in Fig. 7b. All these shifts are quite small in comparison to the unit cell side a ; they are somewhat exaggerated in the picture above.

Note that the symmetry of the unit cell is reduced below the transition; a certain part of the total symmetry which existed above the transition is *broken* below the transition, as was discussed in Sec.1.

When the Ti ion is in the center of the cube, the charges are balanced, and the polarization is zero. If the Ti and oxygen ions move as illustrated in Fig. 7b and in the figure below, a finite *electric dipole moment* p appears in each of the unit cells, and the material becomes polarized; this polarization can be reversed with an application of an external field, as shown in the figure below:

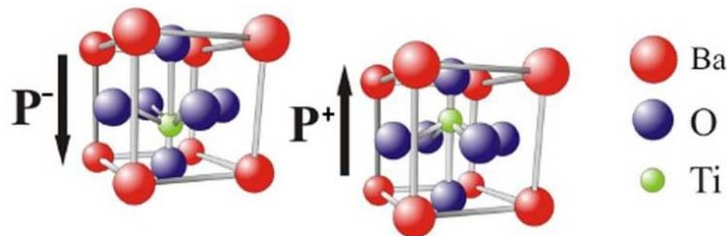


Fig. 8. Unit cells of BaTiO₃ with (green) Ti ion in two off-center positions: with p pointing up (called $p+$ here) and with p pointing down (called $p-$ here). Below T_C , dipole moments $p+$ and $p-$ spontaneously appear without an external field. If there is an electric field E , for example, pointing up, than $p+$ will have lower energy than $p-$.

In this lab we will use a temperature controlled environment to study the dielectric constant as the sample undergoes the phase transition, and we will search for signs of “divergence” in the dielectric constant. We will also attempt to extract some microscopic information from these measurements, as described in the main text of this Write-up and also below.

A note on terminology and notation: The dielectric properties of any material depend on the frequency at which the properties are measured. This is described by the dielectric function $\varepsilon(\omega)$. In our case the properties depend on the temperature as well, and therefore we have a two-variable function, $\varepsilon(\omega, T)$. We use the term “temperature dependent dielectric constant” to represent the low frequency limit of the dielectric function, $\varepsilon(T) = \varepsilon(\omega = 0, T)$. The frequencies applied in our studies (25 kHz) are actually in this low frequency limit. Furthermore, we will distinguish between dielectric function $\varepsilon(T)$ and dielectric constant $k(T)$, which is just the *relative* dielectric function $k = \varepsilon(T)/\varepsilon_0$. We may call it a “constant”, following the normal usage, but in our case, when studying a ferroelectric, we will recognize that it is a temperature-dependent quantity.

2.6 Paraelectric to ferroelectric transition; Curie-Weiss Law

BaTiO₃ undergoes a second-order *paraelectric* to *ferroelectric* phase transition upon cooling through $T_C = 120$ C. Above that temperature it has cubic structure and no local dipole moment; below it transforms to a new, slightly-non-cubic, *tetragonal* structure in which Ti ion and oxygen ions move in opposite directions creating a local dipole moment in each unit cell. In other words, each unit cell becomes electrically polarized and, below T_C , these local dipoles align to form macroscopic polarization P .

As was stated in sec. 1.2.2, electric *susceptibility* is defined as $\chi_E = dP/dE$, or, when P is linear with E , as $\chi_E = P/E$. We mentioned in Sec. 1.2.2 that Landau theory predicts that in the vicinity of a phase transition χ_E is large (in fact, diverging at $T = T_C$) and that it follows Curie-Weiss Law

$$\chi_E = C_{CW} / (T - T_C) \quad (8)$$

The relationship between χ_E and the dielectric constant ε is given by $\varepsilon = \chi_E + 1$ ^(1,2,5,12). Near the transition χ_E is large, and so, to a very good approximation, $\varepsilon = \chi_E$. Therefore Curie-Weiss law (8) can be re-written for the relative dielectric constant as

$$k = C_{CW} / [\varepsilon_0 (T - T_C)] \quad (9)$$

Again, to clarify the meaning of symbols and to avoid confusion, in this equation k is a *relative* dielectric constant, in other words, it is equal to the ratio of capacitance with a dielectric to the capacitance with the same geometry but with vacuum between the plates, $k = C/C_0$. As was mentioned above, although it is called a “constant”, it is not a constant with respect to temperature. The formula for a parallel-plate capacitance with vacuum (or air) between the plates is $C_0 = \varepsilon_0 A/d$, where A is one plate’s area, and d is the distance between the plates. The formula for the capacitance with dielectric is $C = \varepsilon A/d = \varepsilon_0 k A/d$. Here ε_0 , *permittivity of free space*, is equal to $8.85 \times 10^{-12} \text{ C}^2/(\text{Nm}) = \text{F/m}$.

In sec. 1.2.2 we saw that Landau theory predicts this Law both above and below the transition: we had [see formula (6)] $\chi_E = C_{CW1}/(T_C - T)$ for $T < T_C$ and $\chi_E = C_{CW2}/(T - T_C)$ for $T > T_C$, with $C_{CW1} = 1/(2a)$ and $C_{CW2} = 1/(4a)$, a being expansion coefficient. This law is indeed observed in the vicinity of ferromagnetic and ferroelectric transitions, where experimental data mostly concentrates on paramagnetic or paraelectric states at $T > T_C$. You will have a chance to see it in your own measurements, and even to test it on both sides of a transition.

Problems: As was already discussed in the main text of this Write-up, your sample may have a somewhat irregular shape, so you will at best be able to estimate k perhaps within a factor of 2. Additionally, you are measuring a ceramic sample, which will tend to underestimate intrinsic k in your material, as was discussed in “Advice for evaluating the capacitance data” section, also in the main part of this Write-up. The latter problem can be avoided or corrected, at least in part, by making a correction as described in that section (let us call this a *capacitance correction*). You want to use your measured C values to calculate the value of C_{BTO} , which is the intrinsic capacitance of your material. Thus when we will be talking about your capacitance data, we will assume that you made the suggested *capacitance correction*. The k values which you will be using in the k vs. T data should be using this correction.

Keeping in mind these caveats, by fitting your corrected k vs. T data to formula (9) you will be able to find (or at least to estimate) the value of C_{CW} . You may, for example, plot $1/k$ vs. $T - T_C$ and find the slope of that line, which is $1/C_{\text{CW}}$. You will be using your measured T_C here, which is easy to identify as the temperature of the sharp cusp in your data. **Note: you want to make sure that your temperature ramp is slow enough to trust your T_C measurement.** The temperature interval of your data above the T_C will be rather narrow, but you will be able to estimate C_{CW} . It will be interesting to also fit the data below the transition, to see if you can verify the theoretical factor of two difference between $C_{\text{CW}1}$ and $C_{\text{CW}2}$ (see sec. 1.22). As we said above, the largest uncertainties in the value of C_{CW} will come from not knowing the capacitor geometry well, and from measuring a ceramic sample (although the latter may be in part remedied by the *capacitance correction*), rather than from your fits to a Curie-Weiss law.

Why do we want to find C_{CW} ? Because there is a theoretical connection between the value of C_{CW} and microscopic parameters, and we will try to obtain it in the next section.

2.7 Derivation of C_{CW} in terms of a microscopic dipole moment

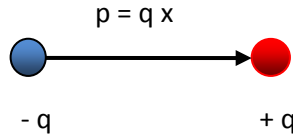
In a ferromagnetic material microscopic magnetic moments μ (fundamentally, spins) are always present; the difference between paramagnetic and ferromagnetic state is merely in the organization, alignment or misalignment of these pre-existing moments. In a paraelectric material, a unit cell is not carrying its p in the same way as an atom carries its spin; in fact the microscopic electric dipole moment p above the T_C should not exist. The lattice symmetry changes at T_C , and above T_C we have a cubic cell with Ti in the center, so that microscopic $p = 0$.

Here we wish to make the following comment: Landau’s potential above T_C (Fig.3a) has only one minimum, and this may suggest that all electrical dipoles p and magnetic moments μ must disappear above T_C . Yet we know that individual spins do not disappear at $T > T_C$ in a paramagnetic state. Thus we realize that the notion of *order parameter* m (in our case, polarization P) is macroscopic in nature; it does not tell us anything about the existence or non-existence of microscopic p ’s above the T_C . Our conclusion about local $p = 0$ in a paraelectric state in BaTiO_3 is based on the known structure and symmetry of its unit cell^(11,12) and not on Landau’s theory.

Another difference between magnetic and electric cases is in that Quantum Mechanics imposes certain requirements on the quantum numbers and energy levels in the magnetic case: as is well-known, a degenerate single energy level in $B = 0$ field will split into two levels for spin $\frac{1}{2}$ electrons in $B \neq 0$ field^(5,6). This may not be the same in case of p ’s and E -field: there is no quantum number associated with p . Yet we will also use the same two-level picture here. In other words, we will assume that microscopic dipole \mathbf{p} can point either along the \mathbf{E} field or against it, as shown in Fig. 8, and will ignore its other possible orientations with respect to the direction of \mathbf{E} .

The analysis of the ferromagnetic case can be found in a number of books, e.g. in Ref. 4,5,12,14. Our derivation here is similar to these treatments, but somewhat simplified.

Let us therefore consider material in a ferroelectric state, at $T < T_C$, where we can assume that one dipole moment \mathbf{p} exists in each unit cell of volume a^3 ; thus in 1 m^3 there are $N = 1/a^3$ microscopic dipoles \mathbf{p} . Let us roughly estimate the value of a microscopic dipole moment in BaTiO_3 . The “generic” dipole is shown below:



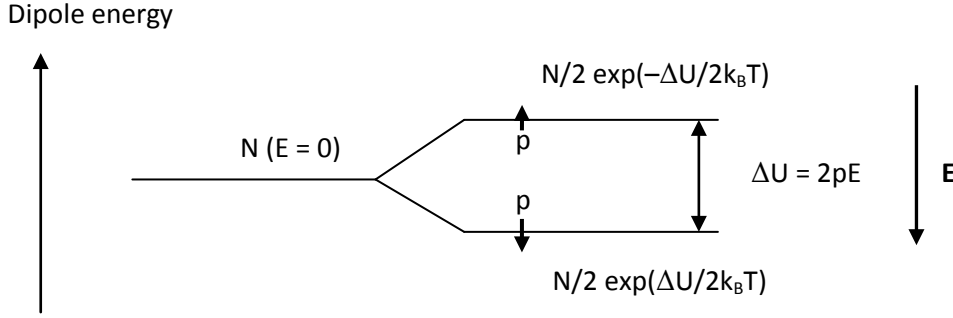
It consists of two opposite electric charges, $-q$ and $+q$, separated by a distance x ; the formula is $\mathbf{p} = q\mathbf{x}$. Whatever ionic displacements happen inside the unit cell of a ferroelectric, and whatever charge imbalance they will create locally, the total charge cannot change, and the unit cell must remain neutral. Thus our *effective* electrical dipole which represents this polarization, must be neutral as well, with $-q$ and $+q$ adding to zero. In this formula we treat q as some *effective* electric charge, which, we recognize, must be of the order of electron charge e , or, at most, it can be several times larger than e . Indeed, we know that in BaTiO_3 , Ti ion carries charge $+4$, and Ba ion carries charge $+2$, while three oxygen ions carry charge -6 , with the sum of all these charges being, of course, zero. We will therefore take $q = 6e$ in our effective dipole.

The second quantity in the dipole formula is x , an *effective* displacement which represents *charge-weighted* displacements of all of the charges in the unit cell. In BaTiO_3 , in the first, rough approximation, we can think of x as a displacement of a Ti^{4+} ion, which in a cubic state resides in the middle of a cell and shifts up or down (Fig. 8) in a ferroelectric state. Therefore, the absolute upper limit on the value of x is $a/2$, half size of the unit cell. In reality we expect x to be much smaller than this upper limit. The lower limit is hard to predict; it can be a small fraction of a . So, we expect to find $0 < x < a/2$. The length of a cubic unit cell side in BaTiO_3 is $a = 0.4 \text{ nm}^{(11,12,13)}$, which gives us $0 < x < 0.2 \text{ nm}$.

Thus we have a rough upper limit estimate $p_{\text{max}} = qx_{\text{max}} = 6ex_{\text{max}} \approx (9.6 \times 10^{-19}) \times (0.2 \times 10^{-9}) \approx 2 \times 10^{-28} \text{ Cm}$.

We also know that in the electric field \mathbf{E} , potential energy of \mathbf{p} is $U = -(\mathbf{p}\mathbf{E})$, i.e. it changes from the minimum value of $(-\mathbf{p}\mathbf{E})$ for \mathbf{p} being parallel to \mathbf{E} to the maximum value $(+\mathbf{p}\mathbf{E})$ for \mathbf{p} being anti-parallel to \mathbf{E} . The thermal energy at room temperature is $k_B T = 1.38 \times 10^{-23} \times 300 = 4.1 \times 10^{-21} \text{ J}$. We see that it will take electric field of $2 \times 10^7 \text{ V/m}$ to make maximum dipole energy equal to thermal energy. We conclude that for any reasonable value of electric field \mathbf{E} , $U \ll k_B T$.

Now, near the transition, where macroscopic polarization \mathbf{P} (the order parameter) is small, in the absence of an external field \mathbf{E} , all microscopic \mathbf{p} 's have the same (zero) energy; once the external field \mathbf{E} is applied, this energy will split into two levels, one above and one below the initial level, the energy difference between these two levels being $\Delta U = \mathbf{p}\mathbf{E} - (-\mathbf{p}\mathbf{E}) = 2\mathbf{p}\mathbf{E}$:



We still have $\Delta U \ll k_B T$, which means that there will be approximately $N/2$ dipoles on each of the two energy levels. Nonetheless the small ΔU is responsible for the population difference between the two levels: the lower level will have $N_1 = (N/2) \exp(\Delta U/2k_B T) \approx (N/2) (1 + \Delta U/2k_B T)$ dipoles, while the upper level will have $N_2 = (N/2) \exp(-\Delta U/2k_B T) \approx (N/2) (1 - \Delta U/2k_B T)$ dipoles, so that $N_1 > N_2$. The total number of dipoles is of course still $N_1 + N_2 = N$. The relatively small difference $N_1 - N_2$ is what is responsible for macroscopic polarization P . So we have: $N_1 - N_2 = (N/2) (2\Delta U/2k_B T) = (N/2) \Delta U/k_B T = (N/2) (2pE) = NpE$. Macroscopic polarization $P = p(N_1 - N_2) = Np^2 E/(k_B T)$. Finally we will get for susceptibility $\chi_E = P/E$

$\chi_E = \frac{(Np^2 / k_B)}{T}$, and for relative dielectric constant $k = \frac{(Np^2 / \epsilon_0 k_B)}{T}$. The one difference between the

last formula and formula (9) is in the temperature dependence: while in (9) we had $1/(T-T_C)$ (or, below the transition, $T_C - T$), here we got simply $1/T$. But we did not develop in this section a full theory of a phase transition; we just assumed that transition takes place at some temperature; in fact, we could assume that it took place at zero temperature; then $T-T_C$ would be the same as T_C . So we should not be very surprised that we did not get $1/(T-T_C)$; the correct temperature dependence was addressed in Landau theory in sec. 1.2.2. You can also find a derivation of these formulas with the correct $1/(T-T_C)$ dependence in Ref. 5.

So we see that we should identify the numerator with C_{CW} ; by doing so we obtain

$$C_{CW} = \frac{Np^2}{\epsilon_0 k_B} \quad (10)$$

or, in terms of a microscopic dipole moment p :

$$p = \sqrt{\frac{\epsilon_0 k_B C_{CW}}{N}} \quad (11)$$

where $N = 1/a^3$ and C_{CW} is a measured quantity which can be obtained from fitting your k vs. T data to the Curie-Weiss Law.

As an exercise, check the dimensionality of C_{CW} in (10): it should come out to be K (Kelvin), so that k in formula (9) will indeed be a *relative* quantity without units.

What you should calculate: You should find C_{CW} derived from your data (again, we assume that you applied the *capacitance correction*), substitute it in formula (11), and calculate p , and then compare it to our upper-limit estimate p_{\max} for this quantity. Once you will know p , you can calculate x , the effective dipole displacement. This last value can be compared to the literature: indeed, transition-induced displacements of various atoms in the unit cell have been studied by X-rays. You should also find in the literature the value of C_{CW} for *single-crystalline* BaTiO_3 , and, using it in formula (11), calculate again p and x . This time the values will be more realistic, as you will be using the correct, intrinsic value of C_{CW} . Compare the results with your own, comment on discrepancies.

References

1. http://en.wikipedia.org/wiki/Lock-in_amplifier About lock-in amplifiers; at the end of this article you will find references to more lock-in tutorials.
2. <http://www.animations.physics.unsw.edu.au/jw/AC.html> A tutorial on AC circuits, with animations
3. http://en.wikipedia.org/wiki/PID_controller#PID_controller_theory About temperature control
4. M. Gitterman, V. Halpern, “*Phase Transitions; A brief account with modern applications*” World Scientific, (2004).
5. P. Papon, J. Leblond, P.H.E. Meijer, “*The Physics of Phase Transitions*”, Springer, (2001)
6. <http://theory.physics.helsinki.fi/~stafyi/SPI06chap10.pdf> About Landau’s theory of phase transitions
7. http://lamp.tu-graz.ac.at/~hadley/ss2/landau/second_order.php About Landau’s theory of phase transitions
8. V. Dobrosavljevic, “*Introduction to Metal-Insulator Transitions*” (2011), [arXiv:1112.6166v1](https://arxiv.org/abs/1112.6166v1)
9. N. Mott, “*Metal-Insulator Transitions*”. Taylor & Francis Ltd. (1974), [ISBN 0-85066-079-3](https://www.tandfonline.com/doi/abs/10.1080/00036817408839503)
10. N. F. Mott, Proc. Phys. Soc. (London) **A62**, 416 (1949)
11. F. Jona and G. Shirane, “*Ferroelectric Crystals*”, Pergamon press, The MacMillan Co., 1962 (a copy of the relevant pages provided in the Lab).
12. J. S. Blakemore, *Solid State Physics*, Second Edition, W. B. Saunders Co., 1974, Chapter 5 “Dielectric and magnetic properties of Solids”.
13. http://instructor.physics.lsa.umich.edu/adv-labs/Ferroelectrics/phase_transitions.pdf About ferroelectric transitions
14. Ashcroft, N. W. and Mermin, D. (1976). *Solid State Physics*. Thomson Learning, Toronto
15. <http://physics.clarku.edu/courses/125/gtcdraft/chap13.pdf> About percolation transition
16. C. N. Bergland and H. J. Guggenheim, “*Electronic properties of VO_2 near the semiconductor-metal transition*”, Phys. Rev. **185** (3), pp. 1022-1033 (1969).
17. M. Gurvitch, Serge Luryi, A. Polyakov, A. Shabalov, “*Nonhysteretic Phenomena in the Metal-Semiconductor Phase-Transition Loop of VO_2 Films for Bolometric Sensor Applications*”, IEEE Transactions on Nanotechnology **9**, pp. 647-652 (2010).

18. For example, see T. G. Lanskaya, I. A. Merkulov, and F. A. Chudnovskii, “*Hysteresis effects at the semiconductor-metal phase transition in vanadium oxides*”, Sov. Phys. Solid State **20**(2), 193 (1978), and a number of more recent papers treating hysteresis in VO₂.
19. http://www.doitpoms.ac.uk/tlplib/ferroelectrics/phase_changes.php About the structure of BaTiO₃, with very nice graphics and illustrations.

Sections relating to Ferroelectricity in BaTiO₃ created by Laszlo Mihaly and in part by Michael Gurvitch; sections relating to Phase Transition Theory and Metal-Insulator transition in VO₂ created by Michael Gurvitch, both in 2012. Thanks to Xu Du, Matthew Dawber, Alexander Shabalov and Carlos Marques for contributions to the design and realization of the measurement and for useful comments.
