1a. Solving Equation 1 for E_f is pretty straightforward, provided we remember the admonition to set $E_V = 0$ and use it at the start. If we don't, it still works, but it takes some less-than-obvious mathematical gymnastics:

$$N_V e^{\left[\frac{-(E_f - E_V)}{kT}\right]} = N_C e^{\left[\frac{-(E_C - E_f)}{kT}\right]} \quad \Rightarrow \quad \ln(N_V) + \left[\frac{-(E_f - E_V)}{kT}\right] = \ln(N_C) + \left[\frac{-(E_C - E_f)}{kT}\right]$$

$$\ln(N_V) - \ln(N_C) = \left[\frac{-(E_C - E_f)}{kT}\right] - \left[\frac{-(E_f - E_V)}{kT}\right] = \frac{-E_C + E_f + E_f - E_V}{kT} = \frac{2E_f - E_C - E_V}{kT} =$$

$$= \frac{2E_f - E_C - E_V + (E_V - E_V)}{kT} = \frac{2E_f - E_C + E_V - E_V - E_V}{kT}$$

$$= \frac{2E_f - (E_C - E_V) - 2E_V}{kT} = \frac{2E_f - E_{gap} - 2E_V}{kT} = \frac{2(E_f - E_V) - E_{gap}}{kT}$$

$$\ln\left(\frac{N_V}{N_C}\right) = \frac{2(E_f - E_V) - E_{gap}}{kT} \quad \Rightarrow \quad kT \ln\left(\frac{N_V}{N_C}\right) = 2(E_f - E_V) - E_{gap}$$

$$2(E_f - E_V) = E_{gap} + kT \ln\left(\frac{N_V}{N_C}\right) \quad \Rightarrow \quad (E_f - E_V) = \frac{E_{gap}}{2} + \frac{kT}{2} \ln\left(\frac{N_V}{N_C}\right) = \frac{E_{gap}}{2} + kT \ln\sqrt{\frac{N_V}{N_C}}$$

Like I said, it's a heckuva lot easier if you set $E_V = 0$ from the get-go. E_V just shifts everything up or down. Alrighty. For Si, $E_{gap} = 1.12$ eV, $N_C = 2.8 \times 10^{19}$ cm^{-3}, and $N_V = 1.0 \times 10^{19}$ cm^{-3}. With E_V set to zero,

$$(E_f - E_V) = E_f = \frac{E_{gap}}{2} + \frac{kT}{2} \ln\left(\frac{N_V}{N_C}\right) = \frac{1.12 \text{ eV}}{2} + \frac{\left(8.617 \times 10^{-5} \frac{\text{eV}}{\text{K}}\right) T}{2} \ln\left(\frac{1.0 \times 10^{19} \text{ cm}^{-3}}{2.8 \times 10^{19} \text{ cm}^{-3}}\right)$$

$$E_f = 0.56 \text{ eV} - \left(4.44 \times 10^{-5} \frac{\text{eV}}{\text{K}}\right) T$$

$$\boxed{E_f(0.001 \text{ K}) = 0.56 \text{ eV} - \left(4.44 \times 10^{-5} \frac{\text{eV}}{\text{K}}\right) 0.001 \text{ K} = 0.56 \text{ eV}}$$

$$\boxed{E_f(150.0 \text{ K}) = 0.56 \text{ eV} - \left(4.44 \times 10^{-5} \frac{\text{eV}}{\text{K}}\right) 150.0 \text{ K} = 0.55_3 \text{ eV}}$$

$$\boxed{E_f(300.0 \text{ K}) = 0.56 \text{ eV} - \left(4.44 \times 10^{-5} \frac{\text{eV}}{\text{K}}\right) 300.0 \text{ K} = 0.54_7 \text{ eV}}$$

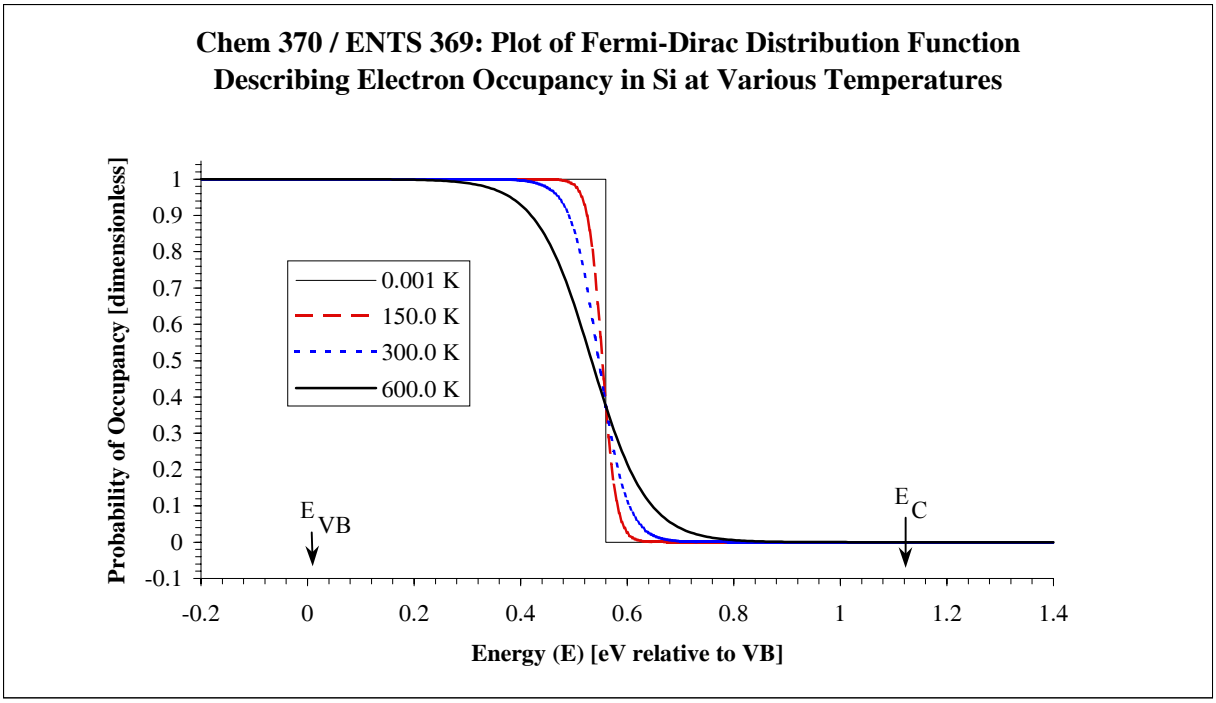
$$\boxed{E_f(600.0 \text{ K}) = 0.56 \text{ eV} - \left(4.44 \times 10^{-5} \frac{\text{eV}}{\text{K}}\right) 600.0 \text{ K} = 0.53_3 \text{ eV}}$$

Note that in Si, E_f moves slightly below mid-gap at elevated temperatures because n_{th} must equal p_{th}, but there's a slightly higher density of states in the conduction band than in the valence band. This is not intuitively obvious to most people! *Be sure you understand and can explain the logic behind this!!!*

b. The Fermi-Dirac distribution is $f_{FD}(E) = \left[1 + e^{(E - E_f)/kT}\right]^{-1}$ = electron occupancy probability

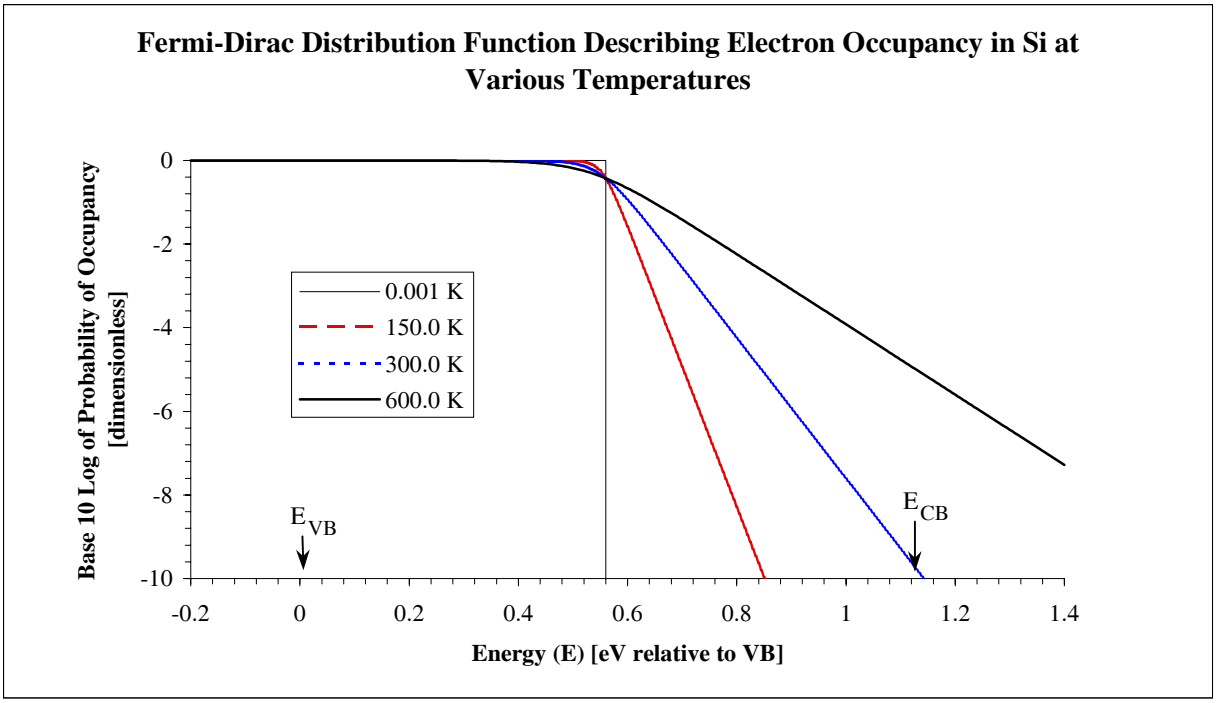
Note that for positive values of $E\text{-}E_f$, the 1 term of the sum quickly becomes negligible.

i. A plot of f_{FD} on a linear scale appears at the top of the following page.



Note that the Fermi-Dirac distribution is pseudo-centro-symmetric about E_f, and is "sharpest" at low temperatures, gradually smoothing out as the temperature rises. Note also that because E_f changes slightly with temperature, the point where f_{FD} equals ½ shifts slightly with temperature. Factoid: The value of f_{FD} that is independent of temperature is analogous to an isosbestic point.

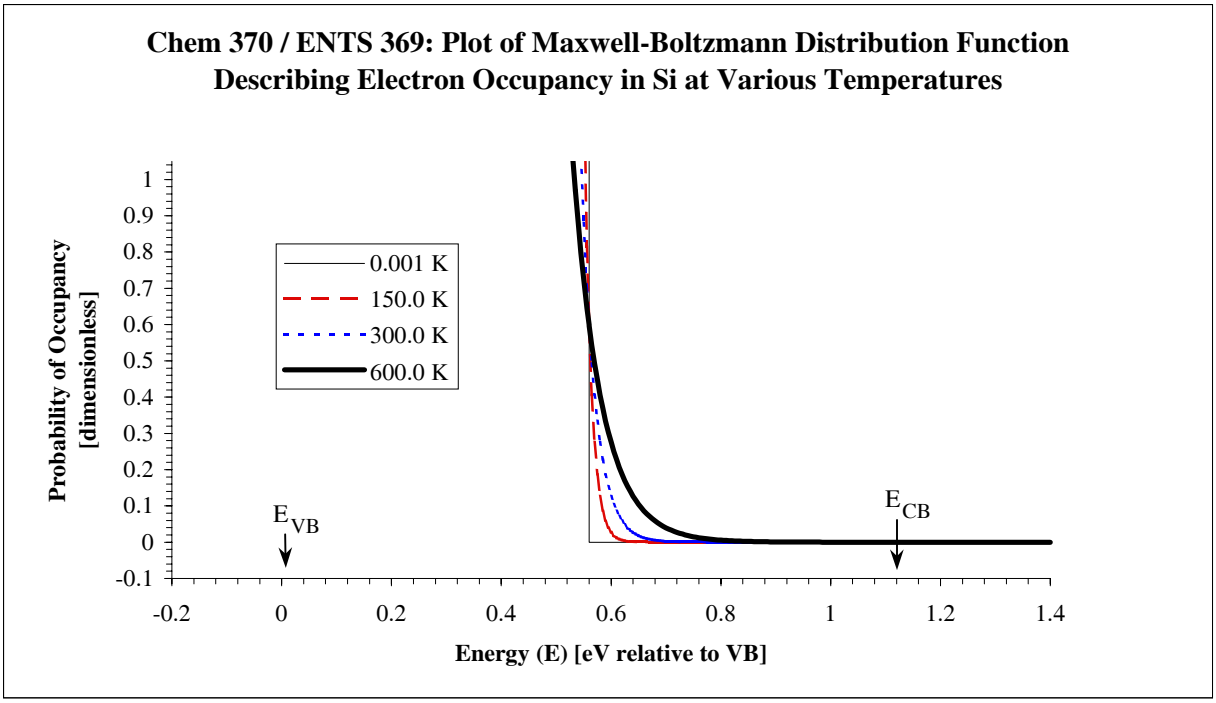
ii. The same information, but now on a log plot:



c. The Maxwell-Boltzmann distribution is $f_{MB}(E) = \left[e^{(E-E_f)/kT} \right]^{-1}$ = electron occupancy probability

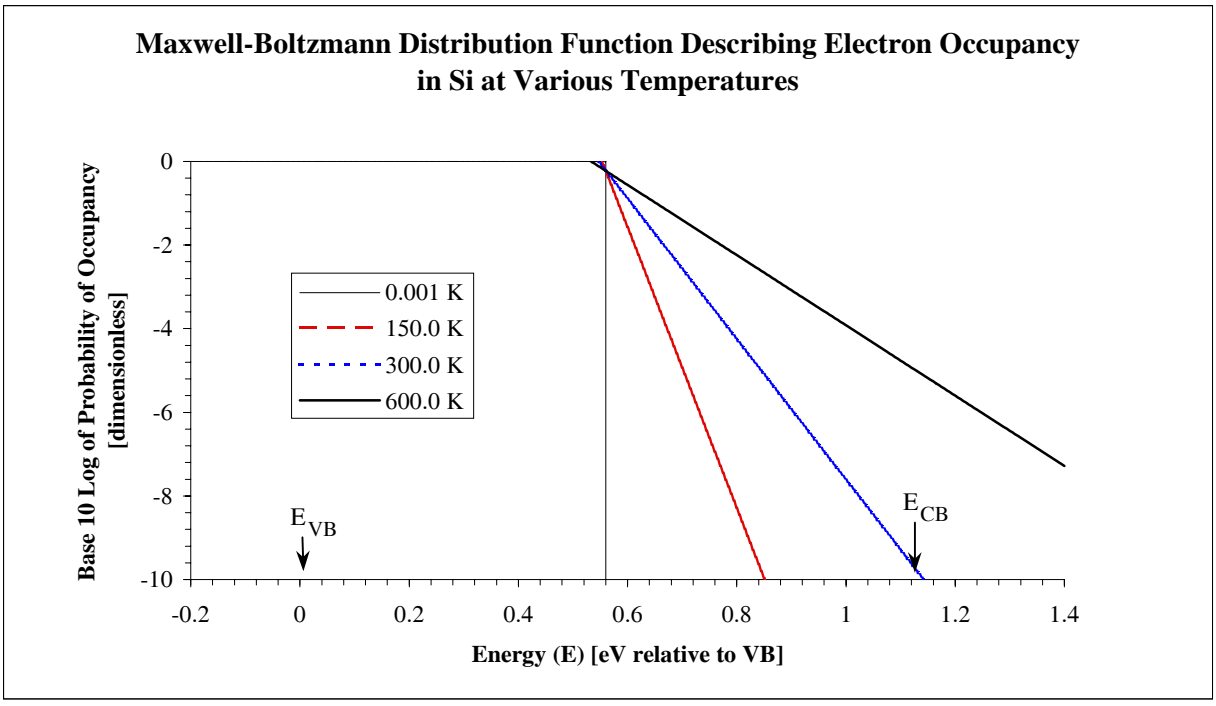
Note that this is the limiting form of f_{FD} for large positive values of $(E-E_f)$!

i. A plot of f_{MB} on a linear scale appears at the top of the following page.



Note that the Maxwell-Boltzmann distribution is nowhere near centro-symmetric, it does not asymptote to one for low energies, and in fact it rapidly approaches infinity for $E<E_f$. However, it does share f_{FD}'s tendency towards increased sharpness with decreasing temperature. Its isosbestic-like point is ½, independent of N_C and N_V (it's blind to occupancy, after all!)

ii. Here's the Maxwell-Boltzmann distribution function on a logarithmic scale:



d. Comparisons of the two distribution functions:

i. Under the conditions considered here, the two distributions are essentially identical for E's ~ 2kT above E_f. At lower energies, however, they very rapidly diverge. This corresponds with the notion that they are in agreement for states that have very low occupation probabilities.

1. d. cont…

ii. Under the conditions considered here, the two distributions are indistinguishable for any energy in the conduction band of Si ($E > E_{CB} = 1.12$ eV). Given that the two functions differ appreciably for E values less than ~2kT above E_f, however, the two distributions might give appreciably different results for states in the conduction band if E_f were within 2kT (we'll play it safe and say 3kT) of the conduction band energy. This occurs in heavily doped n-type samples, in materials with very small bandgaps, and/or at high temperatures.

Problem 1 Postscript:

A little commentary on and explanation of these two distribution functions: The Fermi-Dirac distribution describes the occupation statistics of fermions: particles that can not share states, like electrons. No two electrons can have the same wavefunction, as stipulated by the Pauli Exclusion Principle, and this is why the Fermi-Dirac distribution asymptotes to one for low-energy states. Such states will be occupied by the one electron that they are each allowed to have. [A "State," in the context of an electron distribution, is a *single* unique combination of the principal (n), angular (ℓ), angular momentum (m_ℓ), and spin (m_s) quantum numbers. This differs from the implicit definition chemists are used to using, in which a "state" can hold two electrons, viz., an n, ℓ, m_ℓ combination that allows for both $m_s = +½$ and $–½$.]

The Maxwell-Boltzmann distribution function describes how electrons *would* distribute themselves were they not constrained by the Pauli Exclusion Principle, and could share any given state without limit. Under such conditions, the occupancy of states well below E_f would be far more than one: many electrons would share each low-energy state.

These two distributions become similar for very low occupancies: because when there are relatively few fermions trying to get into a relatively large number of states, the Pauli Exclusion Principle does not come into play. For energies more than 3kT above E_f, electrons do not "bump into" each other, and behave as classical particles.

Comparing the log plots in this problem makes it pretty clear that for occupation probabilities less than about 0.1, we can use the Maxwell-Boltzmann distribution function as an excellent estimate for the Fermi-Dirac distribution function. That might seem a bit restrictive, and for a metal it is. But recall that in semiconductors the Fermi level is typically located within the bandgap, and by the time you get up to an energy where any states actually exist, this condition is generally satisfied. So f_{MB} is a fine estimate for f_{FD} for the states in most semiconductor conduction bands at most temperatures. In the valence band it may seem we have a lot more trouble, but in fact what we are generally interested in is the number of holes in a semiconductor's valence band. There again, they are generally a rather rarified breed and a variant of f_{MB} can be used to predict the probability of a state in the valence band being empty. That's *much* easier to work with than the full-blown Fermi-Dirac distribution, which aggressively resists simple integration tricks.

We won't be doing a lot of heavy-duty condensed phase statistical mechanics in this class, but the general rules learned here turn out to be very widely applicable to the equations describing the operation of semiconductor devices, like the diode equation.

2. Before hunting for any equations, it's worthwhile to think about what's going on here qualitatively. (Shallow) dopants are very easy to ionize, and they will do so almost completely at even very low temperatures. The intrinsic lattice atoms in a moderate-bandgap semiconductor, on the other hand, are difficult to ionize; only a small fraction of them will be thermally ionized at low temperatures. There are many more lattice atoms than dopant atoms, however, and as temperature increases, the fraction of the lattice atoms ionized continues to grow. At a high enough temperature, the number of ionized dopants is overwhelmed by the number of ionized lattice atoms, and the latter dominate the sample's conductivity. Above this temperature, increases in temperature significantly increase the number of free carriers in the solid, and (n+p) varies with temperature. Below this temperature, dopant ionization (which is essentially complete at any temperature above zero Kelvin for very shallow dopants) controls (n+p), and this sum does not change appreciably with temperature. We can see that from the distribution functions we considered in problem 1: they can be used to determine the likelihood of a transition as a function of how much energy it requires, and the temperature. The probability of a transition involving a very small energy increase (ionization of a shallow dopant) is always higher than the probability of a transition involving a very large energy increase (ionization of a lattice atom), but the difference in probability becomes less and less as the temperature rises ($p \approx 1$ at all T for dopant ionization, while p increases with T for lattice ionization.)

   a. Alright, so at low temperature the sum (n+p) is constant, and it reflects the number of ionized dopants in the semiconductor. (if the sample is n-type, doped with donors, then $n \gg p$, $n + p \doteq n$, and $n \doteq N_D$. If the sample is p-type, it is doped with acceptors, $p \gg n$, $n + p \doteq p$, and $p \doteq N_A$.) The natural log of (n+p) is constant at 36.8 at low temperatures (reading across from the horizontal part of the plot on the right side), so $(n + p) = e^{(36.8)} = 9.6 \times 10^{15}\,\text{cm}^{-3}$

   Note that because logarithms basically reduce exponents to unit values, the 6 really isn't a significant figure in 9.6×10^{15}; we used about two of the three sig figs in getting the exponent. The basic rule I use for quick-and-dirty sig fig work is: count the digits in an exponent as sig figs in considering how many sig figs to put in a logarithm. Thus the logarithm of 9.77×10^{15} will produce a logarithm with three sig figs (one for the 9 in the pre-exponential factor and two more for the 1 and the 5 in the exponent), and vice-versa. But you should know how to do this rigorously, if ever you have to. The error after any mathematical operation is obtained from the total derivative of the operation:

$$df(x) = \text{error in } f(x) \qquad dx = \text{error in } x$$

$$df(x) = \frac{df}{dx}dx \qquad \text{so, in the case of } f(x) = \ln x,\ d(\ln x) = \frac{d}{dx}(\ln x)dx = \frac{1}{x}dx$$

   In this problem, we are in a position to wonder what the uncertainty in $e^{(x)}$ is, given dx:

$$d\left(e^x\right) = \frac{d}{dx}e^x dx = e^x dx \qquad \text{Our } x \pm dx \text{ is } 36.8 \pm .1, \text{ so } e^x \text{ should be properly reported as}$$

$$e^{36.8\pm.1} = e^{36.8} \pm e^{36.8}(0.1) = 9.6\times10^{15} \pm 0.96\times10^{15} = 9.6 \pm 1\times10^{15}$$

   Okay, that was a bit of an aside, but important all the same. We're still doing science, and the concept and quantification of uncertainty are still a very important part of what separates us from the television pundits and pie-chart eating folks in business suits. Alrighty, get off your soap box, already, Rob-O. Let's remind ourselves of a few variable definitions:

   n = free electron (e^-) concentration [=] cm^{-3}
   p = free hole (h^+) concentration [=] cm^{-3}
   n_{th} = thermally generated free electron concentration [=] cm^{-3}
   p_{th} = thermally generated free hole concentration [=] cm^{-3}
   N_D = effective (net) donor atom concentration [=] cm^{-3}
   N_A = effective (net) acceptor atom concentration [=] cm^{-3}

At most reasonable temperatures (and at all those considered in this problem, by assumption), essentially all the dopants ionize: acceptors to give holes and donors to give electrons. Thus

$$n = N_D + n_{th} \qquad p = N_A + p_{th}$$

Additionally, note that the thermal generation reaction is:

$$Si^o \xrightarrow[E_{th}]{} Si^+ + e^- \quad \text{or, equivalently,} \quad Si^o \xrightarrow[E_{th}]{} h^+ + e^-$$

Consequently, the number of thermally generated electrons must equation the number of thermally generated holes; they are produced as pairs from an unionized lattice atom: $n_{th} = p_{th}$.

At low temperatures, $(n+p) = N_D + n_{th} + N_A + p_{th} \approx N_D$ or N_A. (n_{th} and p_{th} are negligibly small, and we are told that either N_D or N_A is zero: there's only one type of dopant; we are not looking at a compensated semiconductor. But even if we were, given our use of *effective* dopant concentrations, only one of N_D or N_A could be non-zero.) So we know that for this semiconductor

$$N_D \text{ or } N_A \doteq (n+p) = 9.6\times10^{15}\,\text{cm}^{-3} \quad \Rightarrow \quad \boxed{N_D \text{ or } N_A \doteq 1.0\times10^{16}\,\text{cm}^{-3}}$$

b. At higher temperatures, n_{th} and p_{th} grow large and surpass N_D or N_A. Then

$$\left.\begin{aligned} n &= n_{th} + N_D \approx n_{th} \\ p &= p_{th} + N_A \approx p_{th} \end{aligned}\right\} \text{(high temperatures)}$$

such that, because $n_{th} = p_{th}$,

$$n + p = n_{th} + p_{th} = n_{th} + n_{th} = 2n_{th}$$

Well, whoopie! Provided we are still at equilibrium (we haven't applied a bias or shone any light on our sample), the following equilibrium expression must still be satisfied:

$$n_i^2 = n\cdot p \doteq n_{th}\cdot p_{th} \doteq n_{th}\cdot n_{th} = (n_{th})^2 \quad \Rightarrow \quad n_{th} \approx n_i$$

Thus n_{th} is equal to n_i, which is given by

$$n_i^2 = N_C N_V e^{(-E_{gap}/kT)} \quad \Rightarrow \quad n_i = \sqrt{N_C N_V e^{(-E_{gap}/kT)}} = \sqrt{N_C N_V}\, e^{(-E_{gap}/2kT)}$$

This can be plugged into the (n+p) equation and linearized to match the plot, as follows:

$$n + p = 2n_{th} = 2n_i = 2\sqrt{N_C N_V}\, e^{(-E_{gap}/2kT)} \qquad \text{Remembering that } \ln(a\cdot b) = \ln a + \ln b,$$

$$\ln(n+p) = \ln\left(2\sqrt{N_C N_V}\right) + \ln\left(e^{(-E_{gap}/2kT)}\right) \qquad \text{and because } \ln\left(e^x\right) = x,$$

$$\ln(n+p) = \ln\left(2\sqrt{N_C N_V}\right) + (-E_{gap}/2kT) = \ln(\text{a constant}) + \frac{-E_{gap}}{2k}\frac{1}{T} \cong b + mx$$

This tells us that at high temperatures, a plot of ln(n+p) versus 1/T should be linear, with a slope of $-E_{gap}/2k$ and an intercept that we don't care about in this problem. Drawing a line through the left-hand section of the plot and picking two points, we get the slope:

$$\text{slope} = m = \frac{\Delta y}{\Delta x} = \frac{40.42 - 36.02}{0.0010\ \text{K}^{-1} - 0.0020\ \text{K}^{-1}} = 4400\ \text{K} = \frac{-E_{gap}}{2k}$$

$$E_{gap} = 2\cdot k\cdot(4400\ \text{K}) = 2\left(1.38\times10^{-23}\ \frac{\text{J}}{\text{K}}\right)(4400\ \text{K})\left(\frac{1\ \text{eV}}{1.602\times10^{-19}\ \text{J}}\right) = 0.75_{833}\ \text{eV}$$

$$\boxed{E_{gap} = 0.76\ \text{eV}}$$

(Not too shabby…the actual values I used in preparing the plot were $N_D = 1\times10^{16}$ cm^{-3} and $E_{gap} = 0.78$ eV.)

3. Using the handy table from the reading packet, we have N_C, N_V, and E_{gap} for each semiconductor.
   a. We determine the intrinsic carrier concentration (n_i) at room temperature (300 K) using

$$n_i^2 = N_C N_V \, e^{\left(\frac{-E_{gap}}{kT}\right)} \quad \Rightarrow \quad n_i = \sqrt{N_C N_V}\; e^{\left(\frac{-E_{gap}}{2kT}\right)}$$

$$kT = \left(1.38\times10^{-23}\,\frac{J}{K}\right)(300\ K)\left(\frac{1\ eV}{1.602\times10^{-19}\ J}\right) = 0.0258_5\ eV$$

Germanium: $N_C = 1.0 \times 10^{19}\ cm^{-3}$ $N_V = 6.0 \times 10^{18}\ cm^{-3}$ $E_{gap} = 0.70\ eV$

$$n_i = \sqrt{N_C N_V}\; e^{\left(\frac{-E_{gap}}{2kT}\right)} = \sqrt{\left(1.0\times10^{19}\ cm^{-3}\right)\left(6.0\times10^{18}\ cm^{-3}\right)}\; e^{\left(\frac{-0.70\ eV}{2(0.0258_5\ eV)}\right)}$$

$$\boxed{\text{Germanium}: \ n_i = 1._0\times10^{13}\ cm^{-3} \quad \left(\begin{array}{c}\text{Germanium is pretty darn easy to ionize,}\\ \text{and is passable conductor at room temperature.}\end{array}\right)}$$

Silicon: $N_C = 2.8 \times 10^{19}\ cm^{-3}$ $N_V = 1.0 \times 10^{19}\ cm^{-3}$ $E_{gap} = 1.12\ eV$

$$n_i = \sqrt{N_C N_V}\; e^{\left(\frac{-E_{gap}}{2kT}\right)} = \sqrt{\left(2.8\times10^{19}\ cm^{-3}\right)\left(1.0\times10^{19}\ cm^{-3}\right)}\; e^{\left(\frac{-1.12\ eV}{2(0.0258_5\ eV)}\right)}$$

$$\boxed{\text{Silicon}: \ n_i = 6.5_5\times10^{9}\ cm^{-3} \quad \left(\begin{array}{c}\text{Silicon is appreciably harder to ionize than germanium,}\\ \text{and is a very poor conductor at room temperature.}\end{array}\right)}$$

Gallium Arsenide: $N_C = 4.7 \times 10^{17}\ cm^{-3}$ $N_V = 7.0 \times 10^{18}\ cm^{-3}$ $E_{gap} = 1.42\ eV$

$$n_i = \sqrt{N_C N_V}\; e^{\left(\frac{-E_{gap}}{2kT}\right)} = \sqrt{\left(4.7\times10^{17}\ cm^{-3}\right)\left(7.0\times10^{18}\ cm^{-3}\right)}\; e^{\left(\frac{-1.42\ eV}{2(0.0258_5\ eV)}\right)}$$

$$\boxed{\text{Gallium Arsenide}: \ n_i = 2.1_4\times10^{6}\ cm^{-3}}$$

Cadmium Sulfide: $N_C = 2.4 \times 10^{18}\ cm^{-3}$ $N_V = 1.8 \times 10^{19}\ cm^{-3}$ $E_{gap} = 2.42\ eV$

$$n_i = \sqrt{N_C N_V}\; e^{\left(\frac{-E_{gap}}{2kT}\right)} = \sqrt{\left(2.4\times10^{18}\ cm^{-3}\right)\left(1.8\times10^{19}\ cm^{-3}\right)}\; e^{\left(\frac{-2.42\ eV}{2(0.0258_5\ eV)}\right)}$$

$$\boxed{\text{Cadmium Sulfide}: \ n_i = 3.1\times10^{-2}\ cm^{-3} \quad (\text{Whoa! CdS is mighty hard to ionize!})}$$

Titanium Dioxide: $N_C = 7.9 \times 10^{20}\ cm^{-3}$ $N_V = 1.8 \times 10^{19}\ cm^{-3}$ $E_{gap} = 3.2\ eV$

$$n_i = \sqrt{N_C N_V}\; e^{\left(\frac{-E_{gap}}{2kT}\right)} = \sqrt{\left(7.9\times10^{20}\ cm^{-3}\right)\left(1.8\times10^{19}\ cm^{-3}\right)}\; e^{\left(\frac{-3.2\ eV}{2(0.0258_5\ eV)}\right)}$$

$$\boxed{\text{Titanium Dioxide}: \ n_i = 1._6\times10^{-7}\ cm^{-3} \quad \left(\begin{array}{c}\text{Undoped } TiO_2 \text{ is very much an insulator -}\\ \text{to the extent that it is used as a dielectric}\\ \text{[separating insulation layer] in some capacitors!}\end{array}\right)}$$

Problem 3, continued…

b. In covalent semiconductors, valency and valency preferences provide a good way to predict the behavior of dopant atoms. Essentially, the number of bonds formed by dopant atoms is forced upon them by the lattice: if their valency preferences differ from those of the lattice atom they are substituting for, it is reasonable to predict how they will respond to being forced into a non-ideal situation. Consider a Si atom sitting in a Ga site in a GaAs lattice. Si has a valency of 4, and four valence electrons; whereas Ga has a valency preference of 3, having three valence electrons. The lattice will accept three of the Si's bonding electrons for its covalent bonding network, but not 4: the Si will be forced to have as if it were a Ga atom. It will thus have one left-over electron, which will not be tied up in any covalent bonds and which the Si will not hold on to tightly. This easily-lost electron makes the Si a donor in this situation. If the Si were instead sitting in an As site, the lattice would expect it to provide five valence electrons. Having only four, it would have to share a non-valence electron until it found an extra electron. Very eager to grab an electron from its surroundings, the Si would in this case behave as an acceptor, or p-type dopant. This sort of logic applies to Ge, Si, GaAs, and (to a lesser extent) CdS. Materials scientists have developed a useful shorthand for indicating atomic substitutions in a solid: $A_{B}^{\leftarrow\text{atom that is actually present in this site}}$ $_{\leftarrow\text{atom that would normally be in this site}}$

<u>Germanium</u>

Donors: N_{Ge} (probably too small to actually "fit"), P_{Ge}, As_{Ge}, Sb_{Ge}, Bi_{Ge}, S_{Ge}, Se_{Ge}, Te_{Ge}, etc.

Acceptors: B_{Ge}, Al_{Ge}, Ga_{Ge}, In_{Ge}, Tl_{Ge}, Be_{Ge}, etc.

<u>Silicon</u> (Essentially identical to group IV cohort Ge)

Donors: N_{Si} (again, doesn't actually "fit"), P_{Si}, As_{Si}, Sb_{Si}, Bi_{Si}, S_{Si}, Se_{Si}, Te_{Si}, etc.

Acceptors: B_{Si}, Al_{Si}, Ga_{Si}, In_{Si}, Tl_{Si}, Be_{Si}, etc.

<u>Gallium Arsenide</u> (a bit more confusing)

Donors: C_{Ga} (Again, I doubt this actually "fits," in practice), Si_{Ga}, Ge_{Ga}, Sn_{Ga}, Pb_{Ga} (Pb's probably too large to actually fit into a Ga spot, in practice), S_{As}, Se_{As}, Te_{As}, etc.

Acceptors: C_{As}, Si_{As}, Ge_{As}, Sn_{As}, Pb_{As}, Be_{Ga}, Mg_{Ga}, Ca_{Ga}, Cd_{Ga}, Zn_{Ga}, Cu_{Ga}, Au_{Ga}, etc.

<u>Cadmium Sulfide</u> (quite a bit more confusing!)

Donors: B_{Cd}, Al_{Cd}, Ga_{Cd}, In_{Cd}, Si_{Cd}, Ge_{Cd}, Ge_{Cd}, Sn_{Cd}, Pb_{Cd}, Cl_{S}, Br_{S}, I_{S} (F probably won't play any covalent reindeer games)

Acceptors: Li_{Cd}, Na_{Cd}, K_{Cd}, Ag_{Cd}, P_{S}, As_{S}, Sb_{S}, Bi_{S}, etc.

TiO_2 is a slightly different beast, it being a so-called "ionic semiconductor." The Ti is actually best thought of as a Ti^{4+} ion, the O as O^{2-}. The logic used above still applies, however. Replacing Ti with Al makes the solid short one electron, and thus electron-hungry, or p-type (hole-rich!). Al will have to act like Al^{4+} until it can grab an electron and become the Al^{3+} that it much prefers to be.

<u>Titanium Dioxide</u>

Donors: P_{Ti}, As_{Ti}, Sb_{Ti}, Bi_{Ti}, F_{O}, Cl_{O}, Br_{O}, I_{O}, etc. (Br and I probably won't fit in an O hole.)

Acceptors: Al_{Ti}, Ga_{Ti}, In_{Ti}, P_{O}, As_{O}, Sb_{O}, Bi_{O}, etc.

4. The variables and terminology used in this problem are the same as those used in problem 2. The total free carrier concentration in an intrinsic semiconductor is set by the ionization of the lattice atoms:

$n + p = (n_{th} + N_D) + (p_{th} + N_A) \quad \Rightarrow \quad N_A = N_D = 0$ in an intrisic semiconductor, so $n + p = n_{th} + p_{th}$

We must have $n_{th} = p_{th}$, and so $n + p = n_{th} + p_{th} = n_{th} + n_{th} = 2n_{th}$. Further, because we must satisfy $n_i^2 = n \cdot p = n_{th} \cdot p_{th} = n_{th} \cdot n_{th} = n_{th}^2$, we know $n + p = 2n_{th} = 2n_i$ for any intrinsic semiconductor.

This allows us to mathematically translate the problem's stipulation that we "double the total free carrier (electron + hole) concentration from its intrinsic value: $n + p = 2(2n_i) = 4n_i$.
For any semiconductor at equilibrium (which basically means not illuminated and not under bias), the equilibrium condition $n \cdot p = n_i^2 = N_C N_V e^{\left(-E_{gap}/kT\right)}$ must be satisfied, no matter what dopant soup may be present in the semiconductor. That's going to be key, because we will find that adding dopant reduces the extent of self-ionization, that is, N_D or $N_A > 0$ makes n_{th} and p_{th} smaller, just like adding acid or base to pure water reduces the extent to which it autoionizes.
Let's assume the dopants in our samples are n-type, so that we are trying to determine N_D, and $N_A = 0$. (We'd get an identical numerical result if we assumed p-type, set $N_D = 0$, and determined N_A.) We apply the thermal equilibrium condition again, but this time it doesn't simplify so nicely:

$n = n_{th} + N_D \qquad p = p_{th} + N_A$ but $N_A = 0$, so $p = p_{th}$; note that now $p \neq n \neq n_{th}$, but n_{th} is still $= p_{th}$!

We want to satisfy the criterion $n + p = 4n_i = n_{th} + N_D + p_{th} = N_D + n_{th} + p_{th} = N_D + 2n_{th}$

Consequently, we will want $N_D = 4n_i - 2n_{th}$ Next we turn to our thermal equilibrium equation:

$n \cdot p = n_i^2 \quad \Rightarrow \quad (n_{th} + N_D)(p_{th}) = n_i^2$ (Thermal equilibrium condition)

Plugging in for N_D, and using the fact $n_{th} = p_{th}$ gives us $n_i^2 = (n_{th} + 4n_i - 2n_{th})(n_{th})$

This can be simplified to the following quadratic and solved for n_{th}, which in turns allows solution for N_D:

$$n_i^2 = (n_{th} + 4n_i - 2n_{th})(n_{th}) = (4n_i - n_{th})(n_{th}) = 4n_i n_{th} - n_{th}^2 \quad \Rightarrow \quad n_{th}^2 - 4n_i n_{th} + n_i^2 = 0$$

$$n_{th} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2 \cdot a} = \frac{+4n_i \pm \sqrt{16n_i^2 - 4n_i^2}}{2 \cdot (1)} = \frac{4n_i \pm \sqrt{12n_i^2}}{2} = \frac{4n_i \pm (2\sqrt{3})n_i}{2} = (2 \pm \sqrt{3})n_i$$

At first blush this seems to offer up two possible roots; but n_{th} should never be able to exceed n_i. At most, it can be equal to it, and that only occurs in an intrinsic semiconductor. (If this is unclear, follow the math through with both roots, and you will see that one root gives us a negative dopant density...with which the problem is much more intuitively obvious.) Plugging the smaller root into N_D, we get $N_D = 4n_i - 2n_{th} = 4n_i - 2(2 - \sqrt{3})n_i = 4n_i - 4n_i + 2\sqrt{3}n_i = (2\sqrt{3})n_i$

This is a good deal more than the $2n_i$ we'd expect to need if adding a dopant didn't mess with the autoionization equilibrium! As hinted at above, this is because adding dopants quenches thermal excitation somewhat → thermally, the system obeys Le Châtelier's Principle, and tends to try to keep $n + p$ at $2n_i$. It can be overwhelmed though, just like any buffer. (This question focuses in on the transition region of the plot in problem 2!) Ok, so now let's calculate this for the specific semiconductors in question:

Silicon

$$N_D = 2\sqrt{3}\, n_i = 2\sqrt{3}\, \sqrt{N_C N_V}\, e^{\left(\frac{-E_{gap}}{2kT}\right)} = 2\sqrt{3}\, \sqrt{(2.8\times10^{19}\ cm^{-3})(1.0\times10^{19}\ cm^{-3})}\, e^{\left(\frac{-1.12\ eV}{2(0.0258_5\ eV)}\right)}$$

$= 2.3\times10^{10}\ cm^{-3} \quad \Rightarrow$ It's hard but not inconceivable to get Si pure enough to act intrinsic at 300K!

Titanium Dioxide

$$N_D = 2\sqrt{3}\, n_i = 2\sqrt{3}\, \sqrt{N_C N_V}\, e^{\left(\frac{-E_{gap}}{2kT}\right)} = 2\sqrt{3}\, \sqrt{(7.9\times10^{20}\ cm^{-3})(1.8\times10^{19}\ cm^{-3})}\, e^{\left(\frac{-3.2\ eV}{2(0.0258_5\ eV)}\right)}$$

$= 5.5\times10^{-7}\ cm^{-3} \quad \Rightarrow$ It is nearly impossible to produce TiO_2 pure enough that its resistivity is controlled by self-ionization at 300K! [Oxygen vacancies kill you!]