N. Taylor

Topic 04: Circuit Theorems

Last time, Topic 03, we looked at nodal analysis for systematically writing up a solvable equation system from a circuit consisting of the two-terminal components that we have studied up to now. There were several approaches to this, depending on the circuit and on whether the aim was to do the solution by hand or by computer. This new Topic 04 is not about systematic solutions. It is more like an extension of Topic 02, giving further methods that can be put together as steps to help solving a circuit.

If nodal analysis can solve all our circuits anyway, then why do we have this extra topic ... or even Topic 02? There are at least two reasons. One reason is that we sometimes can get easier calculations, by the simplifications and theorems. This is less important when we have computers to help us, but there are still cases where one can save time by realising that, for example, a circuit of one thousand linear components connected to our two-terminal voltage source can all be modelled as a single source and resistor! Another, probably more important reason is that simplifications and theorems sometimes can help our understanding: they can allow us to make some quick judgements about how a circuit will behave, or what effect a proposed modification can have. For many sorts of analysis and design we have many options of which components and connectivity there are, and what values the components have; it is not feasible to explore all the combinations in detail, so it is important to have a feeling for what effects various changes will have. A lot of ideas and useful proofs are based on other theorems.

The main reasons for leaving the Circuit Theorems until after Nodal Analysis are that: 1) it is often helpful to use nodal analysis to *find* an equivalent circuit; and 2) the equations from nodal analysis help up to justify the superposition and equivalent circuit theorems if we feel like going into that much theory. The main approach in this Topic is axiomatic: a circuit theorem is asserted, and its use explained. Proofs are interesting, and might, some year, get included in the "Extra" part at the end.

1 Superposition

Superposition applies to the currents and the voltages in the *linear circuits* that we consider in this course. Linearity means that there is a direct proportionality between a cause (independent source) and its effect (voltages and currents in the circuit due to just this independent source), and that the effects due to different causes can be summed.

Superposition is an immensely useful principle in circuits, fields, and many non-electrical subjects. Essentially, "when multiple sources contribute to a quantity in a linear system, that quantity is the sum of the separate effects from all the sources". In electromagnetic fields, superposition is a very important tool. For example, the electric field at a point can be seen as the (vector) sum of all contributions from all the different charges everywhere else.

In circuits, any circuit quantity (current, voltage, potential) is the sum of all the values that it would have due to each of the *independent* sources acting alone. A particularly important thing to get right when using superposition is what to do with the *other* independent sources when calculating what one independent source would do by itself. The correct choice is to replace voltage sources by short-circuits, and current sources by open-circuits; but many mistakes are made by doing this the wrong way round.

Why do we not need to look at the separate effects of *dependent* sources too? After all, they are also voltage or current sources that can drive energy into the circuit. The difference is that each one would do nothing by itself: its output is proportional to a controlling current or voltage somewhere in the circuit; with no independent source to drive the circuit, these quantities will all be zero.[1] The dependent sources are rather similar to resistors in this way: their voltage and current can be found as the result of all the separate contributions caused by the separate independent sources. So dependent sources are simply left in the circuit, unchanged, when calculating the effect of each independent source.[2]

1.1 Superposition procedure

Here are the rules for applying superposition.

1. Divide the independent sources into groups, so that each source appears exactly once. For example, three independent sources called a, b and c could be grouped as a and b in one group and c in another, or they could be in three separate groups, etc.

2. For each group, find the circuit solution for the quantity or quantities that you are trying to solve, with just this group of independent sources at their proper values, and all *other* independent sources *set to zero* (nollställd). Note that a voltage source set to zero is a short-circuit, and a current source set to zero is an open-circuit.

[1] Zero values of dependent sources in the absence of independent sources: this is a simplification for 'nice' cases. We can invent cases where a dependent source stimulates its own controlling variable in a way that has a nonzero solution or is unstable. But we tend to ignore these special cases when looking for some simple rules to solve many practical circuits.

[2] This isn't the only choice possible, for getting a circuit solution: see a link in the Extra Section. But it's the common choice when making a list of how to perform superposition in circuit analysis.

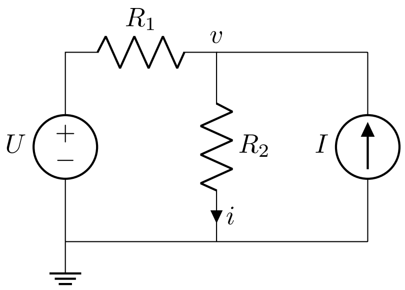
3. For each quantity, add together the solutions found from solving for each group . . . this sum is the solution to the complete circuit where all sources are active.

A more conventional view of superposition tells you that you just have to take each indepdenent source by itself, one at a time. The above procedure is more general: when handling large circuits it can often be useful to do superposition with groups of independent sources instead of always one at a time.

One reason that superposition can be useful is that the replacement of some sources by open and short circuits can make the circuit a lot simpler in each case. If you're lucky, then you might solve for three superposition states, each one with a one-node nodal analysis or divider equation, instead of getting three simultaneous equations when applying nodal analysis without superposition. In other cases it might be quicker to do nodal analysis or other methods directly. However, the usefulness of superposition is even greater for handling special cases in transient and ac solutions that we come to later.

1.2 Superposition example

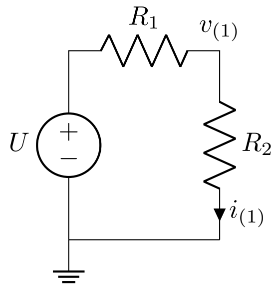
Let's try an example, about as simple as possible without being trivial. All component values are, as usual, assumed by default to be known. We want to find the marked current i. The potential v and earth node have been defined for convenience in the calculations.



If we're going to use superposition at all, in this case with just two independent sources, there's no choice about which groups to use! One superposition state must be with just the voltage source active, and the other must be with just the current source active.

It is useful to be very clear about marking which state is being solved: for example, we can mark the current i as $i_{(1)}$ when it's being solved in the first superposition state, etc.

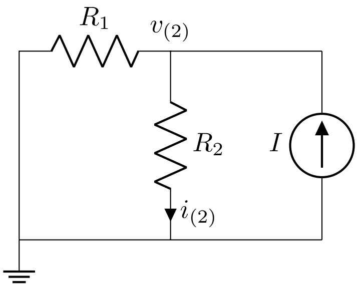
Superposition state 1 Let's choose to have the voltage source active in this first state. Then all other sources, meaning the current source, must be set to zero. A current source set to zero is an *open*-circuit. Therefore, the original circuit can be re-drawn as follows.



This circuit has a very easy solution for $i_{(1)}$, using a series equivalent and Ohm's law, or just KVL.

$$i_{(1)} = \frac{U}{R_1 + R_2}.$$

Superposition state 2 Now the other state is calculated, where the current source is active. The voltage source is set to zero, and thus becomes a *short*-circuit.



This also has a simple solution, using methods from Topic 02. Current division between the resistors gives

$$i_{(2)} = \frac{IR_1}{R_1 + R_2}.$$

Superposition: add the solutions The final step is to combine the calculated values of the desired quantity with the different groups of independent sources acting separately: Hence,

$$i = i_{(1)} + i_{(2)} = \frac{U + IR_1}{R_1 + R_2}$$

1.3 Check the example

It's wise, and educational, to check a supposed solution. Dimensional analysis is a good start. We can usually find several different methods for solving a circuit, without being very confident about which will be easiest until we've tried it. Using an alternative method is a good way to double-check results.

Let's try using nodal analysis directly on the example that we've just solved by superposition. (That's why the potential v was marked!)

We only need to find i, so it is sufficient to find v then use Ohm's law. The supernode method tells us that we can ignore KCL at the node above the voltage source, as we know its potential is U. Just the one node v needs to be considered,

$$\frac{v-U}{R_1} + \frac{v}{R_2} - I = 0 \qquad \text{KCL}_{(v,\text{out})},$$

from which

$$v = \frac{R_1 R_2}{R_1 + R_2}\left(\frac{U}{R_1} + I\right) = \frac{(U + IR_1)\, R_2}{R_1 + R_2}.$$

The desired solution was i, not v, so we divide by R_2.

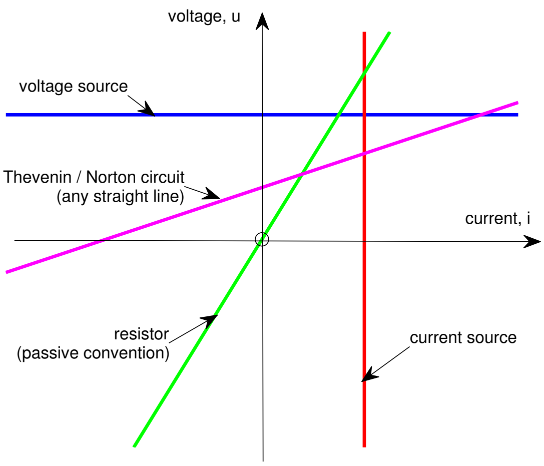
$$i = \frac{v}{R_2} = \frac{U + IR_1}{R_1 + R_2},$$

which matches the solution obtained by superposition. I'd say the nodal analysis was a bit easier this time. But it required a bit more manipulation of the equations, and would quickly lose its ease if there had been more than one equation to solve symbolically (as you know from HW03).

2 Two-terminal equivalents

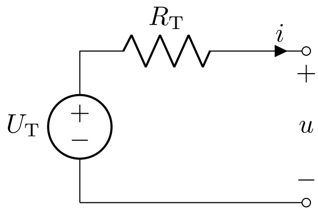
We looked in Topic 02 at all the possible combinations of two components (chosen out of sources and resistors) in series or parallel. The behaviour of each component and combination was considered in the u-i plane, for u and i at the two terminals that connect to the component or combination of components.

Almost all the combinations simply behaved like a single component. For example, two resistors are like another resistor, and a current source in series with anything else would look just like a current source when seen from the terminals.

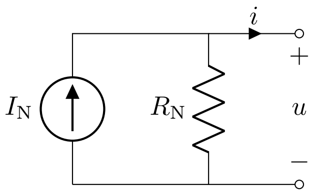


There were however *two special cases* that could not in general be equivalently represented by any single component. These were the two special sources, called the Thevenin and Norton source.

The Thevenin source is a series voltage source and resistor. We previously defined it with a current i going in to the positive reference direction of the voltage u. But now we will define the current in the opposite direction, which is convenient when we are mainly going to think of it as a source that uses its voltage to push a current *out* of it and into a resistor or other simple component connected between the two terminals where u is marked.



The Norton source is a parallel current source and resistance. The resistance is somtimes expressed as a conductance instead, $G_N = 1/R_N$.



We have already seen *source transformation*, whereby the Thevenin and Norton sources are equivalent to each other if $R_N = R_T$ and $I_N = U_T/R_T$.

The equation relating u and i for a Thevenin source, with the directions marked on the diagram above, is

$$u = U_T - iR_T.$$

The Norton source is the dual case, where

$$i = I_N - u\frac{1}{R_T},$$

which could instead be written

$$u = I_N R_N - iR_N$$

to show the equivalence to a Thevenin source. The following analyses will focus on the Thevenin equivalent; we already know from source transformation (Topic 02) that similar properties can be obtained from a Norton source.

In general, U_T and R_T can be positive or negative or zero; a negative resistance is feasible when there are dependent sources in the circuit. Thus, the above equation can describe any straight line "$y = mx + c$". A Thevenin (or Norton) source described in the u-i plane then can therefore have any slope and any offset.

For example, if $U_T = U$ and $R_T = 0$, the Thevenin source behaves like an ideal voltage source U. If instead $U_T = 0$ and $R_T = R$, it behaves as a resistor. If U_T and R_T are very large and have the ratio $\frac{U_T}{R_T} = I$, then it tends to an ideal current source I. Besides these single components, the Thevenin (or Norton) source can fit all the other lines that don't pass through the origin and don't go purely vertical or horizontal.

The **big step** now is the [Helmholtz-]Thevenin theorem, or its current-source equivalent that is sometimes called the Norton theorem. That is, *any linear circuit, when we connect to it by just two terminals, gives some straight-line relation of u and i, and therefore a Thevenin or Norton source can be found that is equivalent to this whole circuit.* That is potentially a very big reduction of complexity, if one wants to analyse what is happening outside the circuit that has been replaced.

The necessary conditions are that: the circuit that is to be replaced with an equivalent must be linear; this circuit must have only the two terminals by which we connect to it; and the controlling variables of dependent sources in this circuit must all be *inside* the circuit, not somewhere outside the terminals.

So, how do we find the *Thevenin equivalent* or *Norton equivalent* of a more complicated circuit?

If we accept that a two-terminal circuit will have a u-i relation that is a straight line, then we just need to define that line. It should then be easy to find the correct U_T and R_T (or I_N) values, by fitting them to the line: $u = U_\mathrm{T} - iR_\mathrm{T}$.

2.1 General equation

One way to find the equivalent parameters is to find the equation that relates the two quantities u and i at the terminals. According to Thevenin's theorem, this relation should be in the form $u = c + ki$. We just need to group all the terms that correspond to the additive and multiplicative constants c and k, which will tell us respectively the Thevenin voltage and resistance.

One way of thinking of this is to take the circuit for which an equivalent is to be made, and connect a current source i to its terminals. Then find the voltage between the terminals as a function of this current. One could alternatively connect a voltage source and find current as a function of voltage.

The same principle applies experimentally, by setting the current or voltage to several different values, and measuring the other quantity that corresponds in each case; for a circuit where Thevenin's theorem applies, this should ideally result in a straight line relation, whose gradient and slope can be calculated.

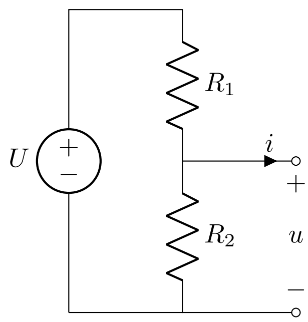
The disadvantage of this general method, for circuit analysis, is that one has to analyse a slighly more complicated circuit than the original one. The method in the next section solves for two special points in the u-i plane, that tend to be easier to calculate.

2.2 Two points: e.g. SC and OC

A straight line can be defined by two numbers: two points on the line, or one point and the gradient (slope).

If two points on the u-i line are to be found, it tends to be easiest to find the points where the line crosses the axes. That's because these points correspond to the short-circuit and open-circuit conditions, in which it is common that one or more of the components near the circuit's terminals can be ignored, thus simplifying the circuit solution.

Let's take this simple example, which comes up a great deal in practical uses of two-terminal equivalent circuits.

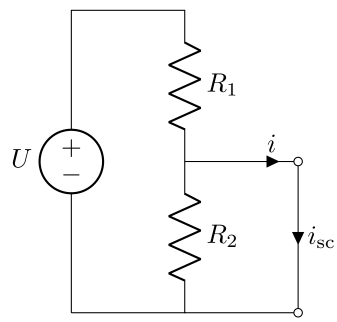


The three components, connected to two terminals shown on the right, are the circuit that we want to replace with an equivalent.

Take the case where the two terminals are left *open*-circuited, as shown above; in this case, $i = 0$, and we need to find the corresponding voltage u. Due to the open circuit, the two resistors are truly in series, so voltage division can be used to find the voltage u in this case, which is called the circuit's *open-circuit voltage* (sv: *tomgångsspänning*),

$$u_\mathrm{oc} = \frac{R_2}{R_1 + R_2} U.$$

Now we can find another point in the u-i plane, where the terminals are *short*-circuited, as shown below. In this case, $u = 0$, and we need to find the corresponding current i.



Due to the short-circuit, the lower resistor can be ignored. Its top and bottom are at the same potential, so no current passes through it. The full voltage U is across the upper resistor, and the resulting current passes between the terminals. Ohm's law provides the circuit's *short-circuit current* (sv: *kortslutningsström*),

$$i_\mathrm{sc} = \frac{U}{R_1}.$$

If we accept the theorem that any two-terminal linear circuit must have a straight line in the *u*-*i* plane, then these two points have fully defined the circuit's *u*-*i* behaviour. The two components of a Thevenin (or Norton) 'equivalent source' now need to be chosen to give identical short-circuit and open-circuit properties to the ones calculated for our circuit above.

Given the Thevenin source equation

$$u = U_\mathrm{T} - iR_\mathrm{T},$$

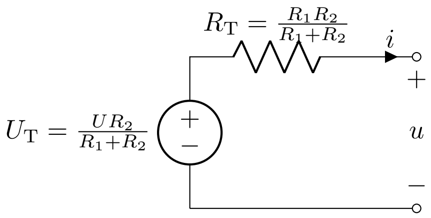
the open-circuit voltage is $u_\mathrm{oc} = U_\mathrm{T}$ and the short-circuit current is $i_\mathrm{sc} = U_\mathrm{T}/R_\mathrm{T}$. The Thevenin source voltage can therefore be directly fillen in as

$$U_\mathrm{T} = \frac{UR_2}{R_1 + R_2},$$

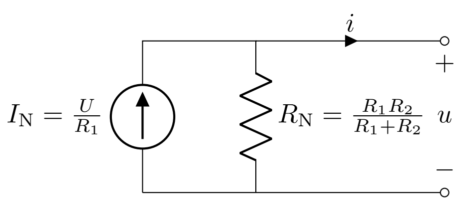
and the Thevenin source resistance is

$$R_\mathrm{T} = \frac{U_\mathrm{T}}{i_\mathrm{sc}} = \frac{\frac{UR_2}{R_1+R_2}}{\frac{U}{R_1}} = \frac{R_1 R_2}{R_1 + R_2}.$$

These can be written into the equivalent Thevenin source. To be extra careful, it would be wise to mark the two terminals, for example as 'a' and 'b', to ensure that the Thevenin source has the correct direction of voltage to match the original circuit; here we will just assume that the upper and lower terminal have stayed in the same positions as in the original.



If a Norton equivalent had been desired instead, its components could be determined by source transformation from the Thevenin equivalent, or directly from the short-circuit and open-circuit properties of the circuit that it is equivalent to. The short-circuit current of a Norton source is the current of its internal source: $i_\mathrm{sc} = I_\mathrm{N}$. The resistance is known to be the same as the Thevenin source's resistance, $R_\mathrm{N} = u_\mathrm{oc}/i_\mathrm{sc}$.



Short-circuit and open-circuit measurements can also easily be made on some physical circuits, using a multimeter to measure current in the short-circuit and voltage across the open-circuit.

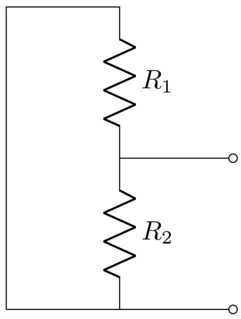
Of course, some physical circuits might have dangerously high open-circuit voltage or short-circuit current, in which case one has to measure other points on the *u*-*i* line instead. A related problem arises when analysing ideal circuits: if a circuit behaves as an ideal source then the *u*-*i* line only passes through one axis. For example, for an ideal voltage source there isn't a nice solution for short-circuit current. In such cases one can fall back to deriving a general equation for the *u*-*i* relation, or can try finding another point that is not a short-circuit or open-circuit. In the symbolic case these options are similar, but when dealing with numbers it might be easiest to look at a particular point.

2.3 Source-resistance directly

If the circuit that is being replaced with an equivalent contains no dependent sources, then the equivalent resistance can be found directly by combining resistors to a single resistor. That can sometimes be a lot easier than calculating the short-circuit *and* open-circuit conditions and then dividing long expressions for $u_\mathrm{oc}/i_\mathrm{sc}$.

To use this method, start by drawing the circuit with all the independent sources set to zero, so that they become short or open circuits as already described in the Section on superposition. There will then be only resistors remaining. Find the equivalent resistance between the two terminals: this is the equivalent source's resistance.

For the above example, there is just the one independent source. Setting it to zero, we replace this voltage source with a short-circuit.



The equivalent resistance of this circuit, between the two terminals shown on the right, is easily found by the rule for parallel resistors: note that R_1 and R_2 are now in parallel, when the voltage source is a short-circuit. The equivalent source's resistance is therefore

$$R_\mathrm{T} \; (= R_\mathrm{N}\;) \; = \frac{R_1 R_2}{R_1 + R_2},$$

as was found earlier by the method of $u_\mathrm{oc}/i_\mathrm{sc}$.

To find the two-terminal equivalent of a circuit that has only independent sources and resistors, it is common to find just the open-circuit voltage if wanting a Thevenin equivalent, or the short-circuit current if wanting a Norton equivalent, and then to find the equivalent source's resistance directly.

When there are dependent sources in the circuit, one cannot just use equivalent resistor rules, but has to consider forcing a particular voltage or current at the terminals and calculating the corresponding current or voltage, respectively. For solving a diagram, this is seldom any better than just using the short-circuit and open-circuit method; but, for solving with a simulation program or by measurement on a real circuit, this method might be useful.

2.4 The most general

The most general way to find equivalent-circuit parameters is to find two u,i points that do not *have* to be the open-circuit voltage and short-circuit current. Then calculate the necessary source and resistance from that line, to make the Thevenin or Norton equivalent.

This can occasionally be useful if dealing with strange circuits that have dependent sources that misbehave when short-circuited or open-circuited.[3] If it feels strange to define some arbitrary current or voltage, try drawing a current or voltage source connected between the pair of terminals where you are trying to find the equivalent. Define the source's value, which will fix either i or u, then calculate the other quantity.

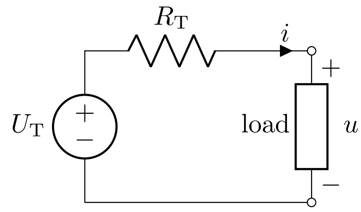
In practical situations one often *measures* the current and voltage at two terminals of a circuit, for two different states of loading, e.g. different resistances connected to the terminals. Then the equivalent parameters are calculated from these two points. It would often be impractical to do a test with a real short-circuit or open-circuit.

3 Maximum power transfer

A common question of interest is what maximum power can be obtained from some circuit that has two terminals. From a circuit theory perspective, this question does not consider such petty details as whether wires will melt: it is idealised, considering just the u-i relation that the circuit has.

The circuit behind the two terminals can be modelled by a Thevenin or Norton equivalent, subject to the conditions described in the previous Section. A Thevenin equivalent, with a generic component connected to its terminals, is shown here. The generic component is called the *load* (sv: *last*). It is common to assume the load to be a resistor whose value we can chose; but we start with the more general case where the load could be a voltage or current source, or even another circuit with a Thevenin equivalent model!

[3]One example of a strange circuit is Question 3c, EI1120 exam 2014-03-20, [pdf].



The question is what power can be got *from* the Thevenin circuit shown at the left of the terminals, *to* this other two-terminal thing that we call the load.

A Thevenin equivalent, with finite positive resistance, *does* have a maximum power that it can supply out of its terminals. Nothing that is connected at the terminals can extract more power than this. (The same applies, of course, to a Norton equivalent.)

Consider the behaviour of the Thevenin equivalent in the u-i plane. From the relation

$$u = U_T - iR_T$$

the power out from the terminals *to* the load can be expressed as a function of the current,

$$P = ui = U_T i - i^2 R_T,$$

which can be seen to be the power provided by the voltage source U_T, minus the power consumed by the Thevenin resistance R_T.

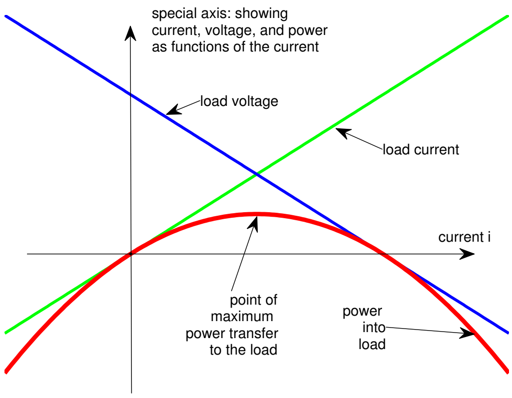
If the load is like an open-circuit, taking no current from the source, $i = 0$, then $P = 0$: there is no power into or out of the load. In this case, there is no power in or out of any of the three components.

If the load is like a short-circuit, taking so much current that there is zero voltage at the terminals, $u = 0$, then there is also no power transfer between the Thevenin equivalent and the load. In this case, all the power out of the Thevenin equivalent's voltage source is then lost in the Thevenin equivalent's resistance.[4]

If we force even more current through the circuit, so that $i > i_{sc}$, by having a 'load' that is a voltage or current source that helps the current to flow, then power flows *into* the Thevenin equivalent.

These situations are shown in the following figure, where the horizontal axis shows the current i. The vertical axis shows three different variables, that are plotted against current in the three curves.

[4]For amusement, notice a difference between the Thevenin and Norton cases: it is in *open*-circuit conditions that the Norton equivalent's current source is feeding a high power to the equivalent's resistance. We only require that the two types of source behave in the same way at their terminals: the internal power consumptions are neglected.



The voltage u decreases from the Thevenin equivalent's open-circuit value at $i = 0$, to zero when $i = i_{\text{sc}}$. A line showing the current i is marked as well, but this is trivial as it is just a plot of i versus i … its purpose is to make clearer how the *product* of the u versus i and i versus i lines gives the quadratic curve of P versus i.

This curve of P, the power delivered to the load, shows that at the open-circuit and short-circuit conditions there is zero power transfer between the Thevenin equivalent and the load, and at currents *between* these two points there is a power transfer from the Thevenin equivalent to the load, with a maximum point in the middle.

For currents outside this range, the power flow is *into* the Thevenin equivalent, which indicates that the load in such cases cannot be a normal resistor! There is no limit to how much power can be pushed into the Thevenin equivalent. The limit on the power *out* comes from the Thevenin resistance consuming the power that the Thevenin voltage-source generates: the power to the resistance increases quadratically with current.

How can we find the current needed for maximum power transfer to the load, without just guessing by looking at the curve? The *maximum point* corresponds to a *zero gradient*, which means that the derivative of P versus i curve must be zero at the point where maximum power is obtained. (We can see from the curve that this curve has a maximum, not a minimum.) This derivative is

$$\frac{\mathrm{d}P}{\mathrm{d}i} = \frac{\mathrm{d}}{\mathrm{d}i}\left(U_{\text{T}} i - i^2 R_{\text{T}}\right) = U_{\text{T}} - 2iR_{\text{T}},$$

which has its zero point when

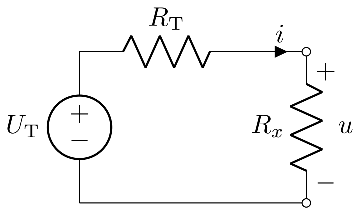
$$i = \frac{U_{\text{T}}}{2R_{\text{T}}}.$$

Thus, the maximum possible power is obtained from the Thevenin equivalent when its current half of its short-circuit. You should also be able to show that the voltage u at this point is half of the open-circuit voltage.

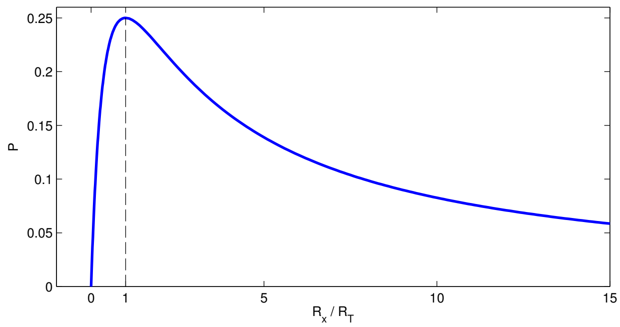
3.1 Maximum power to a resistor

All the above was based on a load of "something" connected to the two terminals of a Thevenin equivalent; this is the general case of getting maximum power, where we even consider negative currents and currents greater than the Thevenin equivalent's short-circuit current.

More traditionally, *maximum power transfer* considers the load to be just a resistor, as with R_x in the following circuit. We can vary the resistance to achieve maximum power transfer to the load, from a given Thevenin equivalent. In this more limited case, negative powers are not possible, as a resistor cannot provide power.



When the resistor is varied from zero to a huge resistance, the situation changes from short-circuit to open-circuit conditions for the Thevenin equivalent. This corresponds to the range of currents for which positive powers were transferred to the load, in the previous Section. The following figure shows the power transferred to the load resistor R_x, as a function of R_x/R_{T}. The numbers shown for power can be assumed to be 'arbitrary units'.[5]



This was calculated by

$$P = R_x i^2 = R_x \left(\frac{U_{\text{T}}}{R_{\text{T}} + R_x}\right)^2.$$

In the previous Section, the maximum power transfer was found to occur when the current to the load was

[5] Assume them arbitrary for simplicity … but in fact, the numbers plotted here are relative to the product $u_{\text{oc}} i_{\text{sc}}$ which is equivalent to $U_{\text{T}}^2/R_{\text{T}}$, which has the physical meaning that it's the power that would go from the Thevenin equivalent's voltage source to its resistor if you short-circuit the Thevenin equivalent. At "maximum power transfer" conditions, only half of this power comes from the source, as the current is half of the short-circuit current; then half of *this* gets wasted in the Thevenin resistance; so only a quarter gets to the load.

$i = \frac{i_{sc}}{2} = \frac{U_T}{2R_T}$. When the load is a resistor, this current will occur when the total circuit resistance is twice the Thevenin resistance; thus, the maximum power criterion is that

$$R_x = R_T.$$

This is the way that the maximum power transfer theorem is usually expressed. Deriving it directly for R_x involves a rather uglier derivative, and does not show the generality beyond resistive loads.

3.2 Who cares?

The interest in maximum power arises in many cases.

In electronic design, one might want to choose a load that will extract the maximum possible signal power from a source of known Thevenin resistance. Or one might want to know the worst case of how much power a given part of a circuit would be able to put into a variable resistor, to check that the resistor won't get burned.

In electric power engineering there is a wide use of Thevenin equivalents to model the rest of the power system that is hidden behind the 230 V socket outlet, or behind the 400 kV busbars in a substation. In some highly stressed cases in the high-voltage system, the loading can approach the maximum power transfer conditions, and some types of load then try to take even more current, leading to "voltage collapse".

3.3 Special cases

A load resistance can be chosen with a value that maximises the power in this resistance *for a given resistance in the source*. Making R_x too high will make the current very low, so $P = ui$ is low. But making R_x too low will make the *voltage* very low, also making $P = ui$ be low. Somewhere in between these extremes is the highest product of the output voltage and current, $P = ui$. The source's resistance causes the terminal voltage to decrease when more current comes out in the direction that can give power to the load.

The other way round is meaningless: if the task is to maximise the power in a given load resistor by choosing the *source* parameters, then the best case is to have a zero resistance of the source, $R_T = 0$. If the source voltage U_T can be chosen, it should be as big as possible.

Another case where there isn't really a well defined maximum power is when it is the *load* that can be chosen, but the source is a pure voltage source or current source. A Thevenin source with zero resistance is an ideal voltage source. It will hold its constant voltage U_T across any load resistor R_x that is connected to it. In this case, we see the power in the load, as a function of the load resistance, is $P_x = U_T^2/R_x \ldots$ which just gets bigger and bigger as $R_x \to 0$. (In its dual, a Norton source with current I_N and infinite internal resistance will supply more and more power to a load R_x as the value $R_x \to \infty$, since in this case $P_x = I_N^2 R_x$.)

See the Extra section for some subtleties of negative source resistance: we have assumed in the above that the source and load have *positive* resistance!

4 — Extra —

4.1 Links

Some interesting history of the equivalent circuits is found here, for `[Voltage-source]` and `[Current-source]`.

There is a far-from-complete page on Wikipedia, summarising and linking to `[CircuitTheorems]`. The `[MaxPower]` theorem is presented in the more classic textbook way on Wikipedia, by differentiating with respect to load resistance, which provides a rather less elegant working than the method we introduced earlier in this Chapter.

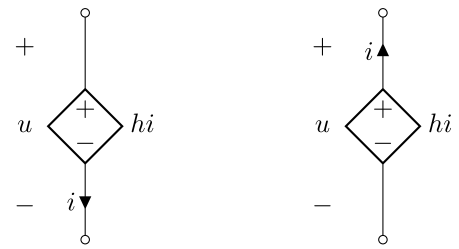
The following is rather advanced. It's an interesting paper about ignoring the usual rules given by circuit courses on how to treat dependent sources when using superposition, `[SuperposDepSrc]`. The principle seems sound to me. It appears that for some of his examples, this alternative method (of treating dependent sources in the same way as independent sources, but with a value that has to be matched with the requirements of a controlling variable) may actually help the solution. We can see that the examples are rather typical small-signal models of transistor circuits. What makes it a particularly interesting paper is its history of being rejected from publication in two journals, in one case for the silly reason that it couldn't solve an undefined problem with series current sources!

More about the `[Wye-Delta]` transform (mentioned in Section 4.3) and its more general form of the `[Star-Polygon]` transform, can be seen on Wikipedia.

4.2 Negative resistance

A resistor is a basic component in circuit theory. When we start dealing with two-pole equivalents, of Thevenin or Norton type, we talk of the equivalent source's resistance. If the circuit that the equivalent is modelling contains just independent sources and resistors, then the source resistance R_{T} or R_{N} will always be positive, just like the resistors that we're familiar with.

However, when dependent sources are included, it is possible to have circuits whose equivalent resistance is *negative*. This is not surprising when one considers a very simple case of a dependent source where the controlling variable is that same source's other variable. For example, take a current-controlled voltage source, and define its controlling variable to be the current *through* this same source.



Depending on the relative direction of the current and voltage definitions, this could look like a positive or negative resistance. In the example shown at the left, the gain h is equivalent to a resistance: the relation of i and u is $u = hi$. By making this value negative, or by swapping the current definition as shown on the right, the component looks like a negative resistor, such that $u = -hi$ when the current is defined into the positive side of the voltage reference direction.

Looking at the equations for Ohm's law or power dissipation, the nature of a *negative resistance* is that:

1. Its current is proportional to the voltage across it, but always in the direction that flows *from low to high* potential; and
2. It therefore never absorbs power, but can only produce power.

I have said that dependent sources and negative resistance are mainly a matter for people dealing with "small signal models" of electronic circuits. I don't retract that claim, but it's only "mainly". There are examples relevant to electric power engineering where sources can have negative resistance, due to some form of feedback.

More than a hundred years ago, dc (likström) generators were being used to supply loads such as lighting. A magnetic field is needed, to allow the generator's rotation to induce a voltage in the moving conductors. This field could be produced by passing a current through a set of thin copper "field coils". The current was supplied from a voltage source, with a variable resistance in series to allow the desired field current to be set and thus the desired output voltage to be generated. The output voltage could even be used as the voltage source for the field current.

However, the voltage supplied to the load will fall when the load draws more current, such as when more lamps are connected. That is like a Thevenin source with a positive resistance. The physical reason is that there is resistance in the wires *in* the generator, and in the wires *between* the generator and load; there is also an effect of the load current producing a magnetic field that reduces the field in the generator.

To help maintain a steady voltage at the load, engineers made the wires carrying the load current out of the generator wrap a few times around the same pieces of iron where the main field coils were producing magnetic field; this was done in the direction so that higher load current would *assist* the field coils by increasing the magnetic field, and therefore would increase the voltage generated! By careful choice of the number of turns of wire in these coils, the generator could be made to have a quite constant voltage at its terminals, in spite of different load currents: that is like a near-zero Thevenin resistance. Indeed, by even more turns of the load current around the field coils, the generator could be made have an *increasing* voltage with increasing load. This was sometimes desired, so that the generator's increase would balance the voltage drop in the line connecting the load to the generator: then the voltage near the loads would remain quite steady. In this case, the generator has a negative resistance, since its output voltage rises when supply more current into a normal load that has positive resistance.

That is called line-drop compensation. The generator's negative resistance can be seen as cancelling the line's positive resistance. It is still done in some modern power distribution systems, but nowadays it is done by controlling variable transformers, usually by switching their turns ratio in discrete steps. With ac, the situation is a little more complicated, too.

Still, the principle is that a source can appear to have negative resistance. This is usually only valid within a narrow range of voltage or current, e.g. $\pm 10\%$ of the normal value, due to the design of the equipment.

4.2.1 Max-power with negative resistance

To use the maximum (delivered) power theorem sensibly, we should really add the condition that the source's resistance R_{T} is positive. Then it makes sense to say that the maximum possible power delivered *by* the Thevenin or Norton equivalent source happens when the load behaves as a positive resistance equal to R_{T}. If R_{T} and R_x are both negative, the maximum power point is still true, but in the sense that it is a maximum of the power *out* of R_x.

In the main Section on equivalent sources, we saw that there was no meaning to a maximum power when the source is an ideal voltage source or current source. The idea of a finite maximum value of power makes sense when the source *limits* the available output power by having a finite resistance. This source resistance, in a Thevenin source, causes the output voltage to decrease when there is an increase in the current coming out in the direction that can provide power from the source U_{T}.

If a Thevenin source has negative resistance, then other strange things can happen. If we just follow the rule of "$R_x = R_{\mathrm{T}}$ for maximum power", this tells us that the load resistance R_x should also be negative for maximum power. But if we plot the function of power into the load versus load resistance, or versus current, we see that this choice of load resistance gives the maximum value of power *out* from the load *into* the source! The maximum power theorem therefore still works when both resistances – source and load – are negative, but the power in question is negative, meaning that it is generated by R_x. After all, a negative resistor *cannot* absorb power, just as a positive resistor cannot generate power.

However, if we allow the source and load resistances to have opposite signs, more interesting things can happen! Choosing R_x very close to $-R_{\mathrm{T}}$ will result in the complete circuit loop resistance, $R_{\mathrm{T}} + R_x$, being very close to zero: the current around the loop is therefore very high. By the relation for load power $P = i^2 R_x$, this results in a very high power to the load: the equivalent source's negative resistance is supplying power to the positive load resistance. Whether the equivalent source's voltage source is supplying or consuming power depends on the direction of current, which depends on which of the two resistances has the higher absolute value. So in this case of source and load resistors with opposite sign, there's again no real meaning to a maximum power: you can have as much as you want!

Looking graphically, one can draw the lines in the u-i plane to represent a source with negative resistance, and a positive load-resistor connected to it. It's convenient to draw the current out of the source and in to the load, so that the current and voltage reference directions are the same for the equivalent source and the load: then the point where the two lines meet in the u-i plane is the solution of the circuit. In this condition of a positive and negative resistance of quite similar magnitude, the lines are nearly parallel, but the Thevenin voltage gives a displacement between the lines. The crossing point of the lines, which is the solution of the circuit for u and i, will therefore occur at a very large voltage and current. This is true whether the load has a positive and the source has a negative resistance, or vice versa. The difference is only in the direction of power flow: remember that a positive resistor can only consume, and a negative resistor can only produce power. On the other hand, if both resistances are, say, positive, then the two lines have very different slopes, one negative and one positive, since we defined the current in the loop to go into the voltage positive reference for the load and out of the voltage positive reference for the equivalent source. There is then a solution at a moderate value of voltage and current, and it cannot be changed much without making the resistances very low. From the above, we note again that there is only a meaning to the "maximum power" when the source and load resistance are constrained to have the same sign.

If you're puzzled about "resistances" producing power, just remember that these resistances are simply a voltage/current relation created by more complex devices that have their own power sources; a real resistor made of a chunk of metal or carbon, etc, only has positive resistance.

4.3 Other Circuit Theorems

The following will not be required for solutions in this course. They are mentioned for “the interested reader”, as just a sample of the various theorems that people have thought it worthwhile to publish: the usefulness is sometimes not obvious unless one has worked with the specific sorts of design or analysis that the theorem was developed for.

4.3.1 Tellegen’s theorem

For each branch of a circuit, take the product of voltage across the branch and the current through it, using the same relation of the directions for each branch (for example that the current reference goes into the positive voltage reference, i.e. ‘passive convention’). The sum of all these values is zero. This is probably “more useful than it sounds”. It may seem obvious, thinking in terms of “power generated equals power consumed”. But its generality is claimed to be surprisingly useful. It is helpful in showing some other theorems.

4.3.2 Reciprocity

In general, reciprocity means that if we know how an input at one point affects a second point, then we also know how an input at that second point would affect the first point.

Reciprocity is found in fields and circuits. In a circuit example, if applying a voltage u across branch 1 results in current i in branch 2 (somewhere in the same circuit), then current i in branch 1 will result in voltage u in branch 2. If there are other sources in the circuit, then the current and voltage mentioned above should be seen as changes (i.e. as superposition of the other sources and the thing we’re adding).

This reciprocity is true for the circuits we’ve seen as long as there are *not* dependent sources. If there are dependent sources, we could easily make a circuit where there is only a one-directional causal relation between two parts, in which case there is very obviously no chance of the ‘backward’ relation happening: think of two separate loops, where one contains the controlling variable and the other contains the source that it controls.

4.3.3 Substitution

Consider a part of a circuit: e.g. a single component or a more complicated two-terminal ‘chunk’ of the circuit. The solution of the complete circuit tells us the current through and voltage across this part. If the part is replaced by a voltage or current source whose value is that voltage or current that the part had in the circuit, then For example, if one knows there is voltage u across an R ohm resistor in the circuit solution, then this resistor can be replaced by a voltage source u, or by a current source u/R, and all the currents and voltages in the circuit will remain the same as before.

Of course, this does not *mean that one can then change* another *component to some arbitrary value.* It has to be a specially chosen value, that fixes one of the two ‘degrees of freedom’ (the quantities voltage and current) to the same as in the original case. And it’s only valid for that particular case of the rest of the circuit: if one does a valid substitution but then changes something elsewhere in the circuit, the value of the substituted component would have to change in order to give the voltage and current that the original component would have in that new situation. Note the big difference compared to, for example, a Thevenin equivalent, which is a genuine equivalent for all cases of what we connect to it.

To try to justify the substitution theorem, recall that linear two-terminal components or circuits all give straight lines in the u, i plane, at varying gradients and offsets: this includes the main components – U, I, R – as well as the Thevenin or Norton sources. For any interface between two parts of the circuit, such as between a two-terminal component and the rest of a circuit, the solution to the circuit is the point where the lines for these two parts intersect.[6] As long as a substituted component has a line that passes through the same solution point as the original component, it will give the same solution. However, if other parts in the circuit are changed, then the line describing these other parts changes, and the new intersection point (solution) will be different depending on what type of component we used for the substitution; this shows how the substitution is only valid for the specific case.

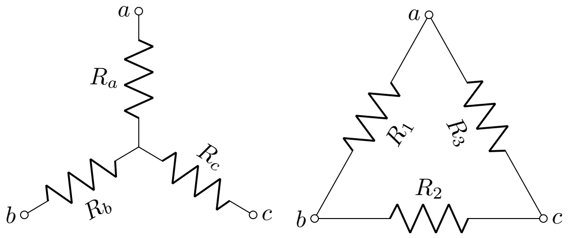
4.3.4 Star-Polygon or Nodal-Mesh (Rosen’s)

Given a set of N nodes, and one other node that is connected to each of the N by an arbitrary impedance (i.e. it can be a different impedance for each node), and equivalent behaviour is given by a set of N nodes in which every node has a direct connection to every other node (a polygon round the outside, then lots of other connections between, as N becomes large!).

4.3.5 Star-Delta transform

The *star-delta* (Y-Δ) transformation is the simplest case of the node-mesh transformation. It has just three nodes that are ‘seen’, and a hidden middle node in the star version of the circuit. It converts between two different ways of connecting three resistors between *three* terminals.

[6] This is assuming that the lines *can* intersect, so they are not parallel. It also assumes current and voltage directions have been suitably defined for both lines. In a nonlinear circuit there could be more than one solution… but we don’t consider that case.



These two connections can be made equivalent, when seen from the terminals a,b,c, as long as component values fit the following conditions.

The required condition for the Y resistance connected to a particular terminal is that the values of the two Δ resistances joining to that terminal are multiplied then divided by the sum of all the Δ resistances. For example, $R_a = \frac{R_1 R_3}{R_1+R_2+R_3}$. Thus, if all the Δ resistances are the same, then all the Y resistances are 1/3 of this value.

For the Δ resistance connected between two terminals, the Y resistance that is *not* connected to either of these terminals is divided *into* the sum of products of Y pairs. For example, $R_1 = \frac{R_a R_b + R_b R_c + R_c R_a}{R_c}$. If all the Y resistances are the same, then all the Δ resistances are 3 times their value.

4.3.6 Source-shift (Blakesley's)

Given a black-box with N terminals, the effect of a series voltage source in one of these terminals is the same as of an oppositely directed source on every one of the other terminals.

In its dual, with current sources, a current source between two nodes can be replaced by a chain of similar current sources in a path between the same nodes but looping over other nodes: at each intermediate node, one source takes out all that the other put in, so there is no change to that intermediate node.

4.3.7 Parallel Generator (Millman's)

In the most general form, Millman's theorem is about calculating the potential at a node where several impedance branches join, where the other end of each branch is at a known potential.

A more specific case is parallel branches, i.e. finding the voltage across the whole set, when some may be sources with impedance, and some just impedances. An example is three-phase sources and impedances, such as for finding the potential of the star-point of a non-grounded Y-connected load.

Probably the most famous context of Millman's theorem, largely synonymous in power engineering, is the "parallel generator" case, where multiple Thevenin-type sources are connected in parallel to a 'bus' (node), in parallel also with a load impedance. This is clearly useful when multiple power sources (e.g. generators or batteries), are connected in parallel to feed a load. Millman's theorem says that the voltage between the terminals is the product of the sum of short-circuit currents of the separate sources, and the total parallel impedance (resistance, in the dc case) of these sources and the load.

The relation can be seen easily from the Norton equivalent of the whole circuit. Between the terminals, the circuit's short-circuit current is clearly the sum of individual short-circuit currents, as all the sources connect between these two nodes. The Norton impedance Z_{N} (R_{N} in the dc case) is the sum of all the resistances of the sources and the load. To see this, consider the method of setting all independent sources to zero, to find the equivalent impedance. The Norton source's open-circuit voltage is known to be $U_{\mathrm{oc}} = I_{\mathrm{N}} Z_{\mathrm{N}}$.