BJT switches

Posted on April 4, 2008, by Ibrahim KAMAL, in General electronics, tagged

*The **B**ipolar **J**unction **T**ransistors (BJT), is one of the oldest but most famous transistors, that still find their place in modern electronics. They come in many classes and sizes to match your application, and with some basic knowledge and some simple mathematical relations, you can build very efficient while very simple switch circuits. Using a transistor in the 'switching' configuration is indispensable in many situation, where it is required to drive a relatively important load, providing important currents, that most controllers are not able to provide.*

Introduction to the BJT

Before going on with the tutorial, let us agree that we are going to discuss one among many other configurations of Bipolar Junction Transistors. The same BJT can be used in an audio amplifier or and FM receiver, however, usually each model of BJTs is manufactured with specific characteristics to make it more adequate to a certain application.

Configuring a transistor, means connecting to it a number of external components and power sources in a very specific way to make if function in a very specific and controllable way.

The BJT has 3 pins. The Base, the Collector, and the Emitter (see ***figure 1***). According to the configuration of the transistor, it can be difficult to precisely say which one of those pins is input, and which one is output, as they are all related, and all affect each others somehow. However, in a switch configuration, this is clearly defined: The Base current controls the flow of current from the collector to the Emitter or vice-versa, depending on the type of BJT.

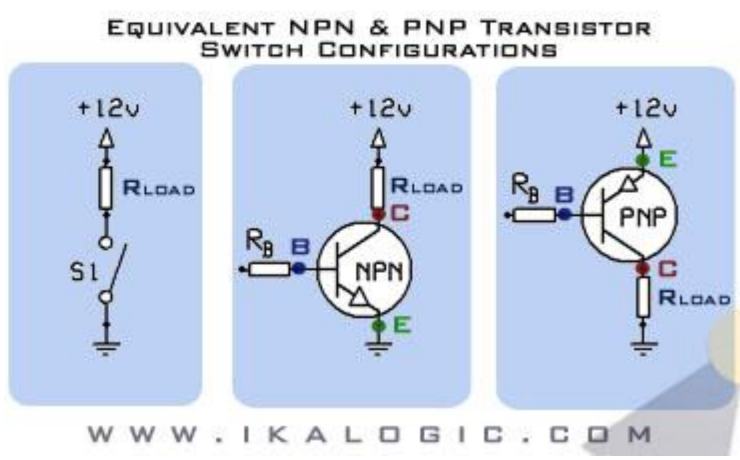

figure 1

There are two main types of BJTs: NPN and PNP. To make this as simple as possible, we can say that the main difference between those two types is the directions of the electric currents. This can be seen in ***figure 1***, where the direction of the arrow points towards the direction of the current; In the NPN transistor, one current flow from the base to 'inside' the transistor, and another current flow from the collector to the Emitter, while in a PNP transistor, all the directions are reversed, one current flows from the 'inside' of the transistor to the base, and another one flows from the emitter to the collector. From a functional point of view, the difference between those two types, is the voltage being provided to the load. As you can see in ***figure 1*** an NPN transistor provides 0V when switched ON, while a PNP provides 12V. You will later understand why this affects the choice of the type of transistor.

For simplicity, we are going to study only one of those two types: the NPN BJT, but the same is applicable to PNP, taking into consideration that all the currents are inversed.

Figure 1 shows a very clear analogy between a real switch (S1) and a transistor switch, where it is clear that the base current will which ON or OFF the path of the current from the collector to the emitter. There are some limitations though as compared to a regular mechanical switch, which are principally the direction of the current which is fixed by the type of transistor and the magnitude of that current which depends on the characteristics (also called parameters and ratings) of the transistor.

Knowing exactly the parameters of a transistor is essential to be able to get the most out of it. The main required parameter is the current gain factor between the base and collector current, usually noted ***Hfe*** or ***β*** (Beta). Then it is important to also know the maximum currents and power the transistor can dissipate as well as the maximum applicable voltages across it's pins.

Those parameters, which can be found in the datasheet of the transistor, will help us to determine the best value of the base resistance, as you will see in the next section of this tutorial.

Basic NPN switch configuration

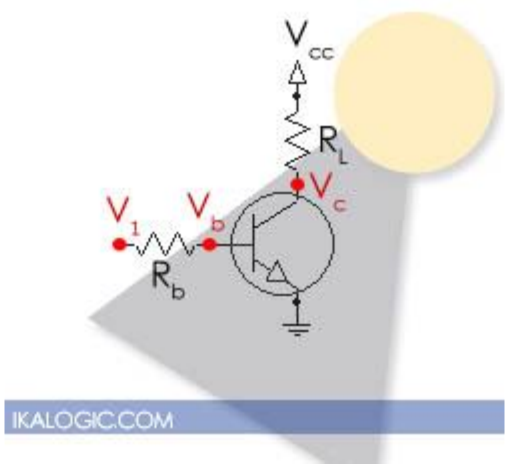

figure 2.A

The diagram shown in ***figure 2.A*** is the basic switch configuration for an NPN transistor. You will encounter this figure very often when analyzing different electronic circuits. We will study how to 'configure' the transistor in switch mode, by choosing the right base resistor with respect to other parameters that we will consider fixed as the Vcc voltage, the Current gain of the transistor, and the Load resistance. To do this, there exists many methods. I am proposing the simplest i've found, which is still fully functional and precise. It goes through the following steps:

1-Assume the transistor is in saturation mode:

By making such an assumption, the mathematical model of the transistor becomes extremely simple, and some node voltages become known like ***Vc***. Actually, we will find the value of the base resistor that justifies such an assumption.

2- Determine the Collector saturation current:

With the assumption above, ***Vce*** (the voltage between the collector and the emitter) is known and can be fetched from the datasheet of the transistor. The emitter being connected to GND, $V_{ce} = V_c - 0 = V_c$. The current can be then calculated using the following relation: $I_c = \frac{V_{cc} - V_c}{R_L}$

Sometimes, the load Resistance (R_L) is unknown or irrelevant, like the resistance of the winding of a relay; In such a case, knowing the required current to activate the relay is sufficient.

If you don't have any mean of precisely knowing the value of the collector current (I_c), then estimate it to the maximum probable value, after all the current will never exceed what the load can support. However, make sure that the collector current caused by the load does not exceed the maximum rated current of the transistor.

3- Calculate the needed base current:

Knowing the Collector current, you can calculate the minimum base current to reach that collector current using the following relation: $I_c = \beta I_b$, Then,

$$I_b = \frac{I_c}{\beta}$$

4- Overdriving:

Once you have calculated a base current, if it's value is far below the maximum rated base current of the transistor (from the datasheet), we ***may*** overdrive the transistor, by multiplying the calculated base current by a factor of 10 for example. This way, the resulting transistor switch will be much more immune again changes in the load. In other words, the performance of the transistor wont be reduced if the load eventually increases due to any unknown reason. ***Be careful not to exceed or even get too close to the maximum base current defined in the datasheet of the transistor.***

5- Calculate the needed value of R_b:

Taking into consideration an overdriving factor of 10, the base resistance (R_b) can be calculated using the following formula: $R_b = \frac{V_1 - V_b}{I_b \times 10}$

Where V_1 is the voltage controlling the transistor (see ***figure 2.A***).

But the emitter being connected to ground, and the base to emitter voltage being known (approximately 0.7V in most switching transistors), and also assuming that V_1 is a 5V TTL voltage, the formula can be simplified to the following form: $R_b = \frac{5 - 0.7}{I_b \times 10} = \frac{4.3}{10.I_b}$

Where you can still see that I_b is multiplied by 10 for overdriving.

When the value of R_b is known, the transistor is 'configured' to function as a switch, also called 'in saturation and cut-off mode' where 'saturation' is when the transistor is fully switched ON, and 'cut-off' is when it is fully turned OFF and no current is passing through the load.

Some common sense about the electric currents in a transistor: When we say that $I_c = \beta I_b$, that does mean that the collector current ***has*** to be equal to 'βI_b', neither that it is forced to be equal to that value. It simply means that the collector of the transistor can sink or source this amount of current. The actual value of the current will follow Ohm's laws, like any electrical current.

Calculation precautions:

When we assumed that the transistor was is saturation mode, we also assumed some of its parameters were constant. This was not exactly true. Actually those parameters change, mainly due to increasing the collector current, and that's why it is safer to overdrive. Each datasheet will present those variation of the parameters as a table or as a graph. Anyway, you have to consider those parameters into your calculations. For example, the table in ***figure 2.B*** shows two parameters (among others) that considerably change:

H_{fe} (β) changes according to the collector current, and the voltage V_{CEsat}. But V_{CEsat} itself changes according to the collector current and the base current.

Another part of the table is highlighted: the V_{BEsat}, which varies depending on collector and base current.

h_{FE}	DC current gain	I_C = 0.1 mA; V_{CE} = 10 V	35	–	
		I_C = 1 mA; V_{CE} = 10 V	50	–	
		I_C = 10 mA; V_{CE} = 10 V	75	–	
		I_C = 150 mA; V_{CE} = 1 V; note 1	50	–	
		I_C = 150 mA; V_{CE} = 10 V; note 1	100	300	

V_{BEsat}	base-emitter saturation voltage 2N2222	I_C = 150 mA; I_B = 15 mA; note 1	–	1.3	V
		I_C = 500 mA; I_B = 50 mA; note 1	–	2.6	V

I_{EBO}	emitter cut-off current	I_C = 0; V_{EB} = 3 V	–	10	nA
h_{FE}	DC current gain	I_C = 500 mA; V_{CE} = 10 V; note 1			
	2N2222		30	–	
	2N2222A		40	–	
V_{CEsat}	collector-emitter saturation voltage				
	2N2222	I_C = 150 mA; I_B = 15 mA; note 1	–	400	mV
		I_C = 500 mA; I_B = 50 mA; note 1	–	1.6	V
V_{BEsat}	base-emitter saturation voltage				
	2N2222	I_C = 150 mA; I_B = 15 mA; note 1	–	1.3	V
		I_C = 500 mA; I_B = 50 mA; note 1	–	2.6	V
V_{BEsat}	base-emitter saturation voltage				
	2N2222A	I_C = 150 mA; I_B = 15 mA; note 1	0.6	1.2	V
		I_C = 500 mA; I_B = 50 mA; note 1	–	2	V
C_c	collector capacitance	I_E = i_e = 0; V_{CB} = 10 V; f = 1 MHz	–	8	pF
C_e	emitter capacitance	I_C = i_c = 0; V_{EB} = 500 mV; f = 1 MHz			
	2N2222A		–	25	pF
f_T	transition frequency	I_C = 20 mA; V_{CE} = 20 V; f = 100 MHz			

figure 2.B

Calculation can become extremely complicated, as all the parameters are tightly and complexly related, so the solution is to assume the worst values. In other words, assume the smallest H_{fe}, the biggest V_{CEsat} and V_{BEsat}, and at last, don't forget to overdrive the transistor, but increasing the base current.

Typical Transistor switch applications

It's a habit in **Ikalogic** to always tightly bond theory and practice, so here are a couple of example applications using the very standard 2N2222 switching transistor.

1. Controlling a relay:

A very classic use of transistor in modern electronics is to control 'electro-magnetic' relays, that incorporates a coil, that sinks important amount of current (up to 200 mA), and generate feed-back voltages when switched OFF. If you want a logic gate or a microcontroller to control a relay, it is indispensable to use a transistor.

In ***figure 3.A***, the base resistance is to be calculated according to the required current to activate the coil of the relay.

The Diode D1 is very important, it protects the transistor from the current surges generated by the coil when switching OFF. It is called a freewheeling diode, and it is a regular rectifier diode like the 1N4007.

You can connect any 220V or 110V appliance to the output of the relay.

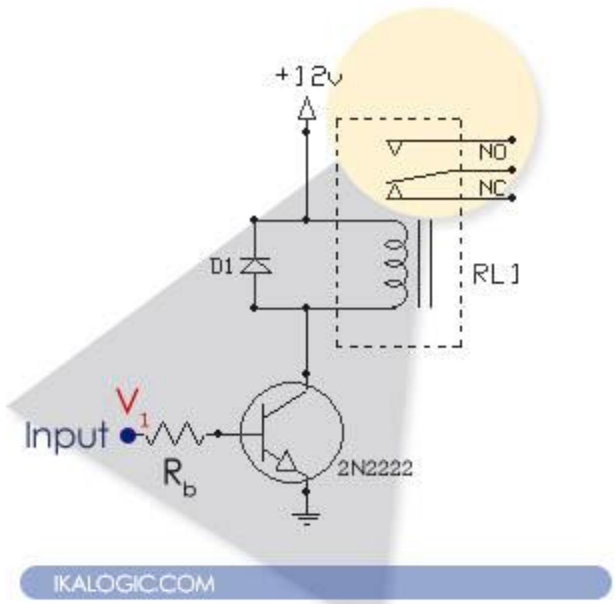

figure 3.A

2. Connecting a transistor to an open-collector output:

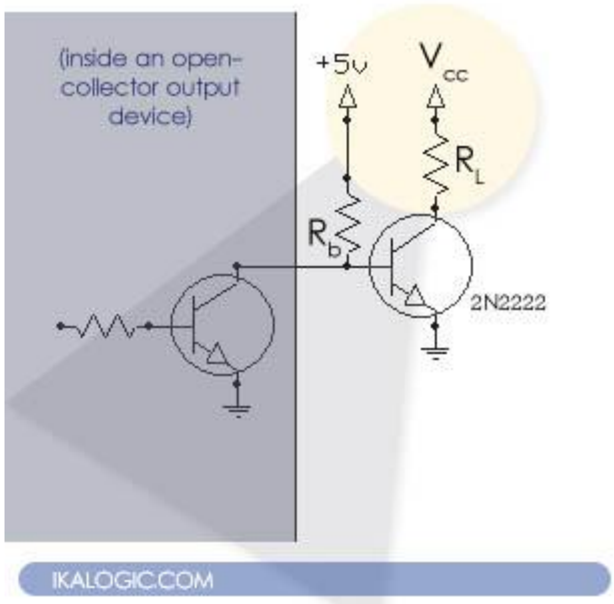

figure 3.B

Many devices like the 8051 family of microcontrollers have open-collector output ports. Such output configuration is unable to 'source' any current, it can only sink current. When connecting a switching transistor to such a device, the position of the base resistance changes, to sink current from the 5V power supply instead of the device that is unable to do so.This way you can calculate the base resistance normally as described in this tutorial.

Note that the representation of the output device is just for clarifying the concept of open collector output, but in general, output ports are more complicated, and often use FET transistors instead of BJTs, and is called 'open drain' output, but the mechanism remains exactly the same and is fully compatible with the configuration given in ***figure 3.B***.

3. Building NOR Logic gates:

Sometimes, on a circuit, you need a single NOR logic gate, but you don't want to use a 14 pin IC containing 4 gates, and only use one of them.

Either due to the cost or due to the foot print (space taken on the PCB), building a gate using a couple transistors can be interesting. However, note that the frequency response of such a gate depends on the characteristics and type of the transistor, but below 100 Khz, you probably wont notice any difference of behavior.

This method also allows you to build transistor with controllable Input and output resistance given in ***figure 3.C***. Decreasing the output resistance (R_o) in this case would increase the power consumption of the gate, but increase the output current.

It's up to you to find the compromise between the different parameters.

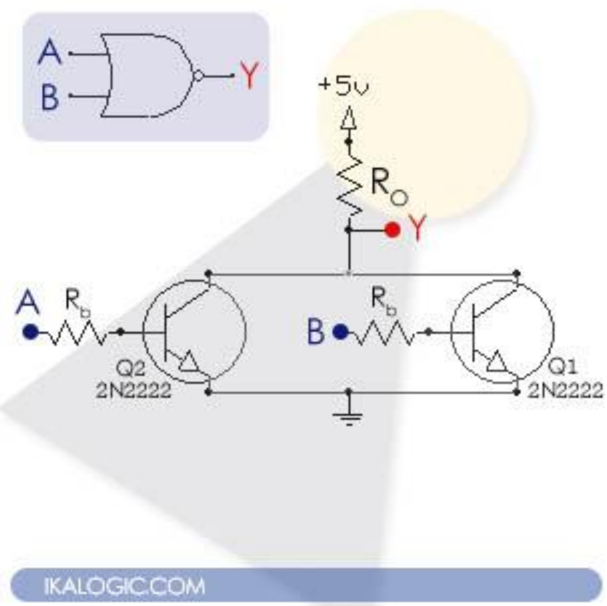

figure 3.C

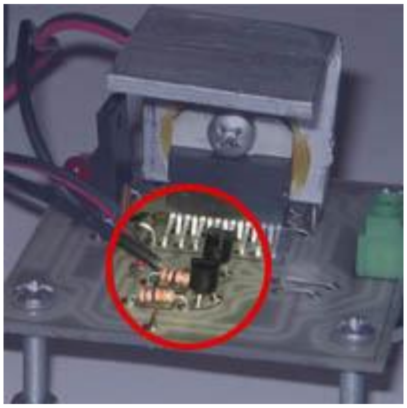

figure 3.D

Figure 3.D shows a NOR logic gate constructed using two 2N2222 BJT transistors. You can see the full schematic of this H-bridge here. You can notice the assembled gate takes relatively less space that a DIP14 logic IC (DIP stands for Dual Inline Package)

Note that you can also build a NAND logic gate using the same two transistors by connecting them in series (not in parallel as in ***figure 3.C***).

All those applications can be built on the complementary PNP transistor 2N2907, with minor modifications. You just have to consider that all the electric currents flow in the opposite direction.

The H-Bridge is a very famous example application to switching transistors, where two PNP and two NPN transistors are used, as in the following schematic, or in our 5A H-Brigde tutorial.

Debugging transistor switch circuits.

When a problem occurs in circuits containing many transistors, is can be very problematic to know which one is defective, specially when they are all soldered. I propose a bunch of tips to help you find the problem in such a circuit in minimum time:

1- Temperature:

If a transistor is getting hot, this probably men there is a problem somewhere. But don't be too confident saying that the problem is coming from the transistor getting hot. Usually a defective transistor don't even get warm any more. This rise of temperature can be caused by another transistor connected to it. Bottom line: The temperature is an indication of the region where things are going wrong.

2- Measure the V_{CE} of the transistors:

If they are all of the same type, and all turned ON, they should approximately have the same collector to emitter voltage. Finding the transistor having a different V_{CE} is a quick way of detecting a dead transistor.

3- Measure the voltage across the base resistance:

The voltage across a base resistance should be relatively important (if the transistor is switched ON). For a TTL (5V) device controlling a NPN transistor, the voltage drop across the resistor should be more than 3V. If there no remarkable voltage drop across the base resistor, then either the transistor is dead, either the device controlling the transistor has a problem. In both case, the base current is equal to 0.

Source: http://www.ikalogic.com/bjt-switches/