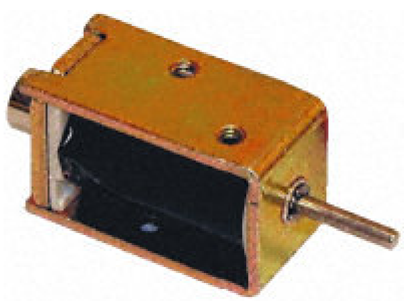
On Using Solenoids in the Derbot

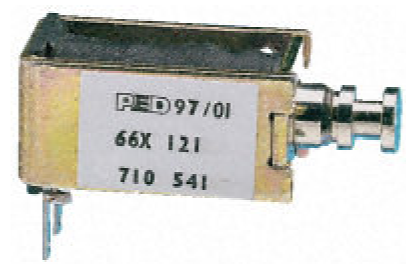
tjw rev. 10.12.09

Introduction

Solenoids are a useful way of achieving linear motion (i.e. movement in a straight line), and it is therefore attractive to use them. On the down side they can be heavy, electrically power-hungry, have limited travel, and exert an impulsive (jerky) force which varies with position of the core. It is therefore well worth understanding how to optimise their application. For the Derbot Challenge two solenoids are on the preferred parts list, as illustrated, though others can be used.



Push Solenoid; RS Part 533-2217

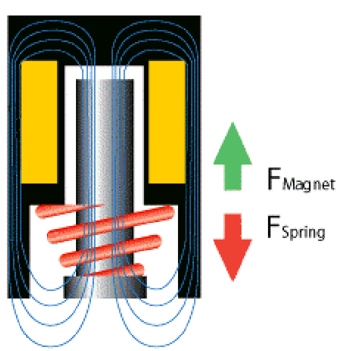


Pull Solenoid; RS Part 250-0653

How They Work

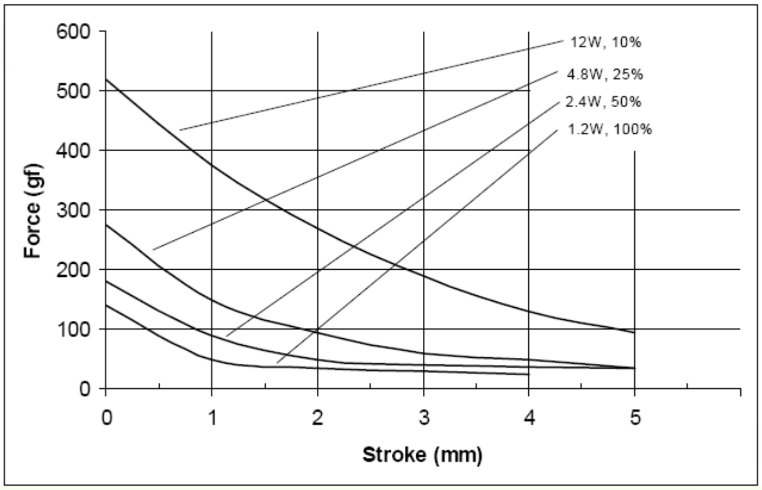
The normal solenoid consists of a moveable soft iron core placed within a coil. When the coil is energised the core is pulled into the centre of the coil. It usually then "bottoms", i.e. makes physical contact with the base of the iron enclosure, and a magnetic circuit is at least partially closed. In this position the force it is able to exert is a maximum.

If the core is a permanent magnet, then attraction and repulsion can be achieved, by reversing the current direction. More often, the core is soft iron and spring loaded. When electrically energised, the core is drawn into the coil. Note that with a soft iron core, direction of current flow does not matter. When not energised, the spring moves the core out, as shown below.



http://www.societyofrobots.com/actuators_solenoids.shtml

A common problem with using solenoids is the variable force they exert, as seen below. When the core is partially out of the coil, the force exerted is comparatively feeble, increasing as the core moves into the coil.



BLP Components Series 134 Data Sheet. www.blpcomp.com

Electrically Driving a Solenoid

The solenoid is a classic inductive load, and hence obeys the equation

$$V = Ldi/dt$$

This equation is primarily of interest when voltage is being applied or switched off. The steady state current is simply I = V/R, where R is the resistance of the coil and anything else in series with it.

In practical terms, this means:

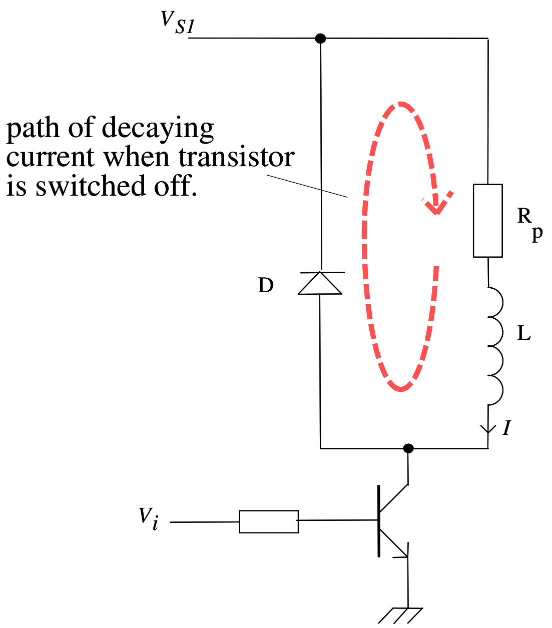
- the current cannot rise instantaneously to its steady-state value, but does so with a time constant L/R , where R is the circuit resistance;
- the current cannot stop instantaneously (which would cause an infinite di/dt term in the above equation), unless you happen to have an infinite drive voltage. There is energy stored in the magnetic field, which must be returned to the circuit (Faraday's Law kicks in);
- the steady-state current, assuming a low coil resistance, might be unnecessarily high.

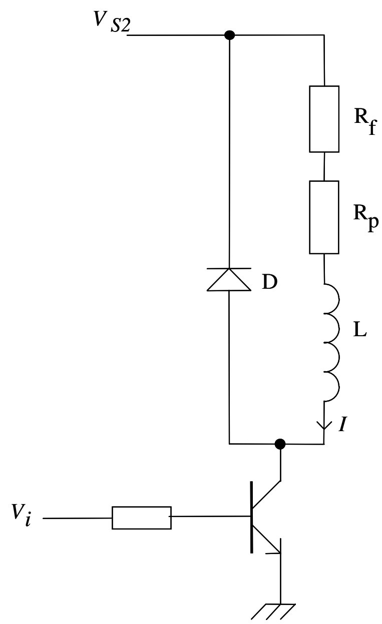
When designing the drive circuit, the common requirements are:

- trying to get adequate current into the coil quickly, fighting against the rise-time issue, and noting that the mechanical pull is likely to be at its weakest;
- limiting the steady state current to something sensible;
- ensuring that the current can decay properly on switch-off.

Drive Circuits

The simplest drive circuit is as shown below left. The transistor (usually a logic compatible MOSFET like the ZVN4206A) is driven from a microcontroller port bit output. When the transistor conducts, the solenoid, indicated by inductance in series with resistance, conducts. Steady state current is V_{S1}/R_P. The “freewheeling” diode, connected as shown, provides a path for the decaying current. Leave it out and your switching transistor ***will*** blow.

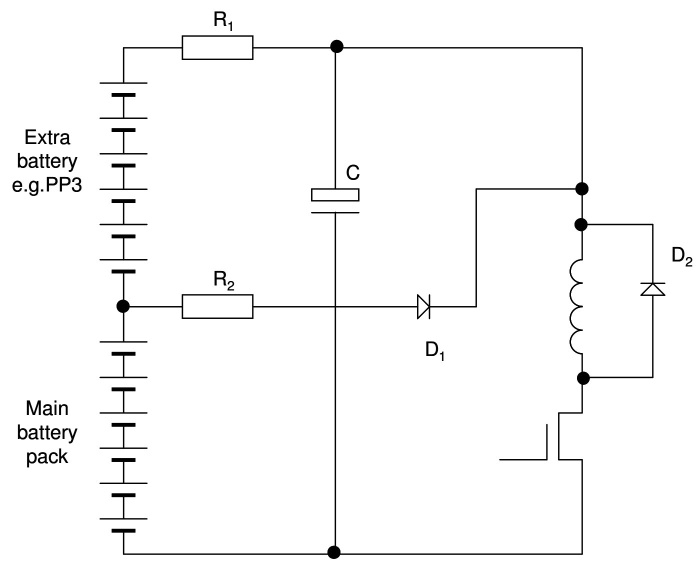




A reduction in time constant is provided by the circuit above right. Scale up the supply voltage and introduce the external series resistance Rf. For example if the supply voltage is doubled, and $R_f = R_p$, then the time constant is halved, with the same steady state current. This is however wasteful of energy, and a lot of battery energy is spent on heating the resistors.

Many dual-voltage solenoid drive circuits have been developed. The higher voltage is used to activate the solenoid rapidly, while the lower maintains the steady-state current. An example, suitable for the Derbot and intended for infrequent solenoid activation, is shown below. A further battery is added in series to the main battery pack. This can be of a much lower energy capacity, e.g. a PP3. The combined battery voltage charges up capacitor C, with time constant R_1C. When the solenoid is activated, the capacitor partially discharges through the solenoid. However when its voltage has fallen to that of the main battery pack, this starts conducting, through D_1 and R_2. The value of R_2 is set to determine the steady-state current, keeping the current as low as is necessary. The value of C is comparatively large, and set to ensure that there is sufficient charge to active the solenoid; 10,000uF is a good starting trial value. The value of R_1 depends on how frequently the solenoid is activated. Clearly C must have time to fully charge before each activation; a useful rule of thumb is that the capacitor is fully charged after 5 time constants. For example, if the solenoid must be activated every 20 s, then the maximum time constant should be 4 s. Hence, noting C = 10,000uF = 0.01F

$$R_1 = 4/10^{-2} = 400\text{ohms}.$$



Alternatives to Solenoids

Note that it is easy to configure a servo to exert a near linear force. This has the advantage of being constant, with a controllable speed, and possibly longer displacement. Of course servos come with their own drive challenges, and can also be bulky and a bit heavy. See Section 8.8 of “Designing Embedded Systems with PIC Microcontrollers”, either edition, for further information.