PWM EXPLAINED

Pulse Width Modulation is a phrase you'll be hearing more and more. This important control technique looks set to revolutionise digital electronics.

LET US delve into the realms of digital control. The particular technique involved is pulse width modulation. This technique is likely to become more and more common in audio amplification and power supplies, especially now that high speed switching devices are becoming widely available.

To understand the principle, look at the simple circuit of Fig. ld. Here a transistor, used in the common emitter mode, has a square wave of equal mark space ratio as an input. The transistor is in saturation or cut off for equal periods, so the average voltage at the collector, as measured with a multimeter will be half the supply voltage (Fig la). If the mark to space ratio is increased as in Fig. lb then the average voltage will rise. Conversly if the mark to space ratio is decreased (Fig. 1c) then the output voltage will fall. Taken to extremes the transistor would be either in saturation or cut off for the whole time and the output voltage would either be zero or supply voltage.

Well, you say, so what? The simple answer is that, unlike an analogue design, a current can be delivered to a load with hardly any power loss or dissipation in the driving device. For instance, look back to the circuit of Fig. 1 and assume that we have a supply voltage of 9 V and a collector resistance Rc of 100R. To maintain a voltage of 4.5V across this resistor there must be a current flow of 45 mA. The power dissipated in the transistor and the resistor is found by multiplying the voltage drop by the current flowing. In this case 4.5V x 4.5 x 10^{-2} A =0.2025 W. In the switching circuit a 1:1 mark space ratio square wave would be used to set the required voltage across the resistor. Ideally the voltage drop across the transistor would be zero when it was in saturation, and the current through Rc is then 9 V I, i.e. OV x 90 mA, zero! When the transistor is cut off the full supply voltage would appear across it and the power dissipation would again be equal to V x I, 9 V x 0 mA, again zero! In reality there will always be a small saturation voltage across the transistor of a few hundred millivolts, and even when the transistor is cut off there will still be a small leakage current flowing, although this will only be in the order of a few microamps.

Although transistors are imperfect devices, you can see that the square wave circuit is many times more efficient than an analogue one.

Square Waves

Before the idea can be used practically, a means must be found of generating a square wave with an easily adjustable mark space ratio. A simple method of doing this is to feed a triangle wave of known amplitude into one input of a comparator, and a control voltage into the other. For those unfamiliar with comparators, a quick description is probably in order. A comparator has two inputs, an inverting and a non-inverting, like an op amp. It also has a high voltage gain but unlike an op amp is operated without negative feedback. It functions as a switch — the output is at zero potential when the non-inverting input is more negative than the ihverting while the output is fully positive when the non-inverting input is more positive than the inverting one. Because the gain of the comparator is very high a voltage difference of a few millivolts at the inputs will be sufficient to ensure switching. Comparators, as the name implies, are used for detecting and comparing voltage levels. If one input is fed with a triangle wave and the other with a variable voltage level the output consists of a square wave whose mark space ratio depends upon the voltage at the input. Figure 2 should make this clear.

Practicalities

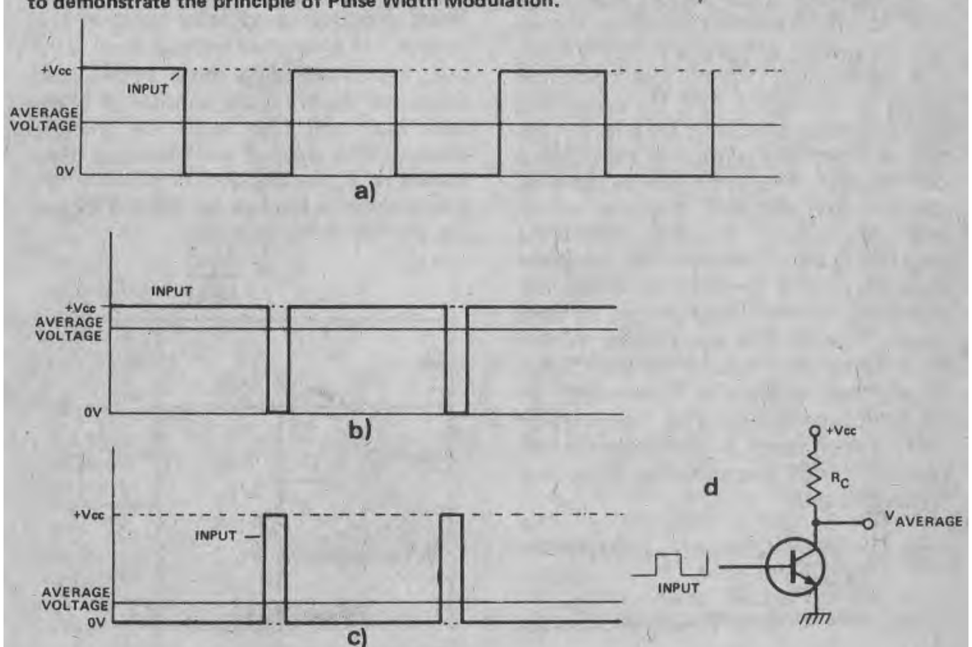
The practical application is shown in Fig. 3, a pulse width modulator that will deliver in excess of one amp to a load (such as a motor or lamp) that can be simply adjusted by a control voltage.

The circuit is an excellent power saving device with an efficiency of 90% and is ideally suited to battery operated equipment.

To keep the circuit simple while providing maximum flexibility, a low cost quad op amp is used as the active device. The type chosen, the LM324, contains four op amps that are similar to the good old 741 but with the advantage that the output can go down to ground even when operated from a single supply voltage.

IC1 a is used as an astable multivibrator that produces a 1:1 mark space spuare wave at its output. These are input to an integrator, which converts the square wave into a linear triangle wave suitable for the pulse width modulator, IClc . ;Cie is used as a comparator with the triangle wave at one input and the control voltage at the other. The resulting square waves at the output of IClc are used to drive the output transistor, a TIP41A which is used in the common emitter mode.

Fig. 1. (a) 1:1 mark-space ratio. (b). Increasing the mark-space ratio increases the average voltage. (c) Reducing the mark-space ratio reduces the average Voltage. (d) Using an NPN transistor to demonstrate the principle of Pulse Width Modulation.



Having outlined the circuit we can consider its operation in more detail.

How It Works

The non-inverting input of IC1a is connected to the junction of R1, R2 and R3. !Cie is being used as a comparator and its output must either be high or low.

When power is first applied C1 is discharged. Thus the voltage at the inverting input is held lower than that of the non-inverting input, and the output of the op amp is at supply voltage. C1 starts to charge up via R4 and when the voltage at the inverting input exceeds that at non-inverting input the output of the op amp goes down to 0 V. Now C1 discharges through R4 and the op amp's output stage.

The non-inverting output is held at a potential that depends on the values of R1, 2 and 3. When the output is high R1 and R3 are effectively in parallel while when the output is low R2 and R3 are in parallel. Since R3 is connected from the output to the non-inverting input a positive feedback loop is obtained. In practice this ensures that the output of the op amp changes from high to low state and vice versa very rapidly. If all three are made equal in value then the potential at the non-inverting input will oscillate between 1/3rd and 2/3rds of the supply voltage.

The frequency at which the circuit runs is determined by the values of R4 and Cl and **can** be calculated from the formula

$$f = \frac{1}{1 \cdot 4 R4 C1}$$

To operate small motors and lamps from the pulse width modulator the actual frequency employed is not critical. The lower limit for reliable operation seems to be 100Hz. At the upper end this particular circuit is limited by the **rate** at which the output of the op amp **can** change. This is known as the slew **rate and you will find** it quoted on op amp data sheets. For the LM324 the slew rate is 0.5 V/microsecond. Another measure of this same effect is the full power bandwidth, also quoted in data sheets. For the 324 this is 6k Hz. Within these limits the values of R4 and C1 can **be whatever happens to be at hand. In the prototype** an operating frequency of 1kHz was chosen. A 10n capacitor was **available for Cl and** so the equation was rearranged thus

$$R4 = \frac{1}{1 \cdot 4fC1} = 1 \div \left(1 \cdot 4 \times 10^3 \times 10^{-8}\right)$$

68k is the nearest value.

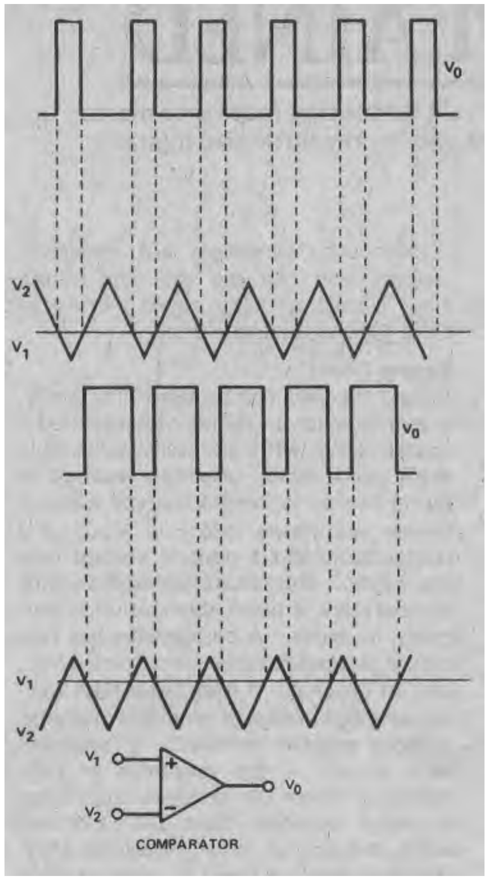


Fig. 2. Using a comparator to produce square wave from a triangle wave and a variable DC level.

Our square wave must now be converted into a triangle wave by a circuit known as an integrator. The output of this circuit is directly proportional to the integral of the input. The integral of a square wave is a linear triangle wave. A positive voltage level at the integrator's input produces a negative ramp at the output. If a negative voltage level is input, a positive-going ramp comes out. Since our square wave consists of alternate high and low levels the triangle -wave at the output will resemble that shown as V_2 in Fig. 2. To calculate the values that we require for R5 and C2 we use the following formula,

$$V_0 = \frac{Vcc}{4RCf}$$

Again a large range of values can be employed and by simply rearranging the equation one can easily calculate the required values. An example of the procedure follows. To get a large adjustment range a fairly large signal swing is required. Assuming that the circuit will be employed over a range of supply voltages it is necessary to ensure that an undistorted and unclipped signal is available. For this reason a peak-to-peak voltage of 7 V was chosen and the required component values were calculated as follows. Choose an arbitrary value for C2 say 100n. The value of R5 can be calculated by rearranging the above equation thus:

$$R = \frac{Vcc}{4V_0Cf}$$

$$= \frac{9}{4 \times 7 \times 10^{-7} \times 10^3} = 3 \cdot 2 \times 10^5$$

The nearest value is 330k.

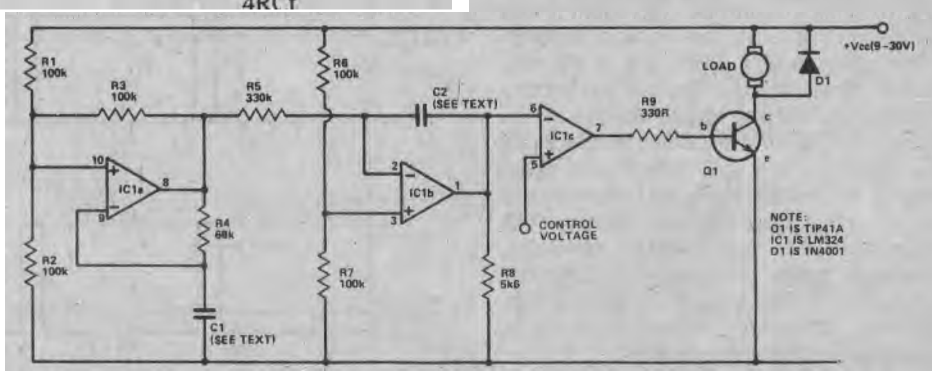
Note that 9 V was taken for the supply voltage. This calculation does not need to be repeated for other voltages as the amplitude of the triangle wave will be in direct proportion to the supply voltage used.

The last part of the circuit is built around IC1c. This is the pulse width modulator proper and its function has already been described. The output square wave drives the power transistor, a TIP41A. Base current is limited by R9.

D1 is included to protect Q1. When a current is fed into an inductive load, for instance, a motor winding, energy is stored in the magnetic field that builds up around the wire. When the supply is interrupted the field collapses and as it does so the magnetic flux produces a large reverse polarity voltage spike.

This spike can have sufficient amplitude to destroy the driving transistor. The diode will short any such spikes to the positive rail thus protecting the transistor. If you glance through any projects in ETI where tran-

Fig. 3. (Below) A practical circuit using Pulse Width Modulation to control a motor.



sistors are used to drive relays, for example, you will notice this feature has been included.

Layout

A simple Veroboard layout for the circuit is given in Fig. 4. The **TIP41A** does not require a heatsink. As it stands the circuit has many applications. A typical application would be a model train speed controller.

Audio Applications

PWM amps are not new. Sinclair marketed a 10 W version in the sixties. What has precipitated the recent interest is the advent of digital recording techniques, and the increasing avail-

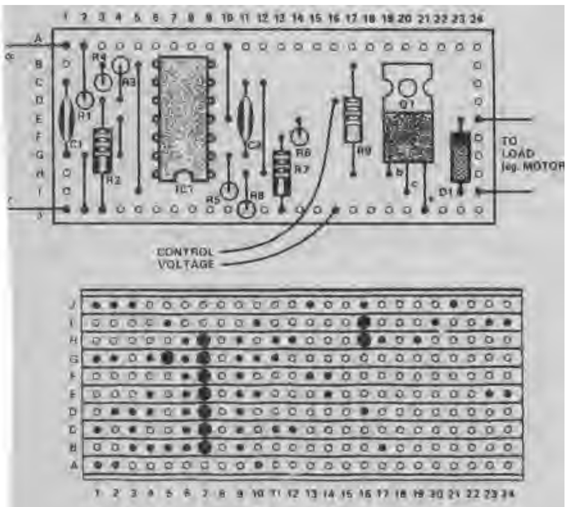


Fig. 4. Veroboard layout for the circuit in Figure 3

ability of fast switching transistors, especially the VFET.

To recap, a pulse width modulator is basically a square wave oscillator whose output markspace ratio can be altered by an external voltage. If this voltage is an audio signal and the output frequency is sufficiently high, the average output voltage of the oscillator will be the audio signal. The obvious question is how high does the frequency have to be to encode the full adulo band from 20 Hz to 20 kHz? Surprisingly, thanks to the work of Nyquist, the answer is only 40 kHz. Nyquist showed that a signal, in the form of modulation oil a carrier wave, could be fully recovered as long as the carrier was at least twice the maximum frequency of the modulating signal.

Given that information it should be possible to produce a PWM amp with a beefy output stage running at well over 40 kHz that can deliver a signal to a normal loudspeaker. Such a digital amplifier has lots of advantages in terms of performance. If a push-pull output stage is employed, the power devices dissipate very little heat since they are either in saturation or in cut-off.

As long as the signal is linearly transferred into digital form no audio distortion can occur. Similarly, there is no crossover distortion or noise added to the signal. What has prevented all audio amplifiers from being built this way is the problem caused by the electromagnetic radiation at carrier frequency from the speaker leads. A second problem has been the shortage of devices.

When Sinclair marketed his PWM amp in the sixties, silicon transistors and logic ICs were expensive novelties. Nowadays there is a proliferation of ICs and transistors capable of being employed in such circuits.

The problem of carrier radiation is a vexed one. Most domestic appliances produce RF radiation, as can be confirmed by anyone whose hi-fi equipment is upset by switching 'thumps'. The simplest way over this problem is to place a low-pass filter between the amp's output and the load. Using screened lead with the screen connected to ground is another precaution. With the amp described here RF radiation is not a real problem as long as the filter is incorporated.

The last circuit employed a comparator to produce PWM. This time the same result is obtained by using the audio signal to alter the switching thresholds of a Schmitt trigger. This trigger is a device that has two switching thresholds, let's call them t1 and t2 **(Fig. 5).** When the input voltage is less than t1 the output is low. As soon as the voltage exceeds tl it goes high.

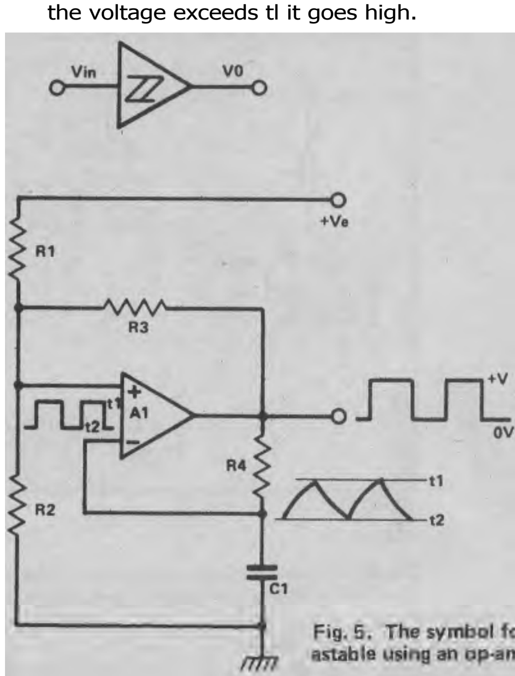


Fig. 5. The symbol for a Schmitt trigger is shown above. Below is an astable using an op-amp operating in a Schmitt trigger action.

Low Down **Volts**

If the voltage is now reduced, nothing happens until it falls below t2. At this point the output goes low again. (Note that the two threshold voltages tl and t2 are not equal.) This characteristic is known as hysteresis. The ZZ indicates **a** Schmitt device or function and these triggers are usually employed in digital circuits to convert slowly rising and falling waveforms to pulse trains suitable for logic systems. A Schmitt trigger can be made from an op-amp or comparator. **Figure 5** also shows an astable similar to one employed last month but this time it is built around an op-amp.

On switch-on, capacitor Cl is discharged, holding the inverting input of Al low relative to the non-inverting input. Since this is so, the output will be high. Now, if we make **R1,** R2 and R3 equal in value, then the non-inverting input will be held at 2/3 Vcc, because R3 is effectively in parallel with R 1 . In consequence, Cl will rapidly charge through R4 until the voltage at the inverting input exceeds 2/3 Vcc. At this point Al 's output will start going negative. Positive feedback through R3 makes the output's transition from high to low extremely rapid.

A second stable voltage will now be found at the non-inverting input, equal to 1/3 Vcc. This time, R3 is effectively in parallel with R2. Capacitor Cl will now discharge via R4 and Al's output stage until the voltage on the inverting input falls below 1/3 Vcc. The output again goes high and the cycle repeats itself indefinitely.

Because of the influence of R3, **Al** acts as a Schmitt trigger, tl and t2 being 2/3 Vcc and **1/3** Vcc.

Figure 6 shows how this simple circuit can be modified to encode an audio signal into PWM. The audio is simply imposed upon the non-inverting input, thus altering the threshold switching voltages. Resistor R5 prevents interaction between the audio signal and carrier as does R6. Capacitor C2 isolates DC voltages from the astable. Unfortunately, Cl charges exponentially via R4 so somewhat less hysteresis is applied by making R3 much larger than either **R1** or R2. This has the effect of making the switching thresholds very close together and I i nearising the triangular waveform across Cl resulting from it being charged through R4.

To produce a good square wave at high frequencies an LF351 op-amp is used. This is JFET device which features a high slewing rate, that is 13V/uS.

rated at 1 A collector current. As you can see they are connected, without base bias, as emitter-followers. When the output of the op-amp is high, Q2 is in saturation and provides current to the load via the low-pass filter L1, C5 and

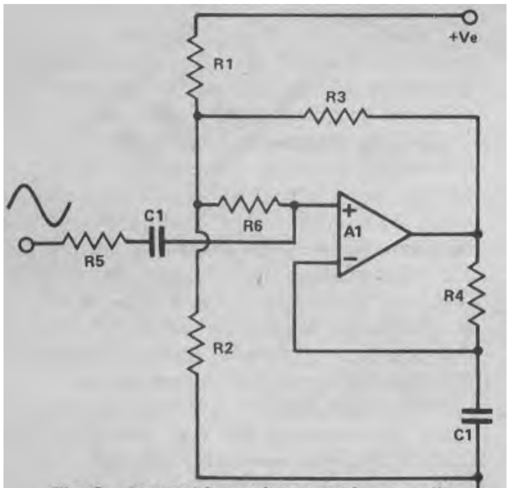


Fig. 6. An astable used to encode an audio signal applied at its input into PWM.

the output coupling capacitor C6. Incidentally, in class B amps the latter component is often of a lower value. This is a pity since the lower — 3 dB point is defined by the size of this capacitor.

As you will remember, the impedance Z of a capacitor is given by:

$$Z = \frac{1}{2\pi FC}$$

where C is in farads and Z is in ohms.

Negative Feedback

The astable has a unity gain and so, to improve the sensitivity, an audio pre-amplifier stage has been added. This is built around Q1 which is used in the commmon-emitter mode. Negative feedback, however, is applied from the collector via R3. A collector current of 1 mA has been chosen to allow adequate drive. Resistor R7, therefore, drops 1 V while C3 decouples line ripple to ground. For linear drive the collector is operated at 17 V/2 8 V. The value of R4, therefore, is given by 8 V/10^3 A=8k, the nearest value being 8k2.

An MPSA16 is used for **Q1** and this has a gain of 200 minimum at 1 mA. Base current IB is therefore equal to:

$$\frac{10^{-3}\,A}{2\times10^{2}} = 5uA$$

12k being the nearest value.

Similarly, the value of R2 will be equal to the collector voltage minus the base voltage divided by 50uA, given by:

$$R2 = \frac{8 - 0.65}{5\times10^{-5}A}$$

$$= 147k,$$

150k being the nearest value.

Virtual Ground

Using a transistor in this way, with feedback from collector to base forms a 'virtual ground amplifier'. This is because the feedback reduces the input impedance. The gain of the stage is set by the ratio of R3 to R1. For a gain of 10, sensitivity 900 mB, a value of 15k was chosen. Capacitor C1 simply isolates the input from DC from previous stages. A suitable Veroboard layout is shown in Fig. 8. In the circuit shown in Fig. 7. we have a classic high-pass filter formed by the speaker impedance Z and C. The lower — 3dB point can be calculated by rearranging the equation for f1 given by:

$$\frac{1}{2\pi CZ}$$

substituting:

$$f1 = \frac{1}{2\times3.14\times2.2\times10^{-3}\times8} = 9.05\text{ Hz}$$

follows that the worst-case minimum current that can be fed into the speaker is:

$$\pm 25\text{ mA}\times 30 = 750\text{ mA}.$$

Since the speaker impedance is 8R, the peak voltage under these conditions is equal to 1 R; that is, 750 mA x 8R = 6 V, or 12 V peak-to-peak. Since the power output is given by V^2/R, one could be forgiven for thinking that the output power would be (12 x 12)/8 = 18W. Unfortunately, you would be wrong!

The output power is the RMS voltage divided by the load. Assuming our output is a sinewave of 12 V peak-to-peak, the RMS value is found by dividing Vpk-pk by 2.8.

The output power (minimum) is therefore given by:

$$\left(\frac{12}{2.8}\right)^2/8R = 2.29\text{W RMS}.$$

Because of the small voltage drop that occurs across a saturated transistor, the output will be slightly less than this, namely 2 W.

Going back to our astable, the operating frequency has been set at 300 kHz. This gives a full power bandwidth of 10 Hz to 150 kHz.

Build It Yourself

The construction is quite straightforward and requires little comment ex-

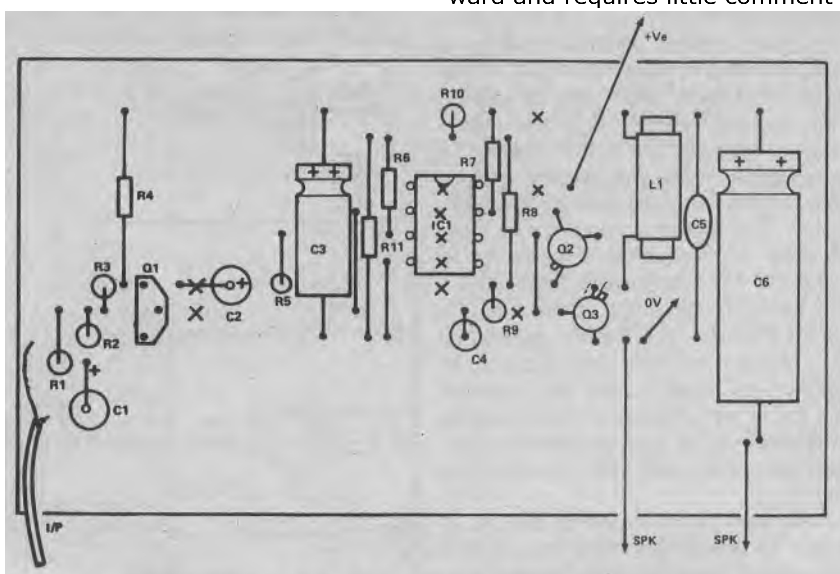


Fig. 8. The veroboard layout of the PWM amp. The small crosses (x) indicate breaks in the copper track underneath the board.

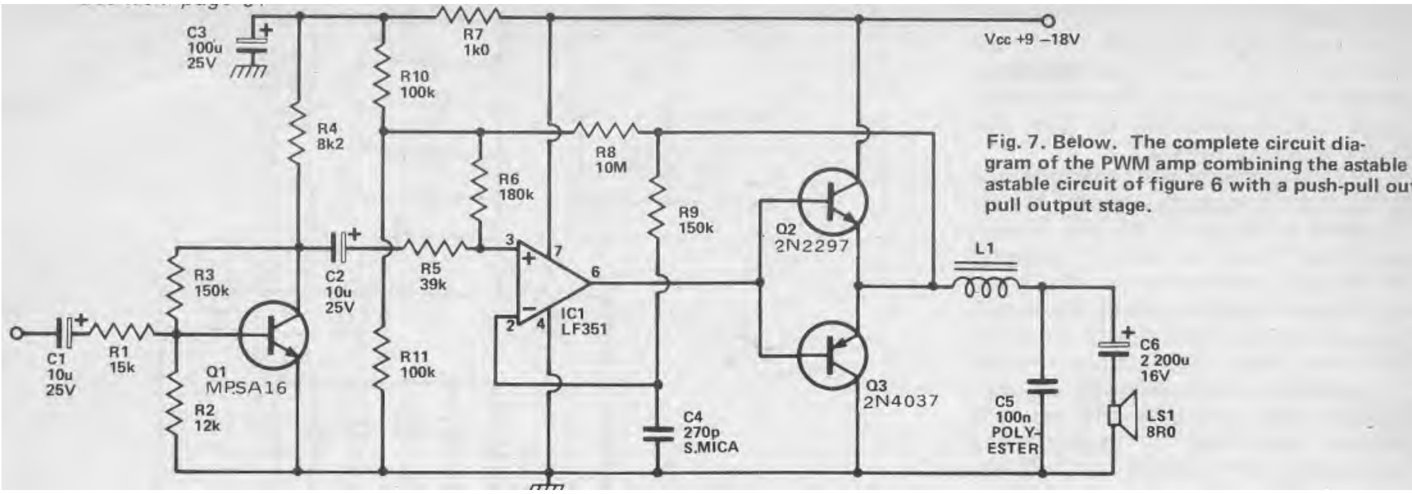


Fig. 7. Below. The complete circuit diagram of the PWM amp combining the astable astable circuit of figure 6 with a push-pull out pull output stage.

cept that it is necessary to ensure that all the semiconductors are correctly orientated and the breaks in the veroboard tracks are not forgotten.

Although the circuit will operate from 9 V, batteries are not really suitable and any line PSU offering an output voltage in the range indicated is better.

L1 consists of 60 turns of 0.56mm enamelled copper wire, pile-wound on a 1 in length of 3/8 in diameter ferrite rod.

When construction is completed, no adjustments need to be made to the circuit. All that is required is an input signal and a speaker.