SINGAPOREAN
SECTION

IRSE - SINGAPOREAN SECTION

How maximum is a "Maximum Safe Speed"

Presented by: David Mott, BA, LLB(Hons), MBA, MSc(Eng), C Eng, FIRSE, MIET, MAPM, CMILT

Presented on 16th November 2007

1 Introduction

During my career in railway signalling there have been significant developments not only in the technology itself, but also in the tools that are available to apply that technology.

We are all familiar with the way that electro-mechanical control systems have given way to processor-based systems. We have all witnessed the transformation that has occurred in Communications Technology –and we have all seen the amazing developments in Semiconductor Technology.

At least as important as these developments themselves, and their application to railway signalling systems, has been the development in the tools available to the engineers who design and develop these applications. These tools, of course, have primarily resulted from the developments in Computer Systems.

It is now possible to carry out design tasks and design simulations very quickly, and to a very high degree of accuracy, by using computer based tools that are very often specifically designed for their particular purpose. In fact, by using computer based tools, it is actually easier to carry out calculations to a high degree of accuracy than it is to carry them out more generally. This is all very well provided that the result of the calculations is interpreted in cognisance of the accuracy of the input data –and of the appropriate accuracy of the results.

Figure 1: Modern Design Tools

I will, later in this paper, give some examples where results have caused practical difficulties when they have been quoted to an inappropriate level of accuracy. Of more concern in the context of this paper is the effect of such values when they are relevant outside the specialist discipline in which they have been calculated. Electronic engineers, mechanical engineers, and civil engineers are all trained to deal with tolerances appropriate to their discipline. Inappropriately applying figures between disciplines can result in difficulties and in sub-optimum performance for the overall railway system.

I have taken, as an example of this distortion, the calculation and interpretation of the maximum safe speed used on an automatically signalled metro system.

The paper starts by explaining examples of inappropriate specifications of detail that have been experienced in practice. It continues by comparing methods used to calculate block lengths and maximum safe speeds before computer based tools were available with those with those methods available today. After that I look at the factors that civil engineers and rolling stock engineers use to define the safe speed for a railway –and specifically the accuracy to which they are able to specify a safe speed. I compare this with the accuracy associated with the measurement and calculation of speeds by an ATP system. I go on to compare how track coded based (FS 2000) and CBTC (TBS) systems deal with the concept of Maximum Safe Speed (MSS), and explain the loss in overall performance that can result from an inappropriate interpretation of the concept of MSS.

2 Inappropriate Levels of Accuracy

Computers can calculate dimensions to extremely high levels of accuracy - usually much more accurately than it is possible to measure the input parameters. Computers are certainly capable of calculating dimensions to much higher levels of accuracy than it is usually possible to install them. Difficulties can also be encountered when dimensions are exchanged across engineering disciplines.

I recall an anecdotal example that occurred during of the construction of one of the stations on phase one of the Singapore metro. Difficulties had arisen between the supplier of the escalators and the civil engineer constructing the station structure. The escalator supplier had quoted by the size of the opening that he required for his escalator. He was clearly used to measuring dimensions in the context of fabricated steel assemblies. The civil engineer, however, was used to measuring dimensions in the context of poured concrete structures. Both parties considered that the size of the opening was correct. Unfortunately the escalator was slightly wider than the opening.

Difficulties can also occur, obviously, within the same discipline. The computer simulations used to determine the block lengths required for Singapore phase 1 calculated the chainage of block positions to the nearest millimetre. The drawings prepared by the design engineers reflected this degree of accuracy. The design engineers also produced a drawing showing how to attach the appropriate block termination equipment to a sleeper.

Figure 2: IRJ Block Termination on Sleepers

The sleepers on the railway had a nominal spacing of 860 millimetres. Clearly a stated tolerance of, say +/- 430mm on the installed location of the block termination equipment was not only reasonable but a pragmatic alternative to moving the sleepers to put them in the right place. It is incredible how much debate this apparently simple problem caused in the design department.

Figure 3: WRSL Design Office circa 1970

Perhaps it is relevant to remind ourselves how we managed to design railways before everybody had personal computers. In about 1973 I was given the task of working out where to put the block joints on the new, track code based, speed signalled, ATP/ATO equipped Madrid Line 7. Nobody in the company had ever done it before – but I had just passed my IRSE examination so I was considered to be the expert.

After some initial research into the equations of motion and traction performance we established some equations that we could use to predict the movement of a train. It soon became obvious that to apply these equations along the length of the line would take several years. Fortunately I was able to make use of the company mainframe computer. Even that took several runs over consecutive nights to produce a numerical printout that could subsequently be used to manually plot the speed distance curves for the line. These curves were then manually converted to time distance curves.

All the braking distances, deceleration distances, maximum speeds and response times were manually applied –using a slide rule and a drawing board –to establish the optimum block joint locations. The slide rule calculations were probably accurate to three figures. The graphical work, on the drawing board, was, I admit, subject to a degree of parallax error. The results themselves were probably accurate to a few percent –say about a metre in the length of most tracks circuits. Bearing in mind the accuracy of the input data –particularly relating to traction and braking performance –this was probably an appropriate degree of accuracy. Certainly the line has worked successfully for more than 30 years without any problems caused by track circuits being too short.

The manual work involved in this process was laborious and time consuming and took a few months. However, at every stage of the process the results were continuously being evaluated and challenged so that there was a highly degree of understanding of the process itself and of any assumptions made.

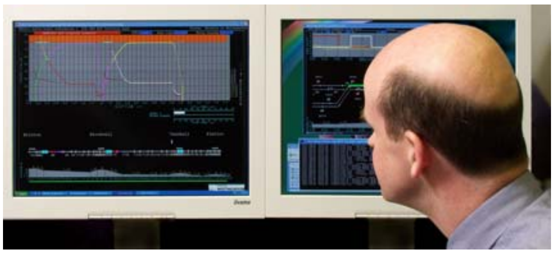

Figure 4: WRSL Multi-Train Simulator

The same basic process is carried out today. The same algorithms are used. However all the calculations are carried out on a PC. The process (apart from inputting the necessary data) is entirely automatic and takes very much less time. Unfortunately, errors in the input data are less obvious and more difficult to detect. Nevertheless, the calculations are carried out with extreme accuracy and the results are often quoted to the nearest 10th of a second or to the nearest millimetre. It is my contention that this implies a degree of precision that is completely inappropriate taking into account the accuracy of the input data and of the practicalities of installing the subsequent system.

The engineers who carry out the calculations exist in a culture where such precision is assumed to exist. I believe that it is worth questioning whether in fact this precision applies to some of the input data that they use. As this paper uses the concept of maximum safe speed as its prime example of this general problem it is necessary to consider how engineers outside the signalling discipline might regard the idea.

3 Factors that Govern the Safe Speed for a Railway

Fi

Figure 5: Safe Trains?

It used to be said that the only is safe speed for a train was zero. Hence, failsafe circuits caused a train to stop if the situation was not safe. Even this simple statement is, of course, questionable. It serves, however, to remind us that safety is not a binary condition. It is erroneous to think that something can either be safe or unsafe. There is a whole range of conditions from "very safe", through "quite safe", "quite dangerous", and "very dangerous" to "certain catastrophe".

As far as a safe speed for a train is concerned (and ignoring complications like braking distances and deceleration rates) there are three generally recognized conditions. One relates to the danger of derailment caused by excessive speed, another relates to the danger of discomfort caused by excessive acceleration forces acting on passengers, and the third relates to excessive wear on either the track or the wheelset.

Let's look at the factors relating to the danger of derailment. The derailment of a vehicle can occur as a result of (Profillidis, 1995):

- The track shifting;
- The wheel rebounding on the rails; or
- The vehicle capsizing.

Consider derailment caused by track shifting:

"under the influence of considerable transverse loading, the track shifts as a whole and causes derailment of the vehicle. This form of derailment mainly occurs at high speeds. The condition for derailment by track shifting is that the transverse force, H, which may cause track shifting, exceeds the transverse resistance, L" (Profillidis, 1995)

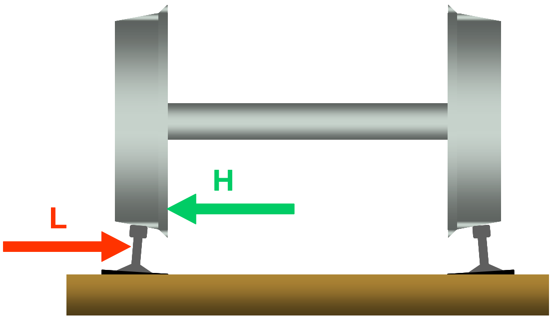

Figure 6: Transverse Loading

Now, H is the sum of the static and dynamic transverse forces (H_s and H_d).

Now, H_s is calculated from the "semi empirical" formula: (Amans, 1969)

$H_s = \frac{P \times NT(mm)}{1},500$, where P is the per-axle load and NT is the "transverse defect", which is effectively the (actual) cant deficiency.

Similarly, H_d is calculated from another "semi empirical" formula: (Amans, 1969)

$$H_d = \frac{P \times V\left(\frac{km}{h}\right)}{1{,}000}$$

, where *P* is the per-axle load and *V* is the velocity.

Also, L is calculated using either the equation (ORE, 1980):

$$L = 0.85\left(1 + \frac{P}{3}\right), \quad L = \left(1 + \frac{P}{3}\right),$$

or $L = \left(1.5 + \frac{P}{3}\right)$, depending on whether the track is timber sleepers with manual track maintenance, timber with mechanical track maintenance, or twin block concrete sleepers.

The use of "semi empirical" formulae that include factors such as "1,500" and "1,000", rather than, say, "1,387" or "1,136" give a reasonable impression of the precision of such calculations. The results are very important, but they are not particularly precise.

Now, consider derailment caused by the wheel rebounding on the rails:

"When the transverse force developed between wheel and rail exceeds a certain value, then the wheel rebounds on the rail (-climbs over) *and causes derailment. This form of derailment mainly occurs at low speeds and the condition that must be satisfied is given by Nadal's formula"*

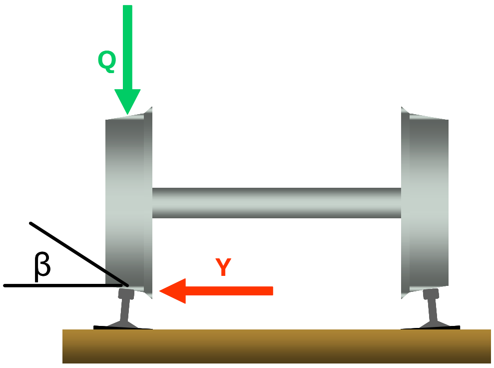

Figure 7: Nadal's Formuls

"Nadal's formula" (Nagase, 2002) states that:

$$Y = Q\frac{\tan\beta - f}{1 + f.\tan\beta}$$

, where Y is the transverse force, Q is the wheel load (static and dynamic) and β is the angle between the horizontal and the wheel/rail contact plane, and f is the wheel-rail friction coefficient. Of course, the dynamic wheel load can be easily +/- 50% of the static load.

Fortuately this theoretical formula can be empirically simplified so that, for vehicles on bogies, $\frac{Y}{Q} = 1.3$. (Amans, 1969).

As in the previous case, the inclusion of factors such as "1.3" give a reasonable impression of the precision of such calculations. The results are very important, but they are not particularly precise.

Finally, consider derailment caused by the vehicle capsizing:

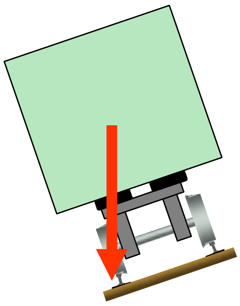

Figure 8: Vehicle Capsizing

According to Profillidis :

"In (the case of derailment caused by overturning of the vehicle) *the vehicle capsizes due to overall unstable equilibrium. It was found that in the worst case (with the centre of mass 2.25m from the track) for standard-gauge track, a vehicle will overturn when the transverse acceleration reaches g/3. … Tracks are laid out for a maximum value of non-compensated centrifugal acceleration ranging between 0.5 and 1.0 m/s*^{2} *and never exceeding a maximum value of g/10. Therefore the safety factor against derailment by overturning is (g/3)/(g/10) >3."*

It is clear that the factors governing derailment are incapable of precise calculation and rely on empirical factors. It is also clear that, for two of the mechanisms at least, the wheel loading is one of the major variables. The dynamic wheel loading at any point depends, among other things, on the design of the vehicle suspension, and the state of wear of the wheel tyres (conicity), and on the speed of the vehicle.

What does all this mean in practice? It means that, provided that the signalling system can take care of braking distances to prevent collisions, trains generally become increasingly likely to derail as speeds increase but it is difficult to predict the precise speed at which derailment will occur. However, there are very large safety factors that are applied. In practice metro trains are most unlikely to derail because of excessive speed – unless there is some serious failure of the train suspension or the track support system. The exception to this is probably derailment due to the wheel climbing over the rail on tight curves.

Apart from actual derailment there remain the considerations related to passenger comfort and rail/wheel wear. The forces experienced by the passengers increase as the train speed increases. It is usual to limit both accelerating forces and jerk for reasons of both passenger comfort and ultimately passenger safety. If the train speed approaches the maximum design speed the suspension system comes closer to the critical speed at which the damping will become ineffective and wheelset and bogie oscillation will begin. This is not only uncomfortable but also dangerous because of the increase dynamic loads.

It is possible to conclude that trains get more uncomfortable, and cause more wear, as speeds increase before they become dangerous enough to derail. This is even true in the low speed case of the wheel mounting the rail on sharp curves – certainly the flange contact noise and the rebounding forces will very apparent before the wheel actually climbs over the rail.

In summary, for any stretch of line there will be a speed at which the comfort levels seriously decline and the rail wear becomes excessive, and there will be a higher speed at which there is a significant danger of derailment. Both these speeds, which I call the "comfort speed" and the "derailment speed" are difficult to define precisely and depend on a number of variables that are difficult to measure. Speed limits quoted by the civil engineers reflect experience in practice and contain significant safety factors.

Figure 9: Tilting Train - showing the difference between a "comfort" speed and a "derailment" speed

If a civil engineer quotes a civil speed limit of, say, 50 km/h he does so on the basis that such a speed will not impose more than about 0.1g of lateral forces on the passengers, that it will not cause undue wear to the wheel or rail, and that there is a sufficient safety factor to ensure that minor problems with the suspension or track fixings are very unlikely to cause a derailment. In the context of a manually driven railway with no continuous over-speed protection, compliance with this speed limit is entirely the responsibility of the driver.

4 Accuracy Associated with the Measurement and Calculation of speeds by an ATP system.

Modern ATP systems are capable of measuring the rotational speed of an axle to a very high degree of precision. They are even capable of compensating for long term variations in wheel diameter. Use of multiple sensors can reduce any errors introduced by wheel slip or slide. It is usual to measure speed to within a few km/h – or, in some cases, to a fraction of a km/h. The calculations performed using this input data can also be carried out to a very high degree of precision. Thus, it is clear that, once a "maximum speed" is defined, the signalling system can measure that speed – and carry out related calculations – to very high degrees of accuracy.

In terms of ensuring that a train always has an available emergency braking distance ahead of it this precision can be important. It allows braking distances from a particular speed, assuming a particular deceleration, to be calculated with precision. This avoids the necessity to use large safety factors and allows the clear distance ahead of a train to be minimised – thus maximising line capacity.

In terms of ensuring that a train does not exceed the civil speed limit such accuracy is of questionable benefit. As demonstrated earlier the accuracy to which civil speed limits are determined is not particularly precise when compared to the accuracy that the ATP system uses – and they are often enforced solely by the driver.

I do not suggest that the ATP should work less accurately – but I do suggest that it is applied appropriately. There is also a need to dispel the culture in some areas that a train speed is "safe" at 83.5 km/hr, but "unsafe" at 84 km/hr. As explained earlier "safety is not a binary condition".

5 Comparison of Speed Signalling and cbtc

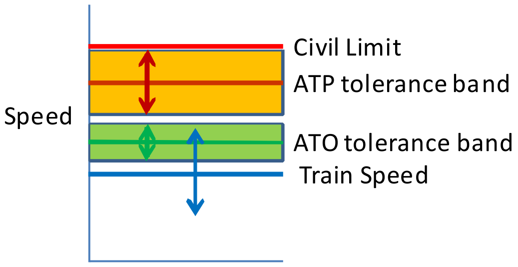

Figure 10: Speed Signalling - Tolerance Bands

The “rules” for a typical speed-signalled system are (simplified) that if the train exceeds the civil speed limit then the emergency brake should be applied – and that it should not suffer from “spurious trips” (i.e. it should not erroneously call for an emergency brake if the speed does not exceed the civil speed limit.

The ATP has a tolerance on its measurement of speed – so the highest tolerance of the ATP speed is placed at or just below the civil speed limit. The lower end of the tolerance is then at a speed lower than this. Typically for a speed of about 80 km/h the ATP tolerance band – assuming 5 steps of wheel diameter correction – will be about 3 km/h.

The ATO, which actually controls the speed of the train, must ensure that the train speed does not exceed the lowest ATP tolerance. Typically the top end of the ATO tolerance band is about 0.5 Km/h below the ATP band and it might be 1 Km/h wide. This is the maximum ATO speed – the average speed is likely to be a few Km/h below that.

Thus, the actual speed at which the train regulates is likely to be 4 or 5 Km/h below the civil speed limit.

This is, arguably, not a very good idea bearing in mind how the civil speed limit is derived. It would probably not reduce the overall safety of the railway if, for example, the ATP nominal speed, or even the ATO nominal speed, was set to equal the civil speed limit. Nevertheless the method described above is standard practice so is unlikely to change.

If the train exceeds the ATP speed (which is lower than the civil speed limit) the ATP will call for an emergency brake application. It may take a second for the ATP to react and it may take another second before the brakes are effective. The train may therefore accelerate for a second or two after exceeding the ATP trip speed.

Even on a level gradient a train (at low speeds) can probably accelerate by at least 1 m/s^2. This would mean that it could actually attain a speed about 3 or 4 km/h above the civil speed limit before the brakes became effective.

This is not the case with a cbtc system. The “rules” for a typical cbtc system are (simplified) that if necessary the emergency brake should be applied to ensure that the train speed does not exceed the civil speed limit. This is a more onerous condition than the one previously discussed for a speed signalled system.

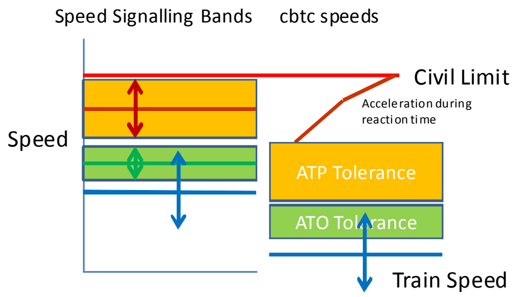

Figure 11: Speed Signalling and cbtc Comparison

A typical cbtc predicts what the train braking profile will be if the emergency brakes are applied, then applies them before the maximum safe speed is exceeded.

Some people view this as much better – since the train never reaches an “unsafe speed” – but remember how the civil speed limit is derived. Remember “safety is not a binary condition”.

It is my premise that this is an inappropriate interpretation of the concept of MSS and that it can result in a loss in performance.

6 Potential Loss of Performance and Possible Remedies.

One of the characteristic features of any metro line is the average journey speed. This is one of the factors by which passengers judge how good the transport system is. It also has a few other implications for the operator.

Let us consider a “typical” metro line of, say 15 stations, each a kilometre apart. Let us further assume that the trains run at a maximum speed of 80 km/h, and that they accelerate and decelerate at 0.8 m/s^2, and that the station dwell is 20s. Let us also neglect any complications caused by reversing trains at the terminal stations.

Then the inter-station run-time is (as shown below)…

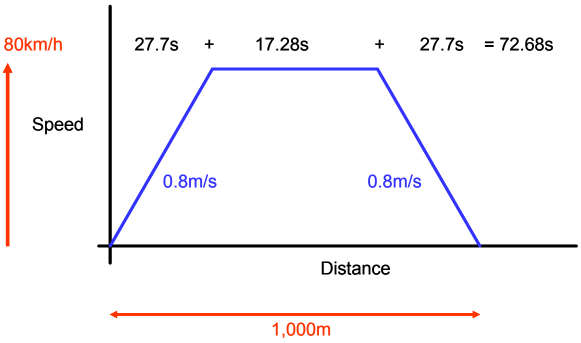

Figure 12: Interstation Profile

… about 73s – to which we can add 20s dwell time. This means that it takes 93s per station. With 15 stations this gives a round trip time of 30 x 93 =2,790s (or just over three quarters of an hour). Since the total distance is 30km this gives an average journey speed of 38.7 km/hr.

Now assume that the train speed is reduced because of the inappropriate levels of accuracy phenomenon that I have been discussing. It could be easily reduced by 10km/h - to about 70 km/h.

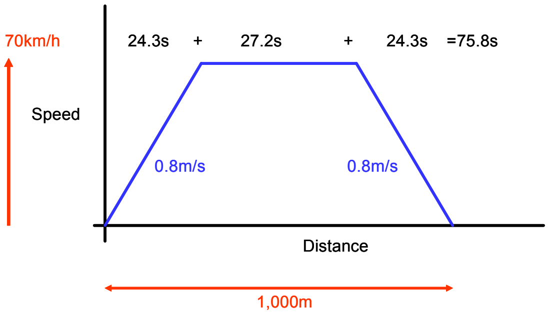

Figure 13: Reduced Speed Proile

The run time is now about 76s + 20s for the dwell, i.e. 96s. This gives a round trip time of 2,880s and a corresponding average journey speed of 37.5 km/hr (1.2 km/h slower – about 3%).

This immediately worsens the “Generalised Cost Benefit” associated with passenger travel on the line.

It can also have a serious impact on the operating cost as shown below:

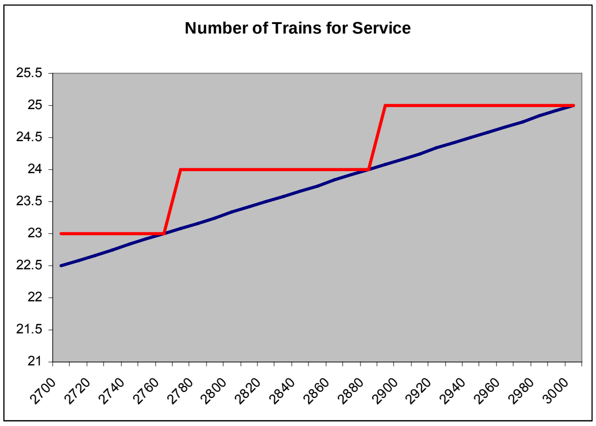

Figure 14: Number of Trains for Peak Service

There is a simple relationship between the average journey time (t), the train interval (i), and the number of trains needed for the service (n):

$$N = t/i$$

Clearly it is difficult to have a fraction of a train in service so “n” needs to be rounded up to a whole number.

As shown in the graph it is not unlikely that the reduction in train speed caused by using inappropriate levels of accuracy will create a need for an extra train to provide the service level. This will cost around £10m - £15m ignoring running and staff costs.

What can be done about this apparent problem? My suggestion would be to tell the civil engineers that we will operate the train, using a SIL2 (ATO) system that will enforce the civil speed limits. This will be more accurate and more reliable than a driver. We can explain that very occasionally (maybe once every few years) an ATO may fail and cause an overspeed condition. If this happens, the ATP (which will NEVER fail wrongside) will limit any such over speed to 10km/h. If this is not acceptable then maybe they would adjust their speed limits appropriately.

7 Summary

I have pointed out that civil engineers work in a culture that depends heavily on empirical formulae and safety factors. This is the context in which they establish the speed limits for a line. Although such speed limits are often enforced solely by the driver, metro systems normally use an ATP system to enforce them. Designers of ATP systems tend to be more pedantic then civil engineers and are used to a much more precise level of accuracy. In interpreting a civil engineers speed limit they have a tendency to interpret the level of accuracy inappropriately, which can result in the overall railway system being sub-optimal.

I have used a particular case to demonstrate this problem – which can be caused either between or within disciplines. I am sure it is not the only example. My plea is that we should all stay aware, and understand the data that we are dealing with.

====================================

References:

Amans,F., Sauvage,R., (1969), “*La Stabilite de la Voie vis-à-vis des Efforts Transversaux Exerces par les Vehicules*”, Annales des Ponts et Chaussees, Vol 1

Nagase, K., Wakabayashi, Y., Sakahara, H., (2002), *“A study of the phenomenon of wheel climb derailment: results of basic experiments using model bogies”*, Proceedings of the I Mech E Part F Journal of Rail and Rapid Transit Volume 216, Number 4.

ORE, C138, RP5 (1980), *“Effect of Train Speed on the Permissible Maximum Value of Load ΣY=S from the Point of View of Track Displacement”*, Utrecht

Profillidis, V.A., (1995), “*Railway Engineering*”, Avebury Technical, Aldershot.