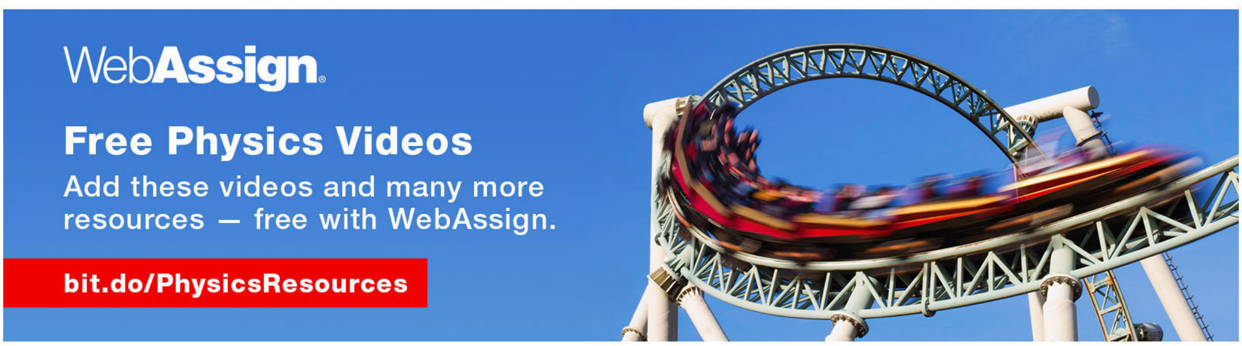
AMERICAN JOURNAL of PHYSICS

Measurement and Its Reliability: An Introductory Laboratory Experiment

Sherman K. Poultney

Citation: American Journal of Physics **39**, 176 (1971); doi: 10.1119/1.1986087
View online: http://dx.doi.org/10.1119/1.1986087
View Table of Contents: http://scitation.aip.org/content/aapt/journal/ajp/39/2?ver=pdfcov
Published by the American Association of Physics Teachers



Measurement and Its Reliability: An Introductory Laboratory Experiment

SHERMAN K. POULTNEY
Department of Physics and Astronomy
University of Maryland
College Park, Maryland 20742
(Received 7 July 1970)

An introductory laboratory experiment about measurement and its reliability has been developed for general science students. The experiment keeps their interest by focusing on the measurement outside the classroom of a quantity of daily interest (speeds of automobiles). It allows simple estimates of direct measurement errors, allows the experiment to be designed accordingly, and stresses the reliability (court room defense) of the measurement of the quantity of interest. At the same time, the kinematical concepts of average velocity, instantaneous velocity, and acceleration can be introduced as can the utility of significant figures and graphical display.

INTRODUCTION

The necessity of introducing the general student to the way physicists design experiments and deal with their measurement errors need not be questioned since physics is an experimental science. The necessity of introducing students to measurement and errors by having them sit in the classroom timing a metronome, for example, can be questioned; especially since the error distribution obtained is not necessarily normal, the estimate of error is not further used to illustrate its effect on a calculated quantity, and the student would probably be more attentive if he were asked to generate a fake distribution. Further sophistication at this point, such as standard error of the mean, is even more inappropriate. Experience with physics majors has indicated that the concepts of error estimate and error propagation are extremely difficult concepts for even them to learn and apply. What is needed is an interesting experiment in which simple estimates of the measurement errors can be made by the introductory student, the experiment can be designed accordingly, and the student is motivated to inquire about the reliability of a quantity he calculates using his measurements. Further sophistication can be introduced after the student understands the basic process.

THE EXPERIMENT

The experiment developed to illustrate the basic process of error estimate and propagation is included here as an appendix. The students are asked to design an experiment to determine the speed of cars on a campus road using stop watch and tape measure. The fact that they will be required to justify the reliability of their speed determination is made quite clear. The experiment occurs at the time that measurement and kinematics is being discussed in lecture. The student will have seen the Modern Learning Aids Film 0105 on Measurement[1] before doing this laboratory. Detailed instructions to the student about how to measure automobile speeds are omitted although he will estimate the timing error involved by measuring the period of a metronome in class and discuss the choice of a timing baseline with his instructor. Brief discussions for the student of the types of errors, how to estimate errors, and how these errors propagate to the calculated quantity are included as Appendices to the experiment. There are also specific points mentioned for discussion by the students in their report on the experiment.

THE EXPERIMENT IN ACTION

The University of Maryland campus is peculiarly suited for such an experiment since the administration has erected a number of highways right through the campus and never enforces the 20 mph speed limit. The students (who are usually also pedestrians) set up their speed traps with glee. They immediately begin to sense the benefits of a more controlled environment when cars turn into parking lots along their baseline (or just stop to let a rider out), when it starts to rain, or when drivers slow down (or speed up) upon sensing they are in a speed trap. In fact, most speeds measured were low for the latter reason.

The dominant error is in the time measurement. Minerva stop watches with 0.1-sec resolution were used in spring 1970 and could be read to about 0.02 sec. Figure 1 shows the distribution of times for the metronome period measurements. The 3.0-sec period was the longest one available. The distribution of 100 measurements is somewhat asymmetric, with a half-width of about ± 0.03 sec. The student is asked to take only the 25 measurements, to calculate the mean, and to calculate the standard deviation. This process itself takes about $\frac{1}{2}$ h. He uses the standard deviation as his estimate of timing error. An error associated with that of measuring the time interval is the one of deciding and signaling when the interval starts. The holder of the stop watch observes the end of the interval but relies on a partner's signal for the start. Let us assume the signaling problem has been satisfactorily worked out. We can check this assumption later. The error in the baseline depends on the method used to measure it but is not usually greater than a half a foot out of 50 ft.

The speed trap measurements are often done over a number of consecutive equal intervals to allow later use for a kinematics laboratory dealing with instantaneous velocity and acceleration. Both of the data tables shown below are for the extended measurements since they will allow us to examine the experiment more closely for appropriateness. In spring 1969, the measurements were done in the mks system. In spring 1970, they were done in the English system because the meter tape had been stolen (a recurrent problem at Maryland). The students were asked to convert their answers to both systems of units.

Tables I and II show extended measurements taken by classes in 1969 and 1970. Most time measurements in spring 1969 were taken with 0.2-sec resolution watches. In spring 1970, either the 0.1-sec stop watches had not completely replaced the 0.2-sec ones or else few of the students estimated to a fraction of the resolution interval. The same problem arose in the timing of the metronome. The data in Fig. 1 were taken during a preparatory laboratory session. Most students

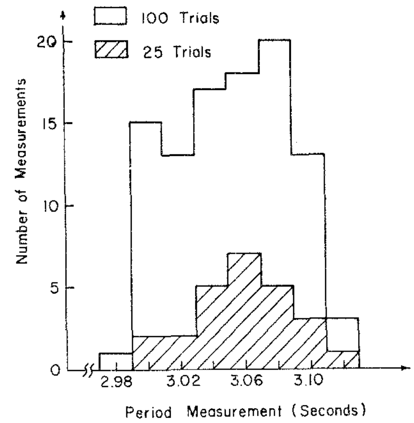


FIG. 1. Distribution of metronome period measurements.

rounded off to the nearest 0.05 sec in the measurement of the variation in their reaction time. The process of estimation will have to be emphasized more in the future.

The choice of baseline is decided in part by how it will be measured and in part by the estimated timing error. A speed of 30 mph (speed limit is 20 mph) equals 44 ft/sec. Thus, for a timing error of ± 0.03 sec, a baseline of about 50 ft would yield an error in speed of $\pm 3\%$. The error in baseline measurement is usually significantly less than 3% (e.g., 1%). The faster the car, the more the timing error will contribute to the error in speed. The error propagation formula used for design purposes is

$$\Delta v/v = \Delta s/s + \Delta t/t, \tag{1}$$

where the speed v is related to distance s and time t by

$$v = s/t. \tag{2}$$

The data in Tables I and II are taken for baselines of 10 m and 50 ft (and multiples thereof), respectively. The average speed over the whole interval is also listed. Its error is much less than the 3% selected above, but then any speed fluctuations on a shorter time scale are averaged out. The fastest car had an average speed of 43 mph over the whole interval. Speeds measured in 1970 are closer to the speed limit due to construction in the vicinity. Inspection of the data will show some highly nonuniform speeds, some timing errors, and some station misses. The baseline unit

TABLE I. Cumulative travel time in seconds for automobiles going north on Regents Drive 17 February 1969 for fixed distance intervals.

Car no.	10 m	20 m	30 m	40 m	50 m	60 m	70 m	80 m	90 m	100 m	$\bar{V}$ (m/sec)	$\bar{V}$ (mph)
1	1.0	2.2	3.6	5.2	6.0	7.0	8.15	9.0	10.7	12.1	8.3	18.8
2	1.6	2.0	3.6	4.6	5.5	6.1	7.0	8.0	9.1	11.2	8.9	20.2
3	1.2	2.3	3.8	5.1	5.2	7.0	8.05	8.8	10.4	11.0	9.1	20.6
4	1.8	1.9	2.9	4.2	5.0	6.2	7.0	8.6	10.6	10.9	9.2	20.8
5	1.1	1.7	3.0	4.5	5.0	6.2	7.10	8.1	9.4	9.7	10.3	23.3
6	0.8	1.1	2.0	1.8	3.2	3.6	3.6	4.2	5.0	5.3	18.9	43.0
7	1.0	1.4	3.2	4.0	5.1	6.3	7.15	8.5	9.6	10.2	9.8	22.2
8	1.3	1.3	2.5	3.7	4.2	···	3.75	4.4	7.2	7.7	13.0	29.5
9	1.0	2.6	3.1	4.5	5.8	7.0	7.7	8.7	10.0	10.3	9.7	20.8

is long enough to give a speed error of 3% or less for each unit. Figures 2 and 3 give the average speed of selected cars from the data tables for each unit plotted consecutively. The students would make such a graph for each car as part of a kinematics laboratory the following week. They would also calculate and graph the acceleration of each car. These graphs would be used to illustrate the concept of instantaneous velocity and acceleration.

The initial synchronization and signaling can cause considerable error; especially in the first interval unit. Each foot of synchronization error would cause a systematic error in the first unit velocity of about 2% for a speed of 30 mph. However, if the light signal reception is uniform, all other interval unit times should be free of this systematic error. About half of the entries in Tables I and II show synchronization errors up to several percent. This conclusion was determined by comparing the average velocity for the $(n-1)$ cumulative interval with that for the similar cumulative interval which omits the first interval unit. Figures 2 and 3 indicate that the signalling was done quite well by visual flagging if human error is taken into account (e.g., interval 2, 4, 5, and 7 of Fig. 2). Light signals, sophisticated electrical switching, or even sound signals could have been used. This discussion shows that it is wise to have several consecutive interval units for the first week laboratory. With correct signalling, the error in the interval unit times is just the probable difference in reaction times measured in the laboratory.

Throughout the experiment, the concept of significant figures should be stressed, especially in the

TABLE II. Cumulative travel time in seconds for automobiles going south on Regents Drive 17 February 1970 for fixed distance intervals.

Car no.	50 ft	100 ft	150 ft	200 ft	250 ft	300 ft	350 ft	400 ft	$\bar{V}$ ft/sec	$\bar{V}$ mph
A-1221	1.1	2.8	4.3	5.57	···	···	8.7	11.8	33.9	23.1
BA-1932	1.3	2.9	4.88	6.75	8.5	10.4	12.1	13.5	29.7	20.2
State-2583	1.27	3.0	5.02	6.45	7.8	9.2	10.6	11.8	33.9	23.1
EH-4312	2.5	4.8	6.49	9.0	11.4	13.2	15.5	18.4	21.8	14.7
EC-9389	1.6	3.3	4.40	5.5	6.9	7.9	9.4	10.5	38.0	25.9
GJ-7957	2.1	4.0	6.08	7.6	9.5	11.1	···	14.7	27.5	18.7
FF-2670	1.3	2.6	4.00	···	6.0	···	···	···	···	···
DT-7546	1.8	4.0	5.9	7.4	8.3	10.6	12.4	13.9	28.8	19.6
NVM-238	1.5	3.7	7.38	9.7	···	14.9	17.0	19.1	21.0	14.3
RCY-509	1.1	2.4	3.38	3.68	5.0	6.1	7.3	8.5	47.0	32.0

calculations and in the statement of results. The students are prone to carrying too many figures and thereby increase their arithmetic problems. This excess calculation also appears in the choice of a baseline. Further sophistication about random errors is postponed until late in the semester when a laboratory concerned with Brownian motion is held.

Many students wonder what their reaction times are with respect to the intervals they are measuring and with respect to the variations in reaction time that they measure. This interest can be channeled into a third laboratory which would be the study of the repeatable physical phenomenon of the falling body. The distance a body falls in a given time can be found using stroboscope and scale or using a sophisticated scheme such as the time-interval meter and light sensor apparatus currently used in Physics 10 at the University. Knowing this law, the student can then proceed by dropping a meter stick for a partner to catch as described by Daw and Fowler.[2]

CONCLUSIONS

The processes of experiment design, making measurements, and stating reliability of results are studied by introductory students by means of a simple experiment outside the conventional laboratory setting. The experiment naturally comes as the beginning laboratory experiment for the general science student. The evaluation of the approach is not yet complete. However, as is typical of evaluations in teaching, I conclude that the content is appropriate, the style interests the student, and the better students do well.

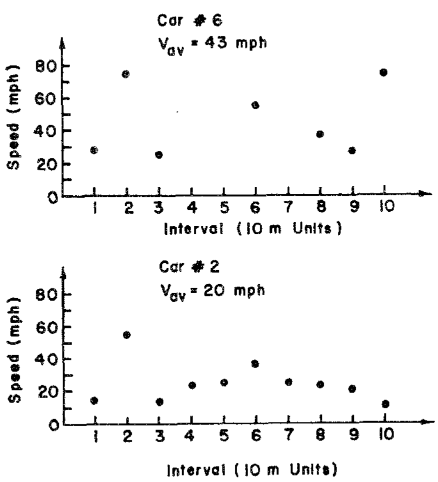


FIG. 2. Average speed of automobile in an interval as a function of interval number, spring 1969.

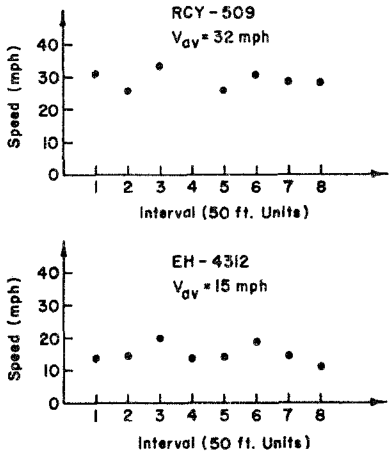


FIG. 3. Average speed of automobile in an interval as a function of interval number, spring 1970.

ACKNOWLEDGMENTS

I would like to thank both the students and the graduate student teaching assistants of Physics 10 for their help in carrying out this laboratory, especially Mark Mussachia and Lauren Rauber.

APPENDIX: PHYSICS 10 EXPERIMENT I: MEASUREMENT AND ITS RELIABILITY

1. Introduction

The ultimate test of a physical theory is based on experimental measurements. However, no measurement is exact. Any measurement (length, time, mass, etc.) has an error associated with it no matter how carefully and painstakingly made. In daily life, a measurement is useful only if experience shows that its error is small. A scientist must be able to estimate the magnitude of the error in each of his direct measurements (e.g., length and time) and to calculate the resultant error in his indirect measurements (e.g., velocity or acceleration). The magnitude of resultant error determines the scientific worth of the experiment. Appendix I discusses types of direct measurement errors, ways to estimate their magnitudes, and how these errors propagate to a resultant error in the calculated quantity.

2. Experiment Synopsis

You must devise a plan to measure the speed of cars traveling on Regents Drive using the equipment listed below. A traffic official of the University expects to receive a report on the speeds of the cars you observe. Moreover, you will also have to present your evidence before a judge in traffic court (he is also an off-duty physicist) about the speed of a particular car that was involved in an accident, and so you must be prepared to make a clear statement as to the reliability of your measurements. Don't forget to get the license plate number.

3. Equipment Available

Stop watches, meter sticks, tape measures, partners, passing cars, and Regents Drive on the west side of the physics building.

4. References [Ford[3]]

a. Measurement of time and length: Ford 3.1, 3.2, 3.4.

b. Uniform velocity: Ford 3.3, p. 150–152, 7.4, p. 180.

c. See also Modern Learning Aids Film[1] 0105 (to be shown in lecture).

d. Experimental error: Ford 7.6 and Appendix I.

e. Probability: Ford 14.1, 14.2, or your own choice.

f. Graphs: Ford 7.4.

g. Instantaneous velocity: Ford 8.2.

5. Suggestions

a. It may be more convenient to calibrate your stride rather than use a tape measure.

b. Do the stop watch test suggested in Appendix I, Pt. D.

c. Take care that you are not the accident.

d. Try to minimize your effect on the driver's speed.

6. Report Discussion

a. Record the results of the stop watch test. Calculate the mean stopping time. Calculate the standard deviation of the distribution. Refer to Appendix II and Appendix III. Do your conclusions depend on the length of time you practice? Why? Does this affect your timing of the cars?

b. Suggest a means of determining your reaction time.

c. Discuss your choice of baseline for timing.

d. Show clearly the reliability of your measurements for the fastest car. Refer to Appendix I.

e. What is the speed of the fastest car? What is its velocity?

f. If you have not used the metric system for measurements, what is the speed in cgs units? mks?

7. Laboratory for Next Week

a. Measure the speed for several different cars over a number of consecutive intervals along Regents Drive and graph the results. Join forces with several other pairs of partners.

b. Relate your graph of speed versus interval on Regents Drive to the concept of instantaneous velocity. Discuss. Graph the instantaneous acceleration of the cars. Show clearly the reliability of your results.

Appendix I: Measurement Errors

A. Types of Errors[4]

Many types of errors in making measurements are imaginable. Two classes of errors are precisely defined: systematic errors and random errors. The remaining kinds of errors include:

Blunders: Misreading a scale, writing down a wrong number, starting a watch at the wrong time, etc.

Chaotic errors: The equipment fails during a measurement, the experimentalist sneezes at the instant he should make a reading, the car you are observing stops or turns into a parking lot, etc.

Blunders obviously should be kept to a minimum and will not be discussed here. If a chaotic error occurs during a measurement, that measurement should be ignored.

Systematic errors are errors associated with the particular instruments and technique of measurement used. For example, a stop watch that is running slow, a tape measure with a foot cut off one end, or a head that is cocked to one side so as to introduce parallax into a scale reading. Systematic errors that are known can be corrected or compensated for. Unknown systematic errors are usually discovered only when someone else does the same experiment and obtains a different result. A good experimentalist is one who is able to eliminate significant systematic errors.

Random errors are unpredictable variations in the results of otherwise identical measurements.

It is these *ever present* errors that lead to the statistical interpretation of measurements. *We will assume that we have been good experimentalists and so must deal only with random errors.*

B. Horseback Estimate of Measurement Error

In each of the measurements that we do (e.g., length), we make some random error. Imagine that your meter stick has interval divisions (centimeter) such that you must estimate length to a fraction of that interval. Your ability to estimate that fraction sets an upper limit on the random error that you make. For example, if you estimate that fraction to the nearest one-fifth, you should quote a measured length of 132.4 ± 0.2 cm. Likewise, the time you measure may be 1.01 ± 0.01 sec.

C. Error Propagation

The question now is how do these errors in the length and time affect the speed that you wish to calculate. A student who has had calculus can easily find this. You will probably have to calculate the speed for 132.4 cm and 1.01 sec, 132.6 cm and 1.00 sec, etc., so that you may quote a speed of 131.6 ± 1.3 cm/sec. However, for small errors, you may use the following formulas:

$$x = ab/c,$$
$$\Delta x/x = \Delta a/a + \Delta b/b + \Delta c/c, \quad (1)$$

where Δa, Δb, and Δc are the errors in length (e.g., ± 0.2 cm), etc., and Δx is the error that you wish to know. Sixty percent of the time x will be within Δx of the "true value."

In the above example, $v = l/t$, $\Delta v/v = (\pm 0.2/132) + (\pm 0.01/1.01)$, $\Delta v = \pm 1.3$ cm/sec, $v = 132 \pm 1$ cm/sec. Which error dominates?

D. More Correct Estimate of Measurement Error

If the meter stick that you used above had had very fine interval markings, you might at first think that you would have no measurement error. However, due to random irregularities of the edge of the object being measured and of your placement of the meter stick, you would actually observe a distribution of lengths. You may more easily believe that the times at which the second hand of a stop watch stops when you try to time a periodic time interval vary randomly about some mean time. These variations are primarily due to slight changes in your reaction time from measurement to measurement. Try it using the electric metronome in the laboratory set on its longest period (~ 2.5 sec). To obtain a better estimate of your time measurement error, you should now determine the distribution of times, calculate the mean time, and calculate the standard deviation of the distribution. Use about 25 readings. The distribution is narrow, but somewhat asymmetric.

You would then use that standard deviation in place of the Δb above and do likewise for the standard deviations of the other measurements. Then you would calculate x and Δx.

NOTE: By repeating measurements of a deterministic physical phenomenon (metronome ticking or ball dropping), you can reduce your random error. This concept may be familiar to you from a course in statistics, and we may have opportunity to use this technique later in the course. In this experiment, the speed of a given car is unique, never-to-be-repeated happening.

E. Correct Calculation of Error Propagation

This subject requires a knowledge of the calculus. The resultant formula is as follows:

$$x = ab/c; \qquad (\Delta x/x)^2 = (\Delta a/a)^2 + (\Delta b/b)^2 + (\Delta c/c)^2. \quad (2)$$

Why is this a better measure of the error?

Appendix II: Mean and Standard Deviation

Assume you measure stopping times of t_1, t_2, t_3, $\cdots$, t_N. The mean time is

$$t_{\text{Av}} = N^{-1} \sum_{i=1}^{N} t_i = N^{-1}(t_1 + t_2 + t_3 + \cdots + t_N). \quad (3)$$

The standard deviation of the distribution is

$$\sigma = [N^{-1} \sum_{i=1}^{N} (t_i - t_{\text{Av}})^2]^{1/2}$$
$$= N^{-1/2}[(t_1 - t_{\text{Av}})^2 + (t_2 - t_{\text{Av}})^2 + \cdots + (t_N - t_{\text{Av}})^2]^{1/2}. \quad (4)$$

Appendix III: Distribution of Measurement Errors

Unlike many other experimental disciplines, physics most often deals with error distributions that are narrow and symmetric. In fact, error distributions in physics are commonly described by the Gaussian or normal distribution. The probability of obtaining a stopping time between t and $t+\delta t$ from a normal distribution described by t_{Av} and σ_t is given by

$$P(t)=[1/\sigma(2\pi)^{1/2}]\exp[(t-t_{Av})^2/2\sigma^2]\delta t. \quad (5)$$

Does the stop watch test yield a normal distribution of errors?

Appendix IV: References for Appendices

J. Topping, *Errors of Observation and Their Treatment* (The Institute of Physics, London, 1957).

D. C. Baird, *Experimentation: An Introduction to Measurement Theory and Experiment Design* (Prentice-Hall, Englewood Cliffs, N. J., 1962).

[1] W. Siebert, *Measurement*, PSSC Physics Film 0105 (Modern Learning Aids, New York, 1958).

[2] H. Daw and J. Fowler, *Variations on the Falling Body Experiment* (Commission on College Physics, College Park, Md., 1968).

[3] K. Ford, *Basic Physics* (Blaisdell, Waltham, Mass., 1968).

[4] Staff, "Physics 10 Laboratory," Department of Physics and Astronomy, University of Maryland.

Relation between the Binding Energy and Scattering Phase Shift*

K. F. CHONG†
Y. NOGAMI
Physics Department
McMaster University
Hamilton, Ontario
(Received 23 July 1970)

It is "well known" in the potential scattering in quantum mechanics that it is, in principle, impossible to obtain information about the binding energy from the scattering phase shift, and vice versa. This is illustrated in a simple way for the case of a nonlocal separable potential. The implications of the "relation" between the deuteron binding energy and the low-energy proton–neutron scattering parameters in the ^3S state are discussed.

In the potential scattering theory of quantum mechanics, it has been accepted for a long time that in principle it is impossible to obtain information about the binding energy from the scattering phase shift, and vice versa.[1] However, we often find statements in the literature which give an opposite impression. For example, many textbooks[2] discuss the "relation" between the binding energy of the deuteron and the scattering length and effective range for the pn scattering in the ^3S state. A typical argument goes as follows: The asymptotic form of the scattering wave function $\psi(r)=u(r)/r$ is given by

$$u(r)\sim\exp(-ikr)-S(k)\exp(ikr)$$

with

$$S(k)=\exp[2i\delta(k)]. \quad (1)$$

Here k is the wave number and $\delta(k)$ is the phase shift. If there is a bound state with the binding energy $(\hbar\alpha)^2/2m$, it will correspond to a pole of the "S matrix" $S(k)$ at $k=i\alpha$. The above $u(r)$ will then behave like $\exp(-\alpha r)$ for $k=i\alpha$, which can be interpreted as the asymptotic form of the bound-state wave function. If the phase shift is given by the effective-range formula

$$k\cot\delta=-1/a+\tfrac{1}{2}r_0k^2+\cdots, \quad (2)$$