

APPLICATIONS OF TECHNOLOGY IN TEACHING CHEMISTRY
An On-Line Computer Conference
June 14 TO August 20, 1993

PAPER 6

INDIVIDUAL COMPUTER-GENERATED GRAPHICAL PROBLEM SETS

Frank M. Lanzafame	Department of Chemistry
Monroe Community College	1000 East Henrietta Rd.
Rochester, NY 14623	(716) 292-2000 Ext. 5130
Internet: flanzafame@eckert.acadcomp.monroecc.edu	

SCHEDULE: Short questions on this paper: July 5, 1993
Discussion of this paper: July 12 and 13, 1993

ABSTRACT:

A computer program has been written to generate individual graphical problem sets for students in general chemistry. Its purpose is to develop graphing skills through take-home assignments requiring each student to do his own work. Data is generated for several problems including: the equation relating different thermometer systems, Gay-Lussac P-T data, vapor pressure versus temperature data, and radioactive decay/first order kinetic data. Experimental statistical fluctuations are simulated. Answers are calculated for questions accompanying the data. Errors in all quantities are provided using regression errors in slope and intercept and the propagation of these errors to calculated quantities.

The generation and use of these problems in freshman chemistry will be discussed. The program runs in QuickBasic 4.5 on IBM PC compatibles. It is expected that the program as well as accompanying questions will be available via anonymous FTP to conference attendees who are interested in obtaining a copy.

I SOME STUDENT DIFFICULTIES IN WORKING WITH GRAPHS:

One of the deficiencies many students bring to general chemistry is their lack of facility with graphing and working with linear relationships. Some of the student problems with graphs are:

- A. The inability to plot linear data, finding slope and intercept.
- B. The inability to use the equation of a line to interpolate data
- C. After drawing the line, students often work with two actual data points rather than the line.
- D. Counting squares of graph paper to determine slopes. This is

sometimes learned in math class.

- E. The inability to conceive of a graph with a slope of several thousand units. In math class it usually ranged between 0.5 and 2.
- F. The failure to realize that slopes and intercepts have dimensions as well as magnitudes.

It was decided to use the computer to generate individual problem assignments involving linear relationships after noting:

1. The difficulties that many students have with graphs and linear relationships
2. The difficulty in thoroughly testing this area on exams due to time constraints
3. The tendency of some students to copy assignments.

The random number generation capability enabled the construction of unique problems with simulated experimental fluctuations in the data points.

II ADVANTAGES OF COMPUTER GENERATED PROBLEMS:

- A. Each student has his own unique numbered data set. Different sets are generated each semester with their own numbers. Since each student has a unique problem set, there is less concern about copying than when a single problem set is assigned.
- B. Grading is simplified since the computer generates all answers along with the standard deviations of slope and intercept, as well as the expected range of error in derived quantities. Alternatively, one can take the student's slope and intercept, and use a programmable calculator such as the HP 28S to generate the actual answers rather than use the range of values provided by the computer. It has been found that by working carefully, students can generally obtain results within one standard deviation of the regression values. It is also possible to grade with a combination of the computer generated results for students that are within the one standard deviation and for which computer answers are appropriate, and use the programmable calculator to check the work of students with larger variations in slope and intercept.
- C. These problem sets can serve to test in a more secure take-home fashion areas which can be a bit long for an hourly exam. Credit can be given as a lab exercise, quiz, or take-home portion of an exam. Having seen copying of laboratory reports over the years, one is reluctant to give take home assignments for credit. This type of individualized problem is less of a problem, although a student might still be able to coerce another into doing his work.
- D. Inclusion of simulated experimental errors results in realistic data which mimics experiments.

- E. While many students believe that they understand how to graphically obtain results from linear data, the results of grading these exercises over a number of years indicates that many overestimate their abilities. The ability to check the results of such data analysis by students is not easily available unless the instructor is willing to analyze each student's results from a lab experiment or endure the copying that takes place when a single data set is assigned.

III USES OF COMPUTER GENERATED GRAPHICAL DATA IN GENERAL CHEMISTRY:

A. TEMPERATURE CONVERSION--FIRST REVIEW OF LINEAR RELATIONSHIPS:

Early in the course, basic linear relationships are reviewed by deriving the equation for conversion between Celsius and Fahrenheit. (This has been done experimentally in our preparatory course using a Fahrenheit and Celsius thermometer.) To emphasize and drill the concepts, equations are derived for various thermometer systems with arbitrary reference points. Finally, the plotting of experimental data is emphasized by using computer simulated data for calibrating arbitrarily referenced thermometer systems. One of the two problems intentionally generates data with positive and negative data to test students' handling of positive and negative numbers.

B. COMPREHENSIVE GRAPHICAL EXERCISE:

Later in the semester, a more comprehensive graphical exercise is done. It is placed after gas laws are completed and before phase relationships are covered. The exercise has three problems.

The following is an example of the student data which is generated for the graphical exercise. It is printed in compressed mode on the left side of a sheet of paper. It is followed by the answers which are printed on the right side of the same sheet of paper. The instructor need only cut the answers away from the questions, and use them for grading student results. They can be returned to the student with his graded exercise.

SAMPLE STUDENT PROBLEM DATA

No. 9303 A. Gay-Lussac P (Torr) Versus T (Trebors)

Pressure	869.2	892.1	941.9	982.8	1036.6	1081.4
Temperature	96.4	103.3	120.3	137.5	154.1	168.3

No. 9303 B. Radioactive Decay Activity versus Time (Years)

Counts	5456	4488	2856	2321	2054	1463	1039
Time	25.1	51.0	77.4	103.3	125.7	152.7	174.1

No. 9303 C. Pvpap (Torr) versus Temp. (deg. C)

Pressure (Torr)	24.4	41.8	51.2	93.3	156.6	200.4
Temperature (C)	7.3	18.8	21.8	35.5	46.2	53.6

SAMPLE PROBLEM ANSWERS

No. 9303 A. Gay-Lussac	P (Torr)	Versus	T (Trebors)
	value	error	(rel.err)
Slope =	2.898	0.062	(0.021)
Intercept =	590.6	8.2	(0.014)
t (P=380) =	-72.7	4.5	(0.063)
t (P=0) =	-203.8	7.2	(0.036)

No. 9303 B. Radioactive Decay	T(1/2) =	64.17 Years
	value	error
Slope = -k =	-.00469	0.00026
Intercept =	3.8617	0.0289
A (t=0) =	7273	484
t (A=500) =	247.9	11.1
		(rel.err)
		(0.055)
		(0.007)
		(0.067)
		(0.045)

No. 9303 C. P _{vap} (Torr)	versus	Temp. (deg. C)
	value	error
Slope =	-4237	76
Intercept	11.65	0.25
t (P=760) =	90.6	1.7
t (P=10) =	-8.0	0.8
dH _{vap} =	35.2 kJoules	dS _{vap} =
		96.9 Joules
		(rel.err)
		(0.018)
		(0.021)
		(0.019)
		(0.103)

1. The A Problem is a Gay-Lussac problem which reviews a laboratory experiment done earlier. The students measured pressure versus temperature in Celsius and then extrapolated to an absolute temperature scale. Here, the computer generates pressure in Torr versus a fictitious temperature called a Trebor. This is named in honor of one of the 'older' faculty members at Monroe Community College, Robert Flanigan. ;-)

FIGURE 1 Graph of sample "A." problem: Pressure in Torr. versus Temperature in Trebor The figure shows the scatter in the points along with the best line (linear regression) through the points.

2. The B Problem utilizes the time integrated radioactive decay or first order kinetics equation. This relationship will not be seen until the second semester, but as with the Clausius-Clapeyron equation, is presented as an equation which can reduce the non-linear data to the equation of a straight line by plotting the log of the activity versus time. Log base 10 is used here for pedagogic reasons even though the equation naturally involves the natural logarithm. Once the equation is determined, it is used to predict the activity at a given time, and the time required to achieve a given activity. The time units are randomly chosen by the program.

FIGURE 2 Graph of sample "B." problem: \log (Activity) versus Time (here, in years) . The figure shows the scatter in the points along with the best line.

3. The C Problem is a vapor pressure versus temperature problem which requires the soon to be covered Clausius-Clapeyron equation. For purposes of this exercise, the equation is a given and the reduction to a linear relationship involving the natural log of vapor pressure in atmospheres as Y and $1/T$ (Kelvin) as X is shown. The slope is presented as $(-dH_{\text{vap}} / R)$ and the intercept as (dS_{vap} / R) . At this point enthalpy has been introduced, but entropy has not. Entropy is merely described as a thermodynamic property which will be covered later. Its magnitude and dimensions are determined from the dimensional analysis of the Clausius-Clapeyron equation. Once the equation is determined, it is used to predict the Normal boiling point and the temperature required to achieve a given pressure.

FIGURE 3 Graph of sample "C." problem: \ln (Pressure in atm) versus $1 / T$ (Kelvin) . The figure again shows the scatter in the points along with the best line for the Clausius Clapeyron problem.

This graphical exercise is sometimes repeated at the beginning of the second semester as a review of graphing skills which will be needed as the student encounters the temperature dependence of equilibrium constants, kinetic rate constants, and the integrated rate expressions. Since not all instructors use this exercise in the first semester, it is a review for some students and new for others.

C. VAPOR PRESSURES FROM MANOMETER MEASUREMENTS:

Another simulation generates data for the measurement of the vapor pressure of a volatile liquid versus temperature. This is an actual experiment which is done in a flask containing air and using a mercury manometer. A volatile liquid is injected into a flask through a septum seal, and the pressure is measured using the manometer. The temperature is controlled with a water bath. Similar to the Clausius-Clapeyron study earlier, this version generates the manometer readings which lead to the flask pressure and, after correcting for air in the system, results in the vapor pressure of the compound. Ideal gas law corrections must be applied to determine the actual vapor pressures of the compound. This experiment is often the last one of the semester, sometimes coming a few days before the final exam. To save time, the pre lab is given normally, initial data taken from the apparatus as a demonstration, and computer generated data is then given to the students. In this way, students can use the time required to collect the data to begin to calculate

the results (which in the worst case may be due the next day). The actual experiment works well, providing heats of vaporization which differ by only a few percent from the literature values for simple alcohols and ketones. When time is short, the computer simulated data saves two hours of data collection which is largely devoted to merely changing the temperature of the system.

IV USES OF COMPUTER GENERATED GRAPHICAL DATA IN ANALYTICAL CHEMISTRY

One simulation used in the analytical chemistry course consists in the potentiometric titration of an unknown polyprotic acid. The students actually do a potentiometric titration of an unknown mixture of phosphoric acid and sodium dihydrogen phosphate. At the same time, they analyze a computer generated potentiometric titration of an unknown mixture. The mixture can be a diprotic acid, its monoprotic salt, or HCl in a compatible mixture. The students must identify the composition of the mixture as well as the concentrations of the components. pKa's are determined from the pH of the buffer regions. The computer selects the mixture randomly from a table of polyprotic acids which have pKa's that favor potentiometric titration. The concentrations are also selected, as well as the sizes of pipets whose sample is titrated. Students not only get a very unique problem on which to work, but one for which it would be difficult to provide an unknown solution because of difficulty in obtaining the variety of compounds of suitable reliability. This simulation complements the experimental work which is done in the laboratory.

NOTE: this simulation has not been ported to QuickBasic on the IBM PC from its original version on the Atari 800. It is therefore NOT part of the package which is available by anonymous FTP.

V APPROACH TO PLOTTING THE GRAPHS:

The following points are emphasized with students in graphing:

- A. Use the 2-5-10 Rule to layout the axes. Calculate the number of units per division, then round to the nearest 2, 5, or 10 units per division to facilitate the plotting of the points.
- B. Label the axes, plot the points, check any outliers, and then fix the points in ink.
- C. With a sharp pencil, draw the best straight line through the points. If the placement of the line is not satisfactory, erase it and redraw it. Extrapolate the line over the axes at the edges of the graph.
- D. Read the two points at which the line crosses the axes. Use these points to determine the slope and intercept of the line.

The use of pen and ink encourages the student to redraw the line until it is the best possible representation of the points. Extrapolation off the edges of the graph encourages the student to take two convenient, widely spaced points for determining the slope

and intercept. This discourages the student from taking two data points which are close to the line, thereby ignoring the rest of the data points. Further, the widely spaced points gives the largest difference between two points, maximizing the significant figures in the calculation of slope and intercept.

VI DESCRIPTION OF THE PROGRAM

A. Random Number Generator:

The random number routines were taken from a collection of random number routines in "Microsoft Quickbasic Programmer's Toolbox" by John Clark Craig, Microsoft Press, ISBN 1-1-55615-127-6, p. 353-364. It was desired to have the randomly generated data be keyed on the problem number as a seed value. This permits one to be able to generate the same sets of data using the same problem numbers. This was deemed useful to reproduce the data in case the keys were lost or some other such disaster struck. This seeding on the problem number was not possible using the random generators provided with QuickBasic. While the RANDOMIZE statement appears to do this, repeated calling of this function does not restart the generator from the new seed value. Thus it was not possible to reliably obtain the desired effect with the built-in function.

B. Generation of Statistical Fluctuations about a Value:

The program uses a function called ErrFactor (relative standard deviation). This function returns a statistically generated multiplier with a mean value of 1.00 and a standard deviation given by the relative standard deviation specified. For example, if it is desired to apply a 5 percent fluctuation to a given value, the function called is ErrFactor (0.05). The function returns a randomly generated value of 1.00 +/- 0.05 which is applied as a multiplier to the value one wishes to randomize. Thus a multiplier between 0.95 and 1.05 is generated approximately 2 of 3 times. Since this follows a normal distribution, occasionally one finds the 2 or 3 or 4 sigma variation. This produces fluctuations with points which are outside the limit (here 5 %) about 1 of 3 times. Students are told that the data is realistic and occasionally points are measured which don't agree with the rest of the data set, and must be discounted in drawing the best line through the data.

C. Modular Program:

The program has been carefully written using a good structured BASIC. Extensive use of Functions and Subprograms provide a modular package of routines dedicated to problem generation. Variable and Constant names have been carefully chosen to be descriptive of the quantities represented. The use of "magic numbers" within the program has been avoided. For example, one finds variables such as SetsPerPage, NumberOfSets, SLOPE, dHvap, dSvap, Temp, Press, etc. A testing capability with on screen graphing has been included to look at the scatter in the points while tweaking the error factors.

It is hoped that the program will be useful to others in its present format, and that those with some BASIC programming experience might be tempted to modify and add their own routines to the basic problem generating environment which has been created.

D. The Clausius Clapeyron Problem:

As an example of the generation of a graphical problem, consider the Clausius Clapeyron Problem:

$$\ln P_{\text{vap}} = -\frac{dH_{\text{vap}}}{RT} + \frac{dS_{\text{vap}}}{R}$$

The program first selects the compound's boiling point in the range of 80 to 120 degrees Celsius which might be typical of compounds to be studied. Next, Trouton's rule is applied to pick a slope for this compound and a 10 percent factor is applied to randomize the slope. An intercept is chosen using the standard molar entropy of vaporization of 88 Joules per mole K also using a 10 percent factor to provide fluctuations. A set of 6 temperatures is randomly chosen in the range of 0 to 60 degrees Celsius which represent the temperatures at which the vapor pressures are measured. The corresponding pressure is calculated using the selected slope and intercept and then applying a 5 percent factor to each point to provide the experimental variation sought. This then is the simulated experimental data which the student must analyze.

Upon generating the simulated data, the program then analyzes the data using linear regression, reporting for the instructor the least squares slope and intercept as well as the standard deviation in both quantities. From experience it has been seen that a student working carefully to draw the "best" straight line through the data can generally come within one standard deviation of the slope and intercept. The intercept is a bit more sensitive, and some allowance should be made for values close to but outside of one standard deviation. It is reasonable to give full credit for a student coming within this range.

The student is then asked to calculate the Normal Boiling Point, the temperature at which the vapor pressure is 10 Torr, dH_{vap} , and dS_{vap} . The generated answer key provides these values along with the propagated error expected from the error in the slope and intercept. These values can be used to evaluate student results or, as was indicated earlier, a programmable calculator can be used to generate the exact values the student should get from his slope and intercept.

E. General Comments Regarding the Program:

Anyone who uses the program, since the source code is available in BASIC, is free to tweak the Errfactors to provide lesser or greater fluctuations. One should be careful in choosing larger factors. Since the specified fractional error is ONLY the

standard deviation, one gets larger fluctuations in the distribution from the random number generator routine.

Occasionally, more than one anomalous point will occur in the same graph. Students are told that this can happen in any experiment, and that they should do the best they can with the data they obtained. Since the program calculates the errors in slope and intercept, those values along with the student's graph serves to flag those few cases where statistics has been unkind. This should be taken into account in grading the student's work.

F. Hardware Requirements:

The program source code runs under QuickBasic 4.5 on an IBM PC compatible. A compiled version, PGEN11.EXE, is also being made available by anonymous FTP. The EXE versions runs under DOS without requirement of having QuickBasic 4.5.

The program has been written to drive an Epson compatible printer. However, only two special printer codes are used during printing:

ASCII 12 sends a FormFeed command to advance to the next sheet.

ASCII 15 sends an initialization command to turn on compressed mode on the printer. The print format assumes 132 character compressed font available per line of output. Student question data is printed on the left side of the page, while the answers are printed on the right side of the same page.

The capability of changing the printer initialization code has been added. The printer must have a compressed font available which provides at least 132 characters per line.

VII CONCLUSIONS:

The computer generated graphical problems has provided the author and other faculty at Monroe Community College with a useful tool with which to help teach chemistry and emphasize the importance of graphing and linear relationships to students of chemistry:

1. Unique problem sets provide plenty of practice without the temptation to copy someone else's work. Thus the instructor can provide some credit for the work as Laboratory, Quiz, or a portion of an Examination.
2. Grading with the combination of answer sheet and programmable calculator provide an easy way for the instructor to check and evaluate student performance.
3. Statistical fluctuations in the data provide practice in working with data which is realistic.
4. It shows the students whether they do (as they often think) or

whether they do not (as is too often the case) have the ability to work with graphical data and linear relationships.

VIII THE PROGRAM PACKAGE AVAILABLE BY ANONYMOUS FTP:

The following files will be available via anonymous FTP from Host:info.umd.edu in the Path:info/Teaching/ChemConference during the conference:

PGENBAS.TXT	The QuickBasic 4.5 source code for problem generation of graphical problems
PGENDOC.TXT	A documentation file describing the requirements of the package.
PGENQTMP.TXT	A text version of the temperature problem questions.
PGENQGRF.TXT	A text version of the graph problem set questions.
PGENQVAP.TXT	A text version of the Vapor Pressure Lab. for which data can be generated.
PGENQTTN.TXT	A text version of the potentiometric titration problem questions.
PGEN11.EXE	An executable, stand-alone version of the problem generator for those without access to QuickBasic 4.5 NOTE, this file is BINARY.

The complete package will be available in a single ZIPped file from the Simon Fraser University anonymous FTP site from Host: truth.chem.sfu.ca in the Path: pub/chemcai

PGEN11ZP.EXE	A self-extracting ZIPped collection of the 7 files described above. NOTE, this file is BINARY
--------------	---

REMINDER: To download a binary file from an FTP site, the BINARY command must be issued before GETting the file.

IX GIF FIGURES:

Three figures accompany this paper. They consist of graphs of the data for the type A, B, and C problems discussed in Section III B. The principal contribution they make to the paper is to give some idea of the statistical fluctuations which is built into the data provided to the student with these generated problems.

The graphs were generated with Quattro Pro, written to a PCX file, and then converted to a GIF file with PICEM. They are 480 X 640 pixels, and were done in black on white to give better quality paper copies than the colored versions provide.

X QUESTIONS:

1. Is this student math deficiency principally a Community College problem, or do Universities see similar problems?

2. Do other participants, particularly from other countries, find a similar math deficiency in their students or is this predominantly a U.S. problem?