DATA ANALYSIS USING A MOBILE DEVICE

DR. ABHIJIT PODDAR, MONALI PODDAR

Abstract: An android app has been developed to facilitate analysis of one or more sets of two-dimensional data using any mobile phone or tablet. The data may be keyed in using the app itself or imported from a file using a file browser. The data may be plotted on a customizable grid as a multi-set graph after sorting each data set if needed. Options to change axes ranges, axes types, graph captions and axes labels have been included. Provision for inserting annotations and other embellishments have also been made. Basic stats comprising minimum, maximum, mean and standard deviation values can be viewed. One, more or all of the data sets may then be least-square fitted to common mathematical functions (linear, polynomial, exponential, Gaussian etc.) or even to a user defined arbitrary mathematical function. The goodness of fit measure (R^2) for each data set can be obtained. From the fitted curves, the peak and bottom points, Y value(s) at any X value and also X value(s) at any Y value can be obtained and viewed on the graph. The slope at any point on the fitted curve may also be calculated using the graph grid, making the use of a graph paper redundant.

Keywords: Multi-setdata plotting; data analysis; least-square-fitting, arbitrary function fitting, Android app.

Introduction:

Motivation: The modern world revolves around 'data'. Knowingly or not, we are always generating and collecting data which needs to be cleaned, analyzed and processed for use.Sometimes the data also needs to be visualized in the form of bar, line or scatter graphs for better understanding and analysis. For many decades, the data visualization and analysis were and are still being accomplished by software applications [1-3] running on personal computers, workstations or servers. But of late, with the coming of age of small mobile devices like mobile phones and tablets, it is only natural that data analysis be made possible in these devices as well. It becomes all the more important in situations like that inside a laboratory class where students are tempted to visualize and analyze their raw experimental data. It is only through the mobile phone or tablet that they always carry in their pockets, can they perform such visualization and analysis on the go to find whether the data conforms to whatever they learnt in theory or is pointing to something new and hitherto unexplained.Till date there are few mobile applications or apps [4-7] which cater to data visualization and analysis. But they suffer from inherent constraints. Either they have command drive user-interface [4-5], have cumbersome navigation [6] or cannot fit multi-set data [7]. Keeping all the above in mind and our past experience in developing java applications [8-11] (Android apps are also built with java but make use of additional Android specific libraries), we have developed an android app, namely 'Lab Plot n Fit' [12] which has a nice button-driven user-interface, intuitive navigation, tackles multi-set raw data and has many new and innovative features which we would discuss in the next section.

Design and methodology: The app has been developed using Android Studio [13], the official IDE for developing android applications. The raw data to be visualized and analyzed may be keyed in as one or more X-Y data sets. Alternately, the single or multi-set data may be assumed to be present inside a file and may be imported inside the app. A file browser cum picker has been integrated to make the process easier.Provision for including a graph caption, axes labels and set distinguishing annotations has also been made. While importing the data from a file and assigning the numeric and non-numeric values to respective variables, pattern matching with regular expressions (regex) [14] has been employed. The data sets may be stacked one over another, or, may reside side by side. The numeric data may be in decimal or exponential format. Figures 1 and 2 show two screenshots of the app depicting how the data may look after being imported from a file. After the data has been incorporated, one may plot one, more or all of the data sets simultaneously on the same screen. A graph grid may drawn in the background with the option of changing the number of horizontal and vertical grid lines so as to mimic an actual graph paper. The axes ranges may be changed to visualize the nature of a select portion of the graph. The axes type may also be changed from linear to log to better visualize the data points which are spread unevenly over the entire axes range. One may also change the number of decimal places in the axes tick labels and the data wherever displayed, or better still, express these numbers in exponential format. Basic stats comprising maximum, minimum, mean and standard deviation values can be calculated and displayed for each data set at the press of a button. If an X-Y data set happens to represent a probability or frequency distribution of the X values, the mean and standard deviation of the X-values may also be found. The generated graphs may be in the form of bar, line or

scatter plots. The symbols depicting the data points or the lines joining them have been made customizable. To have a detailed view, one may pinch and zoom the graph too.Next, one, more or all of the data sets may be fitted to common mathematical functions, namely(i) Linear : $y=a*x+b$(ii) Quadratic: $y=a_0+ a_1*x +a_2*x^2$ (iii) Polynomial of degree>2 (iv) Exponential: $y=a*exp(b*x)$ (v) Logarithmic: $y=a*ln(b*x)$ (vi) Power-law: $y=a*x^b$ (vii) Gaussian: $y=a*exp(-(x-a)^2/(2*b^2))$. For example, Figure 3 depicts linear fitting of each set of data points corresponding to different base-current values in a graph for output characteristics of a bipolar junction transistor obtained from an experiment performed in a basic electronics laboratory class. One may notice that the origin of the X-axis has been shifted to the right so as to bring to the fore the linear portion of each data set. In situations where the plotted data points of a set cannot be fitted satisfactorily with the above stipulated functions, provision has been made to use one's own user-defined arbitrary mathematical function with any number of parameters to fit the data. The arbitrary function has to be entered in a text box and then it is parsed before using it to perform regression on the set of data points.Figure 4 depicts the use of an arbitrary function, viz., the mathematical expression for the current amplitude in an LCR series resonant circuit to fit the data from the related experiment conducted in an undergraduate physics lab. After each data set is fitted with a mathematical function, the 'Goodness of fit' or R^2 value, along with the final values of all the fitting parameters,is displayed. One may double tap on any fitted curve to obtain the top and bottom points of the curve, Y value(s) at any X value for one or more sets (as shown in Figure 3) , X value(s) at any Y value for one and more sets (as shown in Figure 4), as well as slope of the curve at any point. If a graph-grid is employed, one may mimic the calculation of slope on an actual graph paper (as seen in Figure 5) by drawing a tangent to the fitted curve at the point in question, drawing the corresponding right angled triangle and using the values of each smallest division along the X and Y grids and the number of smallest divisions making up the base and height of the triangle. All the data pertaining to the plots including the raw data, basic stats, values of fitting parameters, R^2 values etc. can be saved in a text file. If only the data is saved in a file, the file can later be imported inside the app itself to plot the data all over again. The graphs too can be saved as image files, both prior to and after fitting. One may share the saved data and images with others through email from right inside the app as well.

Conclusion: As can be seen from the above, a mobile phone or tablet can be used just as a desktop computer or workstation to incorporate data, visualize it in the form of plots, fit the data to mathematical functions and extract potentially myriad information from the fits.For a student in a lab class or anyone performing a scientific experiment for that matter, apps such as the one discussed above will prove to be very helpful in collating, visualizing and analyzing the results of his or her experiment.

References:

1. Gnuplot graphing utility, http://www.gnuplot.info
2. Origin graphing program, http://www.originlab.com
3. MATLAB, programming, visualization and analysis utility, http://www.mathworks.com
4. Gnuplot Mobile app on Google Play Store, http://MKLab.org
5. MATLAB mobile app on Google Play Store, http://www.mathworks.com
6. Curve Fitting & Matrix Solver android app on Google Play Store
7. Curve Fitter android app on Google Play Store
8. A. Poddar, "Topic-specific Learning Management Systems through self-contained Applets", International Conference on E-learning in theWorkplace, Columbia University, New York, June 9 to 11, 2010, edited by Dr. David Guralnick, ISBN number (CD): 978-0-9827670-0-9.
9. A. Poddar, "Applet based Learning Management Systems to teachElectronics", International Conference on E-resources in Higher Education, Bharathidasan University, Tiruchirapalli, India, February 19 and 20 2010, edited by Dr. S. Srinivasaragavan and Dr. E. Ramganesh, pp. 111-115, ISBN number: 978-81-908078-9-0.
10. A. Poddar, "Teaching of the phenomena of free, damped and forced oscillations in physics through an all-inclusive java applet", Physics Education Vol.29 No.1, Jan-March issue, article no.1, 22 pgs, 2013.
11. A. Poddar and M. Poddar, "Taking E-Learning to the Lab", Proceedings of the IEEE 2nd International Conference on Contemporary Computing and Informatics (ic3i) 2016, Noida, India, pp. 797-800 ISBN number: 978-1-5090-5256-1.
12. A. Poddar and M. Poddar, "Lab Plot n Fit" android app on Google Play (Latest Ver. 6.1 2017).
13. Android Studio, The official IDE for Android https://developer.android.com/studio/
14. Pattern matching with regular expressions http://www.regular-expressions.info
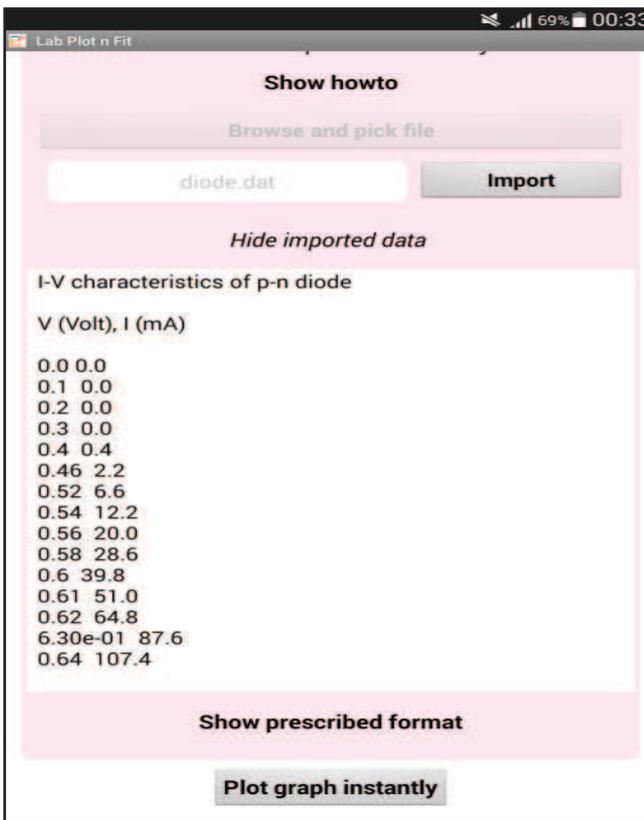
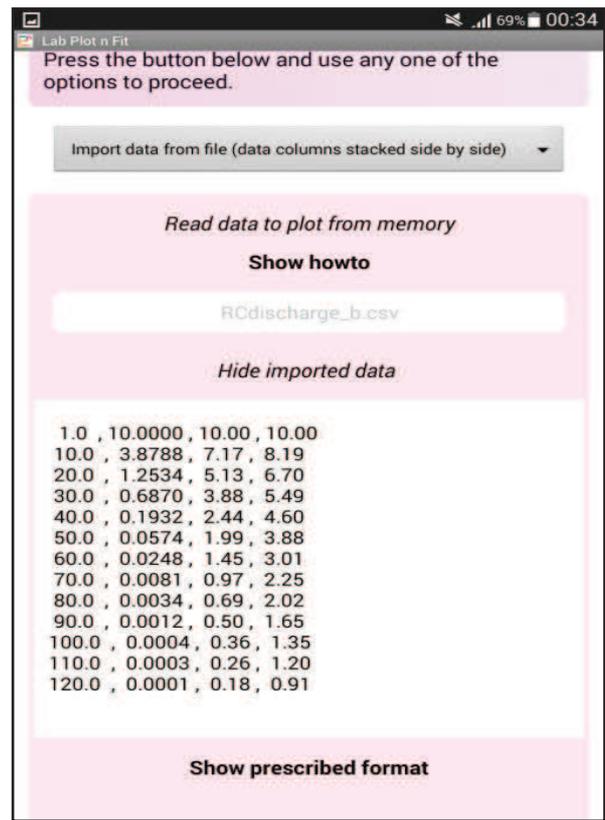**Fig. 1: A screenshot of data imported from a file using a file browser**

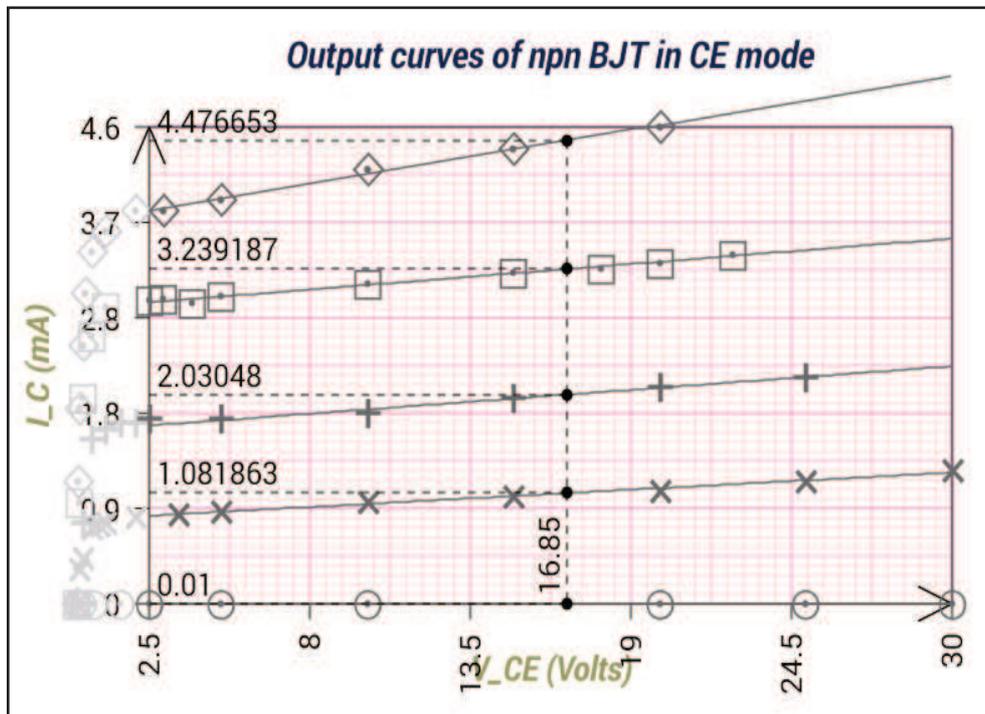**Fig. 2: A screenshot of multi-set data imported from a file**

Fig. 3: Example of linear fitting along with Y values obtained at any X value for different data sets.
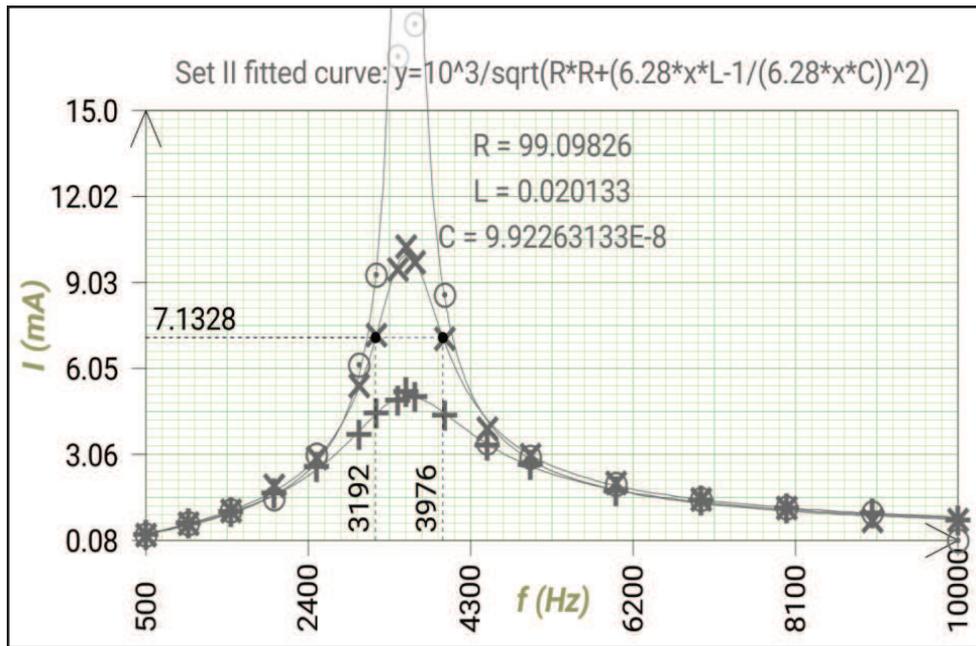
Fig. 4: Example of arbitrary function fitting along with X values obtained at any Y value after fitting
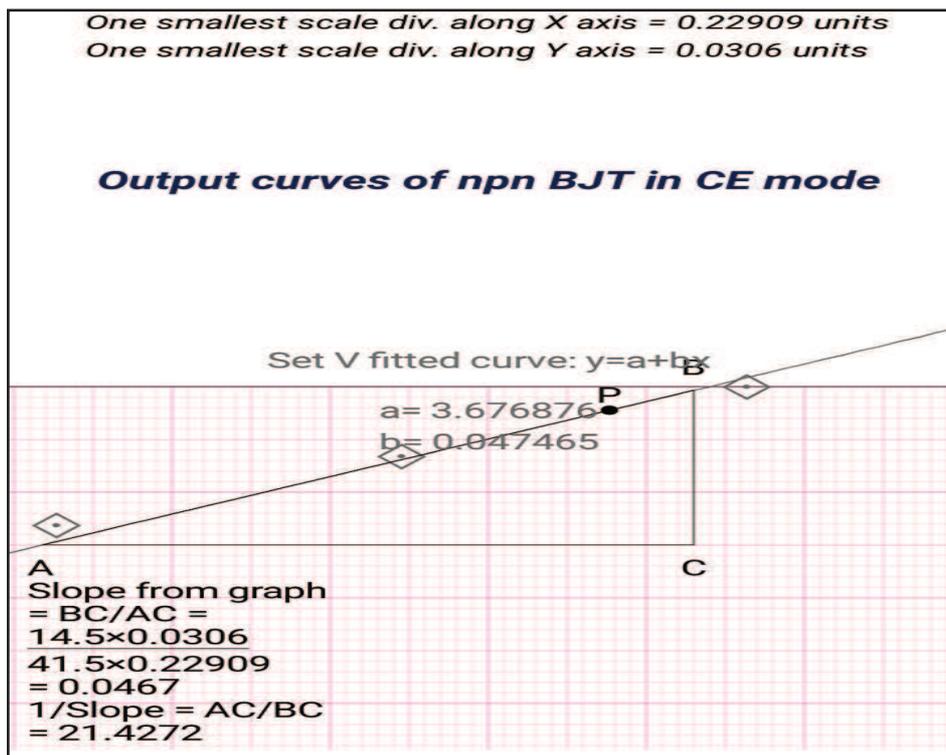

Fig. 5: Slope calculation just as is done using anactual graph paper

Dr. AbhijitPoddar
Dept. of Electronic Science, Surendranath Evening College, Kolkata, India 700009
Monali Poddar
Dept. of Computer Science, Maharaja Manindra Chandra College, Kolkata India 700003.