

LIII. *On Lines and Planes of Closest Fit to Systems of Points in Space.* By KARL PEARSON, F.R.S., University College, London*.

(1) **I**N many physical, statistical, and biological investigations it is desirable to represent a system of points in plane, three, or higher dimensioned space by the "best-fitting" straight line or plane. Analytically this consists in taking

$$y = a_0 + a_1x, \quad \text{or} \quad z = a_0 + a_1x + b_1y,$$

$$\text{or} \quad z = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_nx_n,$$

where $y, x, z, x_1, x_2, \dots, x_n$ are variables, and determining the "best" values for the constants $a_0, a_1, b_1, a_0, a_1, a_2, a_3, \dots, a_n$ in relation to the observed corresponding values of the variables. In nearly all the cases dealt with in the text-books of least squares, the variables on the right of our equations are treated as the independent, those on the left as the dependent variables. The result of this treatment is that we get one straight line or plane if we treat some one variable as independent, and a quite different one if we treat another variable as the independent variable. There is no paradox about this; it is, in fact, an easily understood and most important feature of the theory of a system of correlated variables. The most probable value of y for a given value