Alternate Development of Solutions to Jaynes' Functional Equation 2.45

by

Alan G. Herron, Ph. D
e-mail: A_Herron@telus.net
2004 May 8

Abstract

The following concerns itself with development of solutions to the functional equation 2.45 on page 32 of E. T. Jaynes' book "Probability Theory: The Logic of Science" [2]. But the development given there, however valid, appeared opaque and non-intuitive to this author. It is hoped that the following will provide a proof that is much easier to follow, and more suggestive of the essential concepts required in reaching the result, than Jaynes' work.

In a previous paper [1] it was shown that, in order to satisfy the self-reflective constraint S(S(x)) = x, the "negation" function S(x) must be "even" about the line s = x. It was subsequently determined, however, that this condition is necessary but not sufficient to satisfy Jaynes' equation 2.45. As a result the previous paper was withdrawn. It has been re-issued in view of the relevance of its essential contents to the following. I leave it to the reader to assess which parts of the development and comments therein are still correct and relevant.

The development below shows that the "even" property, although not sufficient, leads to a solution of the functional equation by a far more intuitive route than the conventional solution offered in Jaynes. One need only look to the implications of the "even" property for equation 2.45, along with a symmetry requirement on the s and x variables, to arrive at the general solution offered in Jaynes.

A Preliminary Result

Figure 1 shows a typical candidate for the function S(x), or T(u), a function which is "even" about the line s = x. A second set of coordinates, (u,t), is also shown. The t axis is the line s = x in the (x,s) frame.

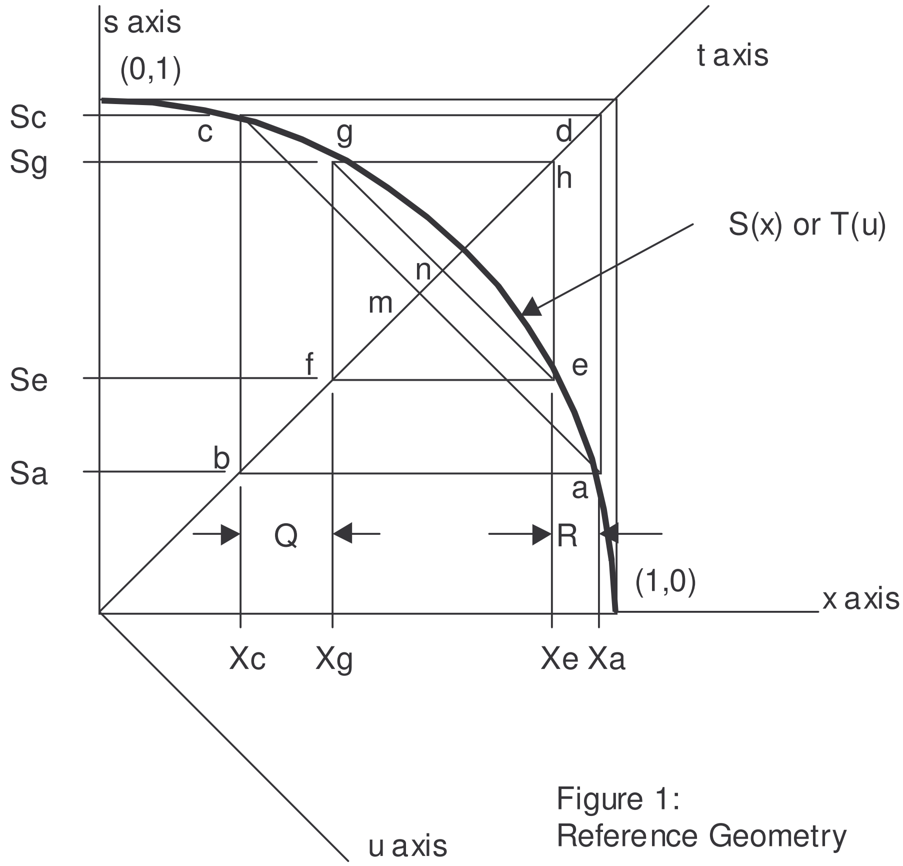

Figure 1:
Reference Geometry

As a brief review of the development in reference [1]…

"Even", in algebraic notation, means

T(u) = T(-u). (1)

The corresponding condition in the (x,s) frame is a "cross" assignment of coordinates. For example, considering the points "a" and "c",

Xa = Sc and Sa = Xc. (2)

The self-reflective points are found easily by "reflecting" an x (or s) coordinate in the line s = x. The reflection points are "d" and "b" in the above example.

Continuing on with the new development…

The centres of the self-reflective squares abcda and efghe respectively are the points "m" and "n", located by drawing the intersections of the other diagonal of the respective squares with the line u = 0 (or s = x).

Points "a" and "g" are chosen arbitrarily. Their self-reflective points, "c" and "e" are determined by the curve T(u). The "x" interval between "c" and "g" is defined as Q. The "x" interval between "e" and "a" is defined as R.

By inspection, these definitions imply that the "s" interval between "c" and "g" is also R, and that between "a" and "e" is also Q. In either case, Q, and, by implication, R, are both considered very small with respect to 1.

In general the intervals Q and R are not equal, with R being the lesser for curves concave to the origin. Their precise difference can be determined as a function of (Tn – Tm), but it turns out this is not important.

Derivation of Jaynes' (Cox's) Solutions

We are now in a position to derive the functional solutions to Jaynes' equation 2.45, viz.

$$Xa\ S(Sg / Xa) = Xg\ S(Sa / Xg) \qquad (3)$$

By inspection of Figure 1

$$Sg = Xe \qquad (4a)$$

$$= Xa - R \qquad (4b)$$

and

$$Sa = Xc \qquad (5a)$$

$$= Xg - Q \qquad (5b)$$

Substituting in (3) gives

$$Xa\ S((Xa - R) / Xa) = Xg\ S((Xg - Q) / Xg) \qquad (6a)$$

$$Xa\ S(1 - R / Xa) = Xg\ S(1 - Q / Xg) \qquad (6b)$$

A general Taylor Series expansion about (x,s) = (1,0) yields

$$S(1 - Z) = S(1) + S'(1)\ Z + O(Z^2) \qquad (7)$$

But one of our constraints is

$$S(1) = 0 \qquad (8)$$

Using (7) and (8) in (6b), and neglecting the high order terms gives

$$Xa\ S'(1) (-R / Xa) = Xg\ S'(1) (-Q / Xg) \qquad (9)$$

We have a problem here in that S'(1) may actually be negative infinity in the (x,s) frame. However it is the SAME S'(1) on both sides of (9), and thus may be cancelled anyway, leaving

$$R = Q \qquad (10)$$

as the requirement on S(x), in order to satisfy (3).

This condition can only be met if the centres, points “m” and “n”, of the self-reflective boxes, are coincident. The only geometry that creates the coincidence makes S(x) the alternate diagonal,

$$S(x) = 1 - x \qquad (11a)$$

or

$$T(u) = 1 / \mathrm{sqrt}(2) \qquad (11b)$$

Our development has thus led us to the “main” solution given in Jaynes, by a far more transparent development. But it seems to have left out the other solutions given by 2.58. Let us consider this further.

The solution, (11a), rewritten in a more symmetrical form, is

$$S(x) + x = 1 \qquad (12)$$

Following Jaynes’ arguments on page 33 concerning loss of generality, but in reverse, there is nothing preventing us from changing the “linearity” of the relation between S and x. Let us define

$$S(x) = y^m \qquad (13a)$$

and

$$x = z^m \qquad (13b)$$

Substituting in (12) then gives

$$y^m + z^m = 1 \qquad (14)$$

This is actually the SAME solution as (12), since any substitution for an independent variable in (3) would force us to adjust the dependent variable to correct for $m <> 1$ in (14). The basic conceptual relation, (11), is not generalized by adding the exponent “m” to the basic requirement.

In other words: by the time one corrects for the exponent to permit use in (3) one has, in essence, reduced (14) to (12) for mathematical purposes. The introduction of “m” is a NESTED change, not one on an equal conceptual level with the basic solution (12). See the Discussion below for more on the issue.

As an auxiliary point here, all solutions must be symmetrical in our independent choices, *viz.* points “a” and “g.” Otherwise the symmetry of (3) would be destroyed, along with the conceptual necessity requiring indifference as to which point is considered independent in setting up the problem.

This implies that the exponents of “y” and “z” must be the same. And the exponent cannot be 0, because this would make (14) an inequality. This conforms to the range specification on “m” associated with equation 2.58.

Thus the remainder of the solution family offered in Jaynes is easily derivable from the “main” solution. The essence of the solution, however, is the single function given by (11).

Discussion

The generalization to include an exponent “m” in (14) raises an interesting idea for algebra as a whole. This conceptual process has a “nested” structure, as opposed to substitution of other functions at the same conceptual level. The process can be applied in any algebraic situation. For example, given a parabola y = x^2, there is nothing preventing the substitution of, say, x = z^4 so our “parabola” becomes the power-eight equation y = z^8. The curve is the same functional relation between y and x even though x has become an “intermediary” between y and z, and has been eliminated algebraically from the relation between y and z. This isn’t a simple linear unit change, like feet to yards, or a “rotation”, but a non-linear change of “sensitivity” of one variable to the other. The power-eight equation isn’t a special case of the parabola. It is a relation between y and a DIFFERENT independent variable, whose definition assumes the relation of y and x in its definition sequence.

If the relation of y and x is conceptually important the “clues” to the difference should be left in the development. Confusion can be avoided if one retains different names for different independent variables, rather than re-defining the “x” in mid-development to serve as the “z” (as in y = x^8) based on an argument that the name is arbitrary. It isn’t arbitrary if the change obscures an essential stage in the exposition.

In the development in Jaynes, for example, this implicit re-definition has obscured an important point, viz. that the basic solution to equation 2.45 and it’s boundary constraints is really a single function S(x) = 1 – x, and not a whole family of functions. The remainder of the family can be obtained through the nested definition process if, and when, necessary.

As further confirmation of this problem, consider equation 2.57. This is a first order differential equation (albeit non-linear), and not a partial differential equation (pde). The solution should have exactly one arbitrary constant, not two, and should not be an arbitrary function (the case for pde’s).

The "true" arbitrary constant was set in Jaynes [2] to 1 to satisfy a boundary constraint (the general solution is $S^m + x^m = k^m$, not $S^m + x^m = 1$). Yet we also have "m" which appeared during the multiple approximations used in the derivation shown in Jaynes. Does this "m" really represent something in the original conceptual situation, or is it a mathematical artifact arising from the solution process? The above development suggests the latter.

Finally, once again I would like to express my thanks, and acknowledge the contributions of, the people credited in [1].

References:

(1) Herron, Alan G., Domain Limits and Functional Forms: Clarification and Extension of E. T. Jaynes' Results, private publication 2004 April 17, available from http://www3.telus.net/public/a_herron

(2) Jaynes, E. T., Probability Theory: The Logic of Science, Cambridge University Press, 2003, ISBN 0 521 59271 2