The Problem of Integrability in Utility Theory

Paul A. Samuelson


Stable URL: http://links.jstor.org/sici?sici=0013-0427%28195011%292%3A17%3A68%3C355%3ATPOIIU%3E2.0.CO%3B2-4

*Economica* is currently published by The London School of Economics and Political Science.

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ionschool.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.
The Problem of Integrability in Utility Theory

By Paul A. Samuelson

I. Historical Survey of the Integrability Issue

A chapter in the history of utility theory has now been brought to a close by Mr. Houthakker's important discussion of integrability. As far back as 1886, G. B. Antonelli had noted that such a problem existed and seems to have given the correct "integrability conditions". But to most economists this problem was first introduced by Irving Fisher's 1892 study—perhaps the best of all doctoral dissertations in economics. Even here it is introduced pretty much as an afterthought so that it is no great wonder that Pareto's Manuale di economia politica (1906) should have neglected this topic—even though Pareto was clearly aware of Fisher's work and had discussed integrability in journal articles as far back as 1893.

Vito Volterra in his 1906 review of the Manuale performed one of the few services professional mathematicians have ever rendered to economic theory: he pointed out that when Pareto treats the case of three or more goods, his discussion of indifference directions is marred by the failure to recognise explicitly the integrability problem. Pareto admitted his mistake and discussed the problem a few

1 Dedicated to the memory of Joseph A. Schumpeter, who had a lifelong interest in the logical foundations of utility theory.
3 G. B. Antonelli: Teoria matematica della economia politica, Pisa (1886). This seems to be a rare item. I know of it only from the extension of Jevons' Bibliography of mathematical economic writings, and from the detailed reference by Wold in a paper later to be cited.
4 Irving Fisher: "Mathematical Investigations in the Theory of Value and Prices", Transactions of the Connecticut Academy, Vol. IX, July 1892, and reproduced in a 1925 Yale Press edition, pp. 88–9. Recognition of non-integrability seems almost an afterthought in Fisher; since he uses his teacher, Gibbs' vector notation, Gibbs may have suggested the problem to him. If in 1892 Fisher was willing to entertain the hypothesis of non-integrability, it none the less seems that in all his later work he really believed in integrability.
5 V. Pareto: Manuale di economia politica, Milan (1906). I have not checked the non-mathematical text to see whether Pareto had really neglected non-integrability completely.
months later in the same journal; most of this discussion is repeated in the improved French version, the *Manuel d'économie politique* (1909) and deals with the puzzling problem of "open and closed cycles" of consumption.\(^1\) Professor E. B. Wilson,\(^2\) when he reviewed the *Manuel* in 1912, expressed the opinion that the economist could throw out of court the integrability problem, however natural it might be for mathematicians to worry about it. He also expressed the interesting opinion that as a result of some remarks of J. Willard Gibbs, Yale's distinguished physicist, Fisher may have changed his mind concerning the admissibility of non-integrability.

W. E. Johnson's classic 1913 paper on utility\(^3\) makes no mention of the integrability problem (nor for that matter of the work of Fisher and other important earlier writers). Eugen Slutsky\(^4\) in his remarkable, but too little known, 1915 Italian article showed that the existence of an integrable utility surface had definite testable implications for cross-elasticities of individual's demand: i.e., a "compensated change" in the price of good \(i\) must have exactly the same effect on the quantity of \(j\) demanded as does a compensated change in the price of \(j\) have on the demand for \(i\).

In his *Mathematical Groundwork of Economics* (1924) Professor A. L. Bowley does not, as far as I can see, anywhere deal with the problem of integrability. Professor G. C. Evans, an eminent mathematician, gives one of the most extensive discussions of this problem in *Mathematical Introduction to Economics* (1930).\(^5\) Most of the arguments for non-integrability are contained in his treatment; these will be discussed in some detail later.

---

We may now move into the modern era, which I arbitrarily date from R. G. D. Allen's foundation article of 1932.\(^6\)

---

5. Ch. IX, pp. 119–22.
This was a necessary preliminary for the later articles of Hicks-Allen and for their separate later writings. Allen was unacquainted with Slutsky's work but refers to most of the other important early writers. He entertains the hypothesis of non-integrability; and if I dare impute any differences to the separate components of the Hicks-Allen composite commodity, I would say that Hicks consistently rules out the non-integrability case, while Allen accepts it as the more general hypothesis. At least Allen in his 1932, 1934, and 1937 treatments deals at length with non-integrability, while Hicks in his 1934, 1937, and 1939 treatments goes out of his way to make it clear that he is against non-integrability. Hicks-Allen also noted an empirical implication of integrability that is the inverse of the Slutsky reciprocity relation: under integrability, when tea is a substitute for coffee vis-à-vis a third commodity, then according to their same 1934 definition, coffee must be numerically equal as a substitute for tea.

As a partial digression, I should mention the 1932 work of Hotelling on the related problem of profit-maximisation by a firm not subject to a budgetary constraint. Integrability conditions arise there which are related to, but distinct from, those that arise in the case of a consumer under budgetary constraint. In this connection, the reader can be referred to Henry Schultz' 1937 book which gives an extensive discussion of the integrability question: while not itself an entirely satisfactory resolution of the problem, Schultz' pages give a very good summary of the uncertainties that pervade the literature.

Professor Nicholas Georgescu-Roegen, now of Vanderbilt University, wrote in 1936 one of the most important

---

3 However, if we are willing to make very special assumptions of a good with constant marginal cardinal utility, we can write $U(x_1, \ldots, x_n) = x_1 + F(x_2, \ldots, x_n)$ for proper choice of units. This is to be a maximum subject to the budget constraint $1 \cdot x_1 + p_2 x_2 + \ldots + p_n x_n = 1$; and thus, with income effects emasculated, the consumer can be thought of as maximising without constraint $F(x_2, \ldots, x_n) - p_2 x_2 - p_3 x_3 - \ldots - p_n x_n + I$ which is precisely of the Hotelling 1932 form. (See pp. 592ff.) Later Hotelling addressed himself directly to the consumer problem and restated definitely some of the reciprocity conditions implied by integrability on the partial derivatives of price ratio functions of quantities of goods. H. Hotelling: "Demand Function with Limited Budgets", *Econometrica*, Vol. 3 (1935), pp. 66-78.

In 1943, Dr. Herman Wold of Uppsala University, Sweden, wrote three illuminating articles on demand analysis entitled "A Synthesis of Pure Demand Analysis".\footnote{H. Wold: "A Synthesis of Pure Demand Analysis", I, II, III: \textit{Skandinavisk Aktuarietidskrift} (1943), pp. 85–118, 221–63, (1944) pp. 69–120. See especially pp. 109ff. and pp. 78ff. The reference to Antonelli is on pp. 115 and 118.} Since these appeared in a rather specialised journal and during wartime, it is to be feared that they may suffer the neglect that was long the fate of Slutsky's fundamental paper. I know in my own case, although I was honoured with reprints, they arrived at a time when I was temporarily out of economics; I am ashamed to say that I have only recently read them after some correspondence with Mr. Houthakker; and I am even more ashamed of failure to cite these important papers in my \textit{Foundations of Economic Analysis} (1947).\footnote{Ch. V. V.}

Although the above list of references cannot pretend to be exhaustive,\footnote{For example, there are papers by L. Court and R. Roy on related problems. Also there is a recent paper by de Ville which I know of only from a brief review by K. Arrow in \textit{Mathematical Reviews}, 1947.} it does bring us down to the present time and to Mr. Houthakker's important contribution.

2. The Meaning of Integrability: Two-Dimensional Case

I do not think there is any simple way of picturing integrability conditions so that we can easily grasp their meaning in common-sense intuitive terms. This is because the picture must be in three-dimensional space at least, and because it must concern itself with subtle "local" conditions that slopes must obey. But I shall make the attempt to convey a notion of what is involved. After this is done, we shall be in a better position to appraise the main arguments in favour of and against non-integrability.
Let us begin in two dimensions where the problem of integrability does not arise. Imagine that you go to the south-west corner of a room. Look down at the floor, and let the corner be the origin from which goods \( x_1 \) and \( x_2 \) are to be measured. On the floor can then be traced the contour lines of indifference looking like Figure 1a. (The little arrows show the direction of preference of the indifference curves and are drawn perpendicular for convenience.) The same thing is pictured in Figure 1b, but as a consumer might reveal his preference field to us if we could observe his demand behaviour only externally. Thus, at point A

![Diagram of indifference curves](image)

the consumer had his choice of all the points on the straight line MN. But he chose A, and we draw a little line through A with a slope equal to the price ratio \(-p_1/p_2\) that brought him into equilibrium at A. Similarly at B or any other observed point we can draw a little "slope-element" (and if we like, a little arrow of preference); the budget lines whose price ratios determine the little slopes are, for simplicity, omitted. Figure 1c is another way of looking at the same problem. Any indifference curve, such as the one through A, can be thought of as the "envelope" or sheath of a family of budget lines, with only the heavy dark points being empirically observed.

It will be noted that any point where the indifference curves are convex rather than concave cannot be observed in a competitive market. (Cf. Fig. 1b.) Such points are shrouded in eternal darkness—unless we make our consumer a monopsonist and let him choose between goods lying on a very convex "budget curve" (along which he is affecting the prices of what he buys). In this monopsony case, we
could still deduce the slope of the man’s indifference curve from the slope of the curved constraint at the equilibrium point. The case of a Gallup-poll questioner who finds out the man’s preference contours by giving him choices of every pair of goods is simply a limiting case of monopolony. And if we experiment sufficiently, we can always find a curved family of unique contours representing his ordinal preference field—if such a consistent field exists. Therefore, in this generalised monopolony we can behaviouristically identify the man’s ordinal preference field if he has one.¹

But I shall confine myself for simplicity to what will reveal itself under perfect competition and shall be concerned only with concave regions. In a recent article on “Consumption Theory in Terms of Revealed Preference,”² I discussed how the little slopes of Figure 1b can be mathematically integrated into the smooth contours of Figure 1a. If the little slope elements are made very small and very numerous, our mind’s eye sees them as joining together into a one-parameter family of contours: it is as if iron filings on the floor line up to reveal the magnetic lines of force determined by a magnet under the floor. If the observed price ratio \( p_1/p_2 \) is given as the following continuous and differentiable function of the two goods, \( B(x_1, x_2) \), then mathematical analysis assures us that the differential equation

\[
(1) \quad \frac{dx_2}{dx_1} = -B(x_1, x_2)
\]

gives rise to a unique family of curves as in Figure 1a. In two dimensions there is no integrability problem.

Before leaving the two-dimensional case, I should give preliminary warning concerning two red herrings that are confused with the integrability question. First, a consumer may be myopic in the sense that he does not know what his tastes are like in situations quite different from his customary position; consequently his short-run demand behaviour may differ from the long-run “normal” behaviour.

¹ In his last and most general formulation, Encyclopédie des sciences mathématiques, Tome I, Vol. IV, Fasc. 4 (1911), Pareto wished simply to assume that an individual’s equilibrium choice would be given as an observable functional of every and any constraint offered him. But Pareto did not sufficiently realise that the functionals so defined would have to be highly restricted so as to give consistent results and define a consistent ordinal preference field. See N. Georgescu-Roegen: “Note on a Proposition of Pareto”, Quarterly Journal of Economics, Vol. 49, 1934-5, pp. 706-14. Professor Kenneth Arrow of Stanford University in a set of as yet unpublished papers has given a formulation of such consistency conditions in terms of logical sets.

which alone is our concern. But to be myopic means to be short-sighted—not zero-sighted or infinitesimal-sighted. We shall see in our multi-variable discussion that finite sight, however local, implies integrability; but the two-variable case drives home the irrelevancy of the problem of short-sightedness for the problem of integrability—since even if we choose to assume literally infinitesimal-sightedness, we still end up with a one-parameter family of contours and no integrability problem is possible.

Secondly, there is the related confusion between the “order of consumption” followed by a consumer and the “dependence of certain line integrals on the path” between two points. It must be emphasised that the paths along which I as an economic scientist choose to evaluate the man’s preference have absolutely nothing to do with the order in which the human guinea-pig consumes the goods. I don’t know whether he drinks his beer before his whisky or his whisky before his beer; I don’t know whether it even makes sense to say that he enjoys his shelter before rather than after he enjoys his food. Note too that in going from A to B the guinea-pig does not eat his way along the path, and in going from B to A regurgitate along the same path. Rather we should always regard the budget of goods at A as a steady flow of consumption per unit time, optimally patterned to the consumer’s tastes. And the flow of consumption goods at B is again a steady flow long maintained. The comparison of A and B (and of intermediate points) is a case of comparative statics. We need not invade the privacy of the consumer’s castle to concern ourselves with the minutiae of his domestic arrangements.

I stress all these complications because they are complete irrelevancies. And yet they arise in the two-variable case where there is absolutely no integrability problem! It should be absolutely clear, therefore, that the problem of “the order of consumption”, in the sense of the path along which the consumer actually moves behind the scenes of the market-place, has nothing to do with the problem of integrability versus non-integrability.

Pareto only confused the issue in his discussion of “open and closed cycles of consumption”. Actually, Pareto seems to end up with the conclusion that the consumer can be thought of as following optimal paths behind the scenes. From this he infers, if I understand his puzzling treatment,
that there is no longer any ambiguity of path and therefore integrability is assured. This is all hopelessly confused: even if the consumer has definite and optimal domestic-housekeeping habits, we shall still have to add strong restrictions on his revealed-preference behaviour once we have more than two goods. To repeat, Pareto’s primary confusion results from his identifying the paths of integration chosen by the economist-observer for statical comparisons with the behind-the-scenes programming of pleasures by the consumer.

In our two-variable case we have already integrated our indifference elements and have arrived at indifference contours. To repeat again: no integrability problem is possible and no integrability conditions need be fulfilled.¹

3. The Three-Dimensional Picture: The Integrable Case

When we bring a third good $x_3$ into the picture, we must use the altitude above the floor as the direction along which we measure the third good. Now we have three axes which meet in the south-west corner of our room. Along the lower edges of the wall, $x_1$ and $x_2$ are paced off as before. But the $x_3$ axis is measured along the vertical of the corner. Figure 2A shows the three axes. Any point in the room represents a combination of all three goods, as shown by A.

If we assume that the consumer has a consistent ordinal preference field, then we shall now have indifference surfaces, rather than indifference curves. How do these surfaces look? Rather like thin bowls inside of bowls, with the bottoms of

¹ See Mathematical Notes below.
the bowls pointing downward and south-westward toward the corner origin. As long as he moves on one bowl, the consumer is indifferent. As he moves toward "outside" bowls, nearer and nearer to the origin of zero consumption, he is worse off. As he moves away from the origin toward inner bowls, he is better off. Of course, this is all an imperfect picture: bowls are too thick, and umbrellas might be better.

Of course, any movement on an indifference surface will involve giving up something of one or more goods and taking on something of the other goods. Any movement off the surface and toward the origin involves a loss of ordinal satisfaction. Any movement off the surface in the opposite direction will involve a preferred move. If the consumer were free as a bird, he would want to get as far from the corner origin as possible, going upward and outward beyond the confines of the room and in the direction of bliss. But if we give him a limited money income, $I$, and prices of all goods, $p_1, p_2, p_3$, then he is not free as a bird. He cannot even move freely within the room. Instead he must stay inside the budget plane defined by

$$p_1x_1 + p_2x_2 + p_3x_3 = I.$$

Figure 2c shows the limiting budget plane beyond which the consumer cannot go. If every good is desired by the consumer, he will never stay inside the budget plane but will instead choose to crawl on its surface seeking his most preferred position. If $A$ is such an optimum position, at that point the budget plane will be touching but not crossing the highest attainable indifference surface; hence there will be tangency, with the "bottom" of the indifference bowl just resting against the plane in the same way that a tea-cup rests against its saucer.

Of course, if income or prices change, then the budget planes will change and a new optimum point of tangency will be found\(^1\): thus we define the general demand functions showing how each good will depend on all prices and income.

Thus far I have sidestepped the problem of non-integrability by assuming it away. I have implicitly assumed that the consumer does have a consistent ordinal preference field: that he can always tell of any two situations either that $A$ is

---

\(^1\) If the indifference surfaces do not have the usually assumed curvature, the consumer may rush away from certain parts of the room; he may even rush to one of the walls or floor and enjoy only a "generalised tangency". This often happens, but need not be studied here.
"better than" B, or B better than A, or that they are indifferent; and that his preferences among three or more situations are transitive in the sense that "if A is better than B and B better than C, then A is better than C", etc. In short, I have assumed that indifference surfaces do exist. But need they?

**The Three-dimensional Case: Non-integrability**

What is it that as scientists we are able actually to observe of a consumer who is in market equilibrium? Only his demand functions in response to all possible prices and incomes. Or, in geometrical terms, all that we can observe are points like A which lie on an observable budget plane and which represent an optimum as compared to other points on or inside that plane. The reader will be tempted at this point to ask: "Just as we traced out indifference curves in the two-dimensional case as the envelope of carefully changed budget lines, can we not hope in the three-variable case to trace out the bowl-like indifference surfaces as the envelope of carefully changed budget planes?"

The answer must be worded cautiously: "Provided that the indifference surfaces exist, yes we can hope to trace them out in this way. But in the three-variable case there may not exist anything corresponding to such surfaces—unless certain extra restrictions are placed on the consumer's demand behaviour."

To investigate further the conditions of integrability and non-integrability, let us paint in just what it is that we can observe in our geometrical room. At any point such as A, we have the budget plane going through A; our stage will get too cluttered up if we leave everything on it, so for simplicity we need at A indicate only a little "piece" of its budget plane. We need only indicate at A a little button, or better still a little thumb-tack, whose back or head lies in the budget plane and whose point tells us which is the "preferred" direction. At every observable point in our room there will be a little thumb-tack suspended in space; its exact orientation in space is indicated by its point, or equally well by the head, which lies perpendicular to the point. Of course, the thumb-tack is just a non-rigorous way of picturing a little "planar element" determined at each point by the two price ratios at which equilibrium takes place: if we let $x_3$ be our *numéraire* or
good whose price is taken as unity, then $p_1/p_3$ and $p_2/p_3$ will be assumed to be single-valued continuously differentiable functions\(^1\) of $x_1$, $x_2$, $x_3$, written as $p_1/p_3 = B^1(x_1, x_2, x_3)$ and $p_2/p_3 = B^2(x_1, x_2, x_3)$.

Our little planar elements or thumb-tacks are often referred to as “marginal rates of substitution” between the goods in question, or as “little indifference elements”. It is said that so long as we take a “little” step on the surface of the thumb-tack, the consumer is just “indifferent”; if we step off the thumb-tack toward its point, the consumer is moving in a “preferred” direction; finally, a move off the tack in the opposite direction is toward a “less-preferred” direction. In the present delicate investigation I must regard this as rather dangerous terminology. We have no right to regard our little elements as anything but shorthand ways of representing the observable price-ratios that give orientation to the observable budget plane passing through the observable optimum point. As behaviourists we have not yet earned the right to speak of “indifference” and “preference”; and we certainly have no right to speak of “indifference directions for infinitesimal or small movements”, especially since the underlined words are by no means unambiguous or mutually equivalent.

But no one can stop us from asking a purely mathematical question: “Can we ‘join together’ the little thumb-tacks to form bowl-like surfaces?” This involves the purely mathematical properties of the partial differential equations

$$-\frac{\partial x_3}{\partial x_1} = B^1(x_1, x_2, x_3), \quad -\frac{\partial x_3}{\partial x_2} = B^2(x_1, x_2, x_3)$$

and of the so-called “total differential expression”

$$B^1(x_1, x_2, x_3)dx_1 + B^2(x_1, x_2, x_3)dx_2 + 1\, dx_3.$$ Can we define unique bowl-like surfaces by setting the last expression equal to zero? The mathematicians know the answer to this,\(^2\) but the meaning of their integrability conditions will not be understood until we investigate the geometrical picture further.

Few people can visualise well in three dimensions. And fewer still can correctly see what must be true about things

\(^1\) We cannot in all cases invert our demand functions to express $p's$ in terms of $x's$. But if appropriate “curvature assumptions” are made, this will be possible within certain regions. See Mathematical Notes below.

\(^2\) See Mathematical Notes below.
infinitesimally close to each other. Therefore, few will be able to visualise the answer to the question of whether the thumb-tacks can be joined together to form surfaces. Let us therefore try to bring the matter down to two dimensions by the following device. Let us see what the planar elements or thumb-tacks look like on any two-dimensional surface. Let us put a thin wafer of wax in our room, bent so as to represent any desired surface. Now put in our whole field of thumb-tacks. Some tacks will not touch the wax surface; but some will have to be pushed sideward into the wax so as to lie in their specified positions. Hence, our surface will be covered by the "tracks" of the planar elements.

These tracks on the wax surface will look like the little line segments of the earlier two-dimensional picture. This is because the wax surface itself has only two rather than three dimensions.

Figure 3a, for example, shows how the little price ratios look on a plane which represents $x_1$ held constant, but with $x_2$ and $x_3$ varying. The little slope elements can obviously from our earlier two-dimension discussion always be joined together by the reader's eye to form a family of contours (which at this stage it would be dangerous to call indifference curves). In our remaining figures we shall assume that all such two-dimensional slope elements have been integrated into smooth contours. Figure 3b shows contour lines observed on a budget plane of a guinea-pig called Jeremy; while 3c shows observations recorded by a scientist under similar conditions for a consumer called Gustav.¹

¹ Of course, Jeremy need not be a hedonist of the Bentham type: he is the 20th century stream-lined version who believes only in ordinal utility; and Gustav is more sophisticated than was the Cassel who rejected utility in favour of demand functions and nothing else, but was never fully aware of what he was thereby assuming or denying about empirical reality.
Note that at a point like A, where the planar thumb-tack is just tangent to our wax surface, we do not get a unique directional track, but instead get a dot with indeterminate direction. This remark will probably give the reader a hint as to how we might try to find indifference surfaces such as that pictured in Figure 2b. Why not cleverly bend a wax surface so that everywhere on it we get dots of indeterminate direction? Professor Georgescu-Roegen in the 1936 article cited earlier was the first to show that this can indeed be done in the case of a consumer like Jeremy, whose tracks on any budget plane around an optimum point are always closed ellipse-like contours as in Figure 3b. But it cannot be done for a consumer like Gustav who leaves spiral-like tracks on his budget planes as in Figure 3c. This, then, is the geometrical meaning of the non-integrable case. Thus, by careful enough examination of the tracings left by a consumer's demand functions, we can determine whether he is a Jeremy or a Gustav.¹

**So-called Open and Closed Cycles**

I shall return to the spiral-like contours left on the budget plane in the non-integrable case later and shall show how this provides us with an alternative proof to Mr. Houthakker's theorem that the "strong axiom of revealed preference" (whereby a consumer never reveals a contradiction of preference in any chain of index number comparisons) definitely rules out non-integrability. But first I should like to use our geometrical model to throw light on Pareto's problem of open and closed cycles.

We have seen that Pareto confused the paths used by the scientist for comparative-statics comparisons with that actually followed by a consumer in consuming his goods. He did something else, less objectionable, but that I should like to avoid. He tried to define numerical or cardinal indexes of utility along such a path; I on the other hand should like to remain in the objective realm of commodity space.²

With two goods there is no problem: every cycle is "closed" in the sense that Pareto can always find a utility index defined by an integral which returns to its old value

¹ See Mathematical Note 6 below.
² The present exposition, which was suggested to me by a reading of Houthakker's paper, was elegantly developed by Dr. Wold, loc. cit., I, pp. 104–17.
once we get back to A. But with three or more goods, it may turn out that cycles are "open": we pass from A to B and then back to A and our integral indicates to us that A is "better than itself".\footnote{See Mathematical Note 4 below.}

That this phenomenon is independent of all cardinal utility considerations can be shown by Figure 4. We examine the contour lines traced in the wax by the little price-ratio planar elements. We have three wax surfaces that form a kind of three-sided tower. (The old Flat-iron building in New York City was actually of this general shape.) One wall rises perpendicularly to the "street" $ab$; another rises perpendicularly to the "street" $bc$. The third wall which is the back of the building could not be seen at all except that I have cut away part of the top to make it visible.

Instead of windows, there appears on each of our walls contour lines. I hesitate to call them "indifference curves", but will do so, putting the expression in quotation marks. Let us start Gustav out at the point A. Let us hold $x_1$ constant and give him more $x_2$; to keep him on the same
contour, we must take away something of $x_3$. Therefore, we move Gustav along the dark contour from A to B. Now we hold $x_2$ constant and decrease $x_1$; to keep Gustav indifferent we must then compensate him by increasing $x_3$, moving him along the dark contour BC. Likewise, by increasing $x_1$, decreasing $x_2$ and increasing $x_3$, we move him along the back wall, along the contour through C all the way to A'.

Gustav is obviously a rather strange chap. We have never crossed an “indifference” contour and yet we end up at the same amounts of $x_1$ and $x_2$ as at the beginning, but with $x_3$ necessarily greater by the amount $AA'$! A man with such woolly preferences can be easily cheated. If Gustav would agree to our moving him anywhere along his “indifference” contours, we could take away $x_3$ from him by walking him around A'CBA, and we can continue the spiral downward until we have deprived him of indefinitely much $x_3$.

This could not happen to a Jeremy. For him, A' and A will exactly coincide: such is the meaning of the integrability condition that his demand function must obey.

But once again I must remind the reader that the paths followed, leading to open or closed cycles, have nothing to do with the process by which a consumer moves toward equilibrium or with the order of his personal consumption agenda.

Revealed Preference and Integrability

We are now in a position to complete the programme begun a dozen years ago of arriving at the full empirical implications for demand behaviour of the most general ordinal utility analysis. My own work in this direction grew out of a remark made to me by Professor Haberler in his 1936 international trade seminar at Harvard. “How do you know indifference curves are concave?” My quick retort was “Well, if they're not, your whole theory of index numbers is worthless”. Later I got to thinking about the implications of this answer (disregarding the fact that it

---

1 This argument is really illegitimate: Gustav, not being a Jeremy, would have no reason to agree that we are free to move him along his “indifference” contours. The word “indifference” is probably not in his vocabulary.

2 See Mathematical Note 5 below.
is not worded quite accurately). Being then full of Professor Leontief's analysis of indifference curves, I suddenly realised that we could dispense with almost all notions of utility: starting from a few logical axioms of demand consistency, I could derive the whole of the valid utility analysis as corollaries.

My fundamental axiom I borrowed from modern index-number theory. I shall call it (for reasons that will be obvious) the Weak Axiom of Consumer's Behaviour:

Weak axiom: If at the price and income of situation A you could have bought the goods actually bought at a different point B and if you actually chose not to, then A is defined to be "revealed to be better than" B. The basic postulate is that B is never to reveal itself to be also "better than" A.

I soon realised that this could carry us almost all the way along the path of providing new foundations for utility theory. But not quite all the way. The problem of integrability, it soon became obvious, could not yield to this weak axiom alone. I held up publication on the conjecture that if the axiom were strengthened to exclude non-contradictions of revealed preference for a chain of three or more situations, then non-integrability could indeed be excluded. At scientific meetings and in correspondence this problem was proposed, both to economists and mathematicians. But no proof was forthcoming for all these years, until Mr. Houthakker's paper arrived in the daily mail. Not only had he provided the missing proof, but in addition he had independently arrived at precisely the same strong axiom as I had hoped would save the day. To his paper which is such a model of logical elegance and compactness, the present historical treatment of the subject can hope to serve only as a supplement.

In words, the strong axiom can be worded roughly as follows:

Strong axiom: If A reveals itself to be "better than" B, and if B reveals itself to be "better than" C, and if C reveals itself to be "better than" D, etc., then I extend the definition of "revealed preference" and say that A can be defined to be "revealed to be better

---

1 For example, in a single step, one arrives at most of the properties of the Slutsky S matrix in the Mathematical Notes below.
than” Z, the last in the chain. In such cases it is postulated that Z must never also be revealed to be better than A.¹

What Mr. Houthakker has shown is the following: Any Gustav, whose demand functions correspond to a non-integrable preference field, can be made to contradict the strong axiom. By judiciously varying prices and income, we can make him (on our strengthened definition of revealed preference) assert that A is better than Z and Z is at the same time better than A. Houthakker has shown how we should pick prices and income so as to reveal this contradiction. He has given us the long-sought test for integrability that can be formed in finite index-number terms, without need to estimate partial derivatives.

Figure 4 shows the same thing, but not so elegantly in that quantities and not prices are the independent variables. As Houthakker, Little and I have shown, we can pick a sequence of points very near to the path ABCA' but lying just inferior to it, along which the individual is asserting himself to be getting worse off. We finally end up just below A', but definitely above A—and it is obvious at this point that our consumer has revealed himself to have become “better off” since A' and anything near it costs more than A. Here, we have a contradiction: the point between A and A' is revealed to be both better and worse than A as a result of assuming non-integrability. Therefore, every cycle must be closed, and integrability is assured by our strong axiom. This is a variant of Houthakker’s proof.

A second variant uses the spirals shown in Figure 3c. Just as there passed an optimal budget plane through A, so too there will pass an optimal budget plane through the point B. This will show up on the original budget plane of A as a line, tangent to B. Any point, like C which lies beyond this line on the side opposite to A, will lie inside the new budget plane and will reveal itself to cost less than B and thus be inferior to B. But we can find a sequence of points (such as C, D, E, . . . etc.) which hug near to the spiral

¹ I have glossed over a few delicate points. Thus, not all of the situations have to be different ones. Also, there is the question of how to handle two situations which at A’s prices cost the same, but where A was chosen. If we rule out the possibility that B is “indifferent to A”, then we rule out some realistic cases of multiple equilibrium points. Even when B costs less than A at A’s prices, in ruling that A is “better than B” we are making some implicit assumptions about the absence of “saturation-effects”. Throughout this paper I have dodged delicate problems of this sort, taking refuge in overly-strong assumptions about regularity of curvature. Since all this needs a definitive treatment, I must apologise for glossing over the difficulties.
through C but which march clockwise all the way around to just to the right of A and which always reveal themselves to be "worse than" C. But it is also clear that all points near to A can reveal themselves to be "better than" B. Thus, we have a point near A revealed to be both better and worse than C. This is our desired contradiction, and again we end up with the inadmissibility of non-integrability.

For purposes of identification, let us call a man a Paul if he satisfies the weak axiom for single-pair comparisons, whether or not he satisfies the strong axiom for any chain. Without going into next section's arguments concerning the merits of integrability v. non-integrability, this much seems clear to me: if a man is not a Jeremy with an integrable preference field, that is understandable; but why should anyone be a semi-irrational Paul without going all the way and being a Gustav? Or alternatively, why should an irrational Gustav make a bow in the direction of being consistent and stop short at being a Paul, refusing to take the final step toward becoming a Jeremy? In short, I should expect there to be few Pauls reported by experienced observers.

**Summary of Arguments for Non-Integrability**

Let us now use the above analysis to appraise briefly the main arguments that have been advanced for non-integrability.

(1) There is the celebrated argument of Pareto, already mentioned, concerning the order of consumption and the question of open and closed cycles. Enough has been said already to indicate that I agree with the majority of writers that this is a confused discussion in at least the following respects: (a) comparative statics has little to do with the behind-the-scenes order of consuming; (b) even if the individual is presumed to follow the optimal order of consumption behind the scenes, there still remains the problem of integrability; (c) all of the problems raised in connection with the order of consumption arise just as essentially in the two-variable case where there is no problem of integrability; (d) there is the realistic possibility that the approach

---

1 Thus, in G. C. Evans: *op. cit.*, p. 121, the example of the carpenter who is not indifferent as to the order in which he renders services, receives wages, and consumes commodities would make as much (or little) sense if formulated in terms of two variables instead of three. On p. 120, this same author adds a new confusion: the path along which the line integral between two points is evaluated is called the path along which [market] exchanges take place.
to equilibrium may alter the final equilibrium point reached, because of irreversible *hysteresis* effects; but again, this has nothing to do with the integrability problem.

(2) The second main argument for non-integrability involves the question of whether an individual may be able to form preferences in connection with so-called "small changes" even though he is incapable of making consistent choice judgments about far-away situations. There is indeed an air of realism about this observation, but as far as integrability is concerned this is also a complete red herring based upon logical confusion—at least such is my personal judgment after a careful re-reading of all the relevant literature. It is true that one cannot accurately predict how he would spend his income quantitatively if he became fabulously wealthy or extremely poor; but the same would no doubt still be true if he consumed only two goods, and here the problem of integrability cannot arise.¹

The most devastating weakness in this view is its failure to mean by *small* what the mathematician should mean by small in this context—namely, *finite* movements not greater than some specifiable amounts. The instant the movements become finite, however small, the integrability conditions (which are themselves of a "local character") must become applicable or we cannot speak of local preferences at all. If we try to get out of this dilemma by interpreting small to mean infinitesimally small in a rigorously defined sense, then the argument loses all its flavour of empirical realism.

I conclude this second issue by adding two red herrings of my own: (a) Georgescu-Roegen, Weber-Fechner, and others tell us that there is a "threshold effect" of perception which surrounds every situation in much the same way that static friction follows a particle on a rough table; as Wicksell and others have recognised, this means that it is precisely for very *small* changes that the individual cannot have consistent preferences. (b) Since in the real as compared

¹ With hesitation, I question the opinions on this point of two such distinguished authorities as G. C. Evans and R. G. D. Allen. The former (loc. cit.) speaks of "infinitesimal loci of indifference" (p. 119) and of "small changes \( dx_1, \ldots dx_n \ldots \) being made" (p. 122), and of an "approximate value function for small changes [which cannot, without extra assumptions, be extended] \ldots beyond a merely local field". In the cited 1937 publication, Allen assumes "that the individual has a scale of preferences for small changes from a given set of purchases \ldots but this assumption \ldots does not imply that a complete scale of preferences exists. The consumer \ldots need not be able to discriminate between widely different sets of purchases" (pp. 440–1).
to the idealised world, all goods are ultimately not divisible beyond certain quanta, what happens to the non-integrability argument based upon the rigorous interpretation of infinitesimals? Leaving all levity aside, I must agree that it is only realistic, in the two-good or n-good case, to distinguish with Marshall between short-run and long-run preference behaviour—and for that matter between ex ante conjecture behaviour and ex post actual behaviour.

(c) There is the argument made ever since the time of Fisher, that non-integrability is the more general case of which integrable utility is but a special case. Since I have myself in the past stressed this point, I should like to make clear the sense in which I now think it valid. Generality pursued too avidly leads to emptiness. As scientists we must be willing to live dangerously. What we must seek is no inadmissible specialisations and no unnecessary generality. Newton did well to speak of the "inverse-square law of gravitation" rather than the "n"-power law"; at the same time, modern physicists do well to modify his law in the field of high-speed particles in favour of the more general Einstein relativity transformations.

As I now look at the matter, why should a Gustav, who has no mind to make up, so to speak, be expected not to be changing his demand behaviour constantly and capriciously? In other words, where a man does not exhibit the behaviouristic traces of Jeremy-like consistent ordinal preferences, why should his demand have any time invariance?

At least two answers might be given to this implicit defence of integrability as the only interesting hypothesis worth making. First, a man might display consistent demand behaviour through habit or crude rules-of-thumb not consistent with an ordinal preference field. Second, what is a man? Or a consumer? I am not so much concerned with the problem of Dr. Jekyll and Mr. Hyde, but with the problem of Dr. Jekyll and Mrs. Jekyll. Much consumption behaviour is family rather than individual behaviour. Now a family must be quite sophisticated indeed to end up with a consistent set of collective preferences: e.g., if they set up the rule that the wife will always spend 99 per cent (or 50 per cent.) of the income on her needs and the husband 1 per cent. (or 50 per cent.) on his quite different needs, this will not be consistent with an integrable set of
price-ratio elements. Only if the family acts in terms of a Bergson Social Welfare Function will this condition result. But to explain this farther would take me into the frontier of research in welfare economics.\(^1\)

(4) A last argument might be built up against non-integrability: if people lack the consistency of behaviour that integrability implies, then that attractive branch of individualistic welfare economics which says people’s tastes should count loses most of its content; hence, we should rule out non-integrability. I am afraid that this is an illegitimate intrusion of wishful thinking by the would-be political philosopher into the facts of life. If people do not behave as if it matters to them just what they consume, that is a weakness (but not necessarily a fatal one) for the Pareto-type compensationist new-welfare economics.\(^2\) We must not bias our view of the facts to fit our wishes and prejudices, however pretty their pattern. On the other hand if integrability should turn out to be the best hypothesis to explain the empirical facts of the market-place, this makes a belief in individualistic ethics possible but still not mandatory. In the last analysis the definition of a welfare function is not an empirical problem.

In this connection it is interesting to note that Professor D. H. Robertson has put forward in his recent presidential address the view that Pareto and subsequent theorists were led through their fancy for fancy methods into a fanciful wild-goose chase in which they dethroned utility by the verbal act of giving new names to old terms—but that in recent years, Hicks and other devotees of the black arts of mathematical economics have completed the cycle and substantially returned to the common-sense orthodox utility views held all along (in quiet) by cooler heads, according to which utility is needed for welfare economics. I am not confident that I have correctly interpreted Robertson’s meaning since his remarks were necessarily brief and purposely couched in charming ellipsis. But on the assumption that this is close to his intended meaning and that this one-fifth of his address was intended to be taken seriously, I must record some personal puzzlement.

In re-reading the cited thirty-odd basic papers of the last 60 years, I have not been able to discern trends of

---

\(^1\) Cf. *Foundations of Economic Analysis* (1947), Ch. 8.

development along the lines of the propounded thesis. Nor am I aware that Hicks has anywhere in print reverted to a pre-Paretian position on utility; on the contrary, in 1939 Hicks independently arrived at an almost too-Paretian view of welfare economics. And other modern theorists (such as Lange, who once expressed the view that cardinal utility was unnecessary for demand studies but needed for welfare economics) were easily converted by Professor Bergson to the view that only ordinal comparisons are relevant.

However, Robertson can be interpreted to mean that the quest by Pareto and later writers for non-integrable utility was both factually unnecessary and destructive to much in welfare economics. Such a viewpoint on this more specialised and technical matter has its merits, and the present paper has the humble purpose of clarifying some of the issues.

We have now completed our main task. We know what it is that integrability implies, and what non-integrability implies: we know the full empirical implications on the demand functions of being a Jeremy or a Gustav. Deductive analysis can carry us no farther. Observation of reality must be the decisive test as to which hypothesis is the more fruitful—or whether neither is very fruitful.

MATHEMATICAL NOTES ON INTEGRABILITY

1. The observable $n$ demand functions are assumed to be single-valued and continuously differentiable in terms of prices and income, of the form $x_k = D^k(p_1, p_2, \ldots, p_n, I)$. Changing all prices and income in the same proportion will, by hypothesis, cause no change in real quantities; in consequence of this assumption of homogeneity of order zero, Euler’s theorem tells us that the matrix $[D^k_j + D^iD^k_j]$ has a zero determinant—it being understood that subscripts stand for differentiation so that $D^k_j = \partial x_k / \partial p_j$, etc.
I now make the special additional hypothesis that these demand functions are reversible and can be inverted to express relative prices in terms of quantities. This will be assured in any local region around a point where the determinant of \([D^k_i]\) is not zero. A symmetrical representation of relative prices could be provided by defining new variables \((Q_1, \ldots, Q_n) = (p_1/I, \ldots, p_n/I)\). But it is more traditional to set the price of some numéraire good, say \(x_n\), as \(I\), and to express income and all prices in terms of the \(n^{th}\) price; hence, I define \([B_1, \ldots, B_{n-1}, I] = [p_1/p_n, \ldots, p_{n-1}/p_n, I]\), and after reversing our demand functions, we have the observable price-ratio functions (or marginal rates of substitution)

\[(1) \quad B_k = B^k(x_1, x_2, \ldots, x_n) \quad (k = 1, 2, \ldots, n-1)\]

A similar function exists for \(I/p_n\), which we might call \(B_n = + B_{1}x_1 + \ldots + B_{n-1}x_{n-1} + x_n\) after noting that there is no need to waste this symbol on \(p_n/I\).

Our reversibility hypothesis can be expressed in terms of the Jacobian of the B’s being non-zero. Specifically, let us define the matrix \(\mathcal{J}_n\) to be the square matrix that results when we set \(p_n = I\) and exclude the \(n^{th}\) price-column from \([D^k_i; D^k_j]\). Then the interested reader can verify the following partitioned matrix identities.

\[(2) \quad \mathcal{J}_n^{-1} = [B^k_i] = \begin{bmatrix} B^k_{ij}, B^k_{in} \\ B^o_{ij}, B^n_{in} \end{bmatrix} = \begin{bmatrix} I, \circ \\ x_j, I \end{bmatrix} \begin{bmatrix} B^k_i, B^k_n \\ B^o_i, I \end{bmatrix} = \begin{bmatrix} I, \circ \\ x_j, I \end{bmatrix} \begin{bmatrix} B^k_i - B^o_i B^k_n, B^k_n \\ \circ , I \end{bmatrix} \begin{bmatrix} I, \circ \\ B_j, I \end{bmatrix}\]

where \(I\) has 1’s in its diagonals and \(\circ\)’s elsewhere and where the upper left-hand sub-matrix is always \((n-1)\) by \((n-1)\).

It is sufficient for reversibility in some small region around an observed point that the determinants of any one of these matrix products be non-zero. If we wish the B’s to be reversible even when we hold some subset of \(x\)’s constant, then the principal minors derived by striking out some of the first \((n-1)\) rows and columns must also be non-zero. The non-vanishing of all these principal minors throughout an extended domain will assure reversibility of our functions—not just locally—but throughout that domain.
2. So far nothing has been said about integrability conditions. Slutsky proved that if there does exist an integrable utility function \( U = F[V(x_1, \ldots, x_n)] \) so that the \( B \)'s can be written as the ratio of \( U_k/U_n \), then the matrix

\[
S = \begin{bmatrix}
D^k_{i} + D^lD^k_l, & D^k_n + D^nD^k_l \\
D^n_{i} + D^lD^n_l, & D^n_n + D^nD^n_l
\end{bmatrix} = \begin{bmatrix}
S_n, & s_{kn} \\
s_{nj}, & s_{nn}
\end{bmatrix}
\]

is symmetric, in consequence of the symmetry of the cross-derivatives, \( U_{kj} = U_{jk} \). This is a singular matrix in virtue of the Euler-theorem relation \( S\varphi = 0 \); hence, the Slutsky integrability conditions involve only \((n-1)(n-2)/2\) independent conditions: e.g., the symmetry of the upper-left matrix, \( S_n \), implies full symmetry.

Slutsky also showed that the elements of \( S \) are observable empirical invariant quantities, not dependent on the form of the arbitrary numerical indexes of utility. He interpreted the elements of \( S \) as compensated changes, for which ordinal utility is constant because the change in price is just offset by a change in money income. Slutsky showed that the diagonal or "own-elasticity" elements of \( S \) must be negative; the off-diagonal elements provide the Slutsky-Schultz-Hicks definition of complementarity, which is of greatest qualitative richness only in the \( n > 2 \) case. It may be added that \(-S\) is positive semi-definite, being of rank \( n-1 \), and with its principal minors being all positive. \( S \) constitutes the most important part of consumption theory.

Not only are the Slutsky integrability conditions necessary, they can be shown as well to be sufficient to insure integrability. By somewhat tedious manipulation, we can show that

\[
\begin{bmatrix}
D^k_{i} & D^k_n \\
D^n_{i} & D^n_n
\end{bmatrix} = J_n \begin{bmatrix}
I & 0 \\
x_j & I
\end{bmatrix} \text{ or from (2)}
\]

\[
= \begin{bmatrix}
I & 0 \\
B^i_j & I
\end{bmatrix}^{-1} \begin{bmatrix}
B^k_{i} - B^iB^k_n, & B^k_n \\
0, & I
\end{bmatrix}^{-1}
\]

\[
= \begin{bmatrix}
(B^k_{i} - B^iB^k_n)^{-1}, & \varphi_k \\
Z_j, & W
\end{bmatrix}
\]

where it is unnecessary to write out the expressions for the \( W \), \( \varphi \) and \( Z \)’s.

If the Slutsky sub-matrix \( S_n \) is symmetric, then \( (B^k_{i} - B^iB^k_n)^{-1} \) and \( (B^k_{i} - B^iB^k_n) \) are symmetric. The symmetry of the last-written matrix we might call the Antonelli-Hicks-Allen conditions, since the former announced them in 1886 and the last two gave an interpretation of
them in terms of symmetry of the Hicks-Allen (1934) defined measures of complementarity.¹

3. These last conditions can now be related to the mathematicians' discussion of integrability. We associate with (1) the partial differential equations

\[- \frac{\partial x_n}{\partial x_j} = B'(x_1, \ldots, x_n) \quad (j = 1, 2, \ldots, n - 1),\]

and the total differential equation

\[B^1 dx_1 + B^2 dx_2 + \ldots + B^{n-1} dx_{n-1} + 1 dx_n = 0.\]

Any path in space along which the latter is satisfied might be termed an "indifference path", but with emphasis on the quotation marks; along such a path, we can replace the \(B^i\)'s by any set of proportional functions

\[Q^1 dx_1 + Q^2 dx_2 + \ldots + Q^{n-1} dx_{n-1} + Q^n dx_n = 0\]

where \(Q^n\) is an arbitrary non-vanishing function and \(Q^i = B^i Q^n\).

An expression like \(\Sigma Q^k dx_k\) may be satisfactory as a local measure of preference, being positive when a "better" move is being made, being zero when there is "indifference", and negative when a "worse" move is being made. But its integral along any finite path between two points will generally not provide a consistent numerical indicator of ordinal preference—even if such a consistent preference field exists, and certainly not if such a field does not exist. Only if \(\Sigma Q^k dx_k = dQ\), an exact differential with \(Q^k = \partial Q/\partial x_k = Q_k\) and \(dQ^{ik} = \dot{Q}^{ik}\), will such an integral be always the same for different paths between two specified end-points; and only in this case can we be sure that if we go from point \(A\) to point \(B\) by one path and return back to \(A\) by another path, the value of the integral will none the less be zero over the round trip, indicating that \(A\) is exactly as good as itself and not better.²

¹ In 1959 Hicks seems to have abandoned this definition in favour of the Slutsky-Schultz definitions. For \(n = 3\), the results of either definition are qualitatively the same. For \(n > 3\), this is not true. If we define all but the two goods in question to be a Hicksian composite commodity, then the Slutsky-Hicks definition can be cast in Hicks-Allen terminology. From this the specialist will gather that I no longer think that earlier criticisms of logical inconsistency between the text and appendix of Value and Capital are valid.

² The line-integral \(\int [\Sigma Q^k dx_k] dt\) along a path between two points \(a\) and \(b\) is defined as the simple integral \(\int [\Sigma Q^k dx_k(t)] dt\) where \(x_k = x_k(t)\) is the equation of the path, and where the appropriate limits for \(t\) are understood. See, for example, E. B. Wilson: Advanced Calculus, Ch. XI and also Ch. X on partial differential equations.
The above condition on the cross-derivatives will rarely be met; but we can ask whether there is not some other set of functions proportional to the $B$’s for which we do get an exact differential. For this to exist, we must be able to find a so-called integrating factor $i(x_1,x_2,\ldots,x_n)$ such that

$$iQ^1dx_1 + \ldots + iQ^ndx_n = dV$$

an exact differential.

This is possible only if $V_{ki} = V_{jk}$, or if we have $n(n-1)/2$

\begin{equation}
(4)\quad o = \frac{\partial (iQ^k)}{\partial x_j} - \frac{\partial (iQ^j)}{\partial x_k} = i_jQ^k - i_kQ^j + i(Q^k_j - Q^j_k) \quad j \neq k.
\end{equation}

When we examine these conditions, they are seen not to be all independent. Also they are not conditions on the $Q$’s alone but involve $i$ and its partial derivatives, where $i$ is a function not even known to exist. For any triplet of variables, say $j, k, r$, we can eliminate $i$ and its derivatives by algebraic manipulation, to get the following relations on the $Q$’s alone

\begin{equation}
(4)’\quad o = Q^k(Q^j_r - Q^r_j) + Q^r(Q^k_r - Q^k_j) + Q^j(Q^k_r - Q^j_r) \quad , j \neq k \neq r \neq j.
\end{equation}

When $n = 2$, there are no integrability conditions; when $n = 3$, there is 1; when $n = 4$, we can write down 4 such conditions; and, in general, $n(n-1)(n-2)/6$ equations. But these are easily seen not to be all independent; at most $(n-1)(n-2)/2$ are independent and all the rest follow by algebraic manipulation.

A convenient way to get a set of independent conditions is to hold $r = n$, and work with $B^k = Q_{ki}/Q_n$. Equation (4)’ then becomes

\begin{equation}
(4)’’\quad B^k_j - B^j_k = B^k_j - B^j_n. \quad n > k \neq j < n.
\end{equation}

But these are the Antonelli integrability conditions, already shown to be equivalent to the Slutsky symmetry conditions.¹

4. It has been fairly easy to show that the above conditions are necessary if there is to be a family of integrable surfaces. They are also sufficient to guarantee the existence of such surfaces; but this is not so easy to show. We may, of course, always accept the guarantee of the pure mathematician that there is a general existence theorem on partial

¹ Jevons, Walras, and Marshall did not have to be concerned with the integrability problem primarily because they implicitly assumed it away, and because they assumed that the marginal utility of any good depends on its own quantity alone. In this case (4), (4)’, (4)’’ are always satisfied.
differential equations assuring us that the above equations do define a family of integrating factors, \( i \).

However, it may add to our intuitive grasp of the problem if I hastily sketch a method of actually constructing a single indifference surface. Let us start from any given point \((x_1^0, x_2^0, \ldots, x_n^0)\). Let us now make all but the last variable \((x_n \text{ or } x_3 \text{ as the case may be})\) move in a fixed ratio to moves in \(x_1\); as we move \(x_1\) from our beginning point, then \(x_2 \ldots\) will move in this determined fashion; and now we require the last variable, \(x_n\), to move as determined by

\[
B^1 dx_1 + B^2 dx_2 + \ldots + 1dx_n = 0
\]

where

\[
(x_k - x_k^0) = y_k(x_1 - x_1^0) \text{ or } X_k = y_k X_1 \quad (k = 2, \ldots, n - 1)
\]

and where the \(y\)'s are arbitrary proportionality constants.

In the three-dimensional case, we are swinging a vertical door whose vertical axis of hinging goes through the original point. When the door is at any angle, there is on its surface a unique "indifference" contour through the original point. If now we let the door swing through all angles, these indifference contours will sweep out a two-dimensional surface.

Analytically, for given \(y\)'s, we can eliminate all the \(x\)'s but \(x_1\) and \(x_n\), and end up with a differential equation for \(x_1\) and \(x_n\). This equation always has one solution through the initial point which we may write as follows

\[
X_n = g(X_1; \ y_2, \ldots, y_{n-1})
\]

where the \(y\)'s are parameters. Now if we let the \(y\)'s vary and recall the relation \(X_k = y_k X_1\), we have defined a hypersurface

\[
(5) \quad X_n = g(X_1; \ X_2/X_1, \ldots, X_{n-1}/X_1)
\]

which of course goes through the original point and can be written in terms of the original small \(x\)'s.

But will any movement along this relation satisfy \(B^1 dx_1 + \ldots + dx_n = 0\)? By tedious substitution, we find that the answer is generally no—unless the integrability conditions are satisfied, in which case the answer happens to be yes. Only in this case can we term our surface an indifference surface; and only in this case shall we arrive back at the same surface if we repeat the process starting out from any other point satisfying (5). For a description of this process, called Mayer's method, see any mathematical text such as Wilson's *Advanced Calculus*, loc. cit.
5. By similar reasoning, we can show that if the integrability conditions are not satisfied, there will always be "open cycles" as shown in Figure 4. Suppose the indifference movements there described, such as from A to B, always lie on a definite surface which can have but one value for \( x_n \) in terms of the other \( x^i \)'s. Call this surface

\[
. x_n = G(x_1, x_2, \ldots, x_{n-1})
\]

where of course

\[
-\frac{\partial G}{\partial x_k} = B^k[x_1, x_2, \ldots, G(x_1, x_2, \ldots x_{n-1})].
\]

Then

\[
-dx_n = B^1[x_1, x_2, \ldots, G]dx_1 + B^2[x_1, x_2, \ldots, G]dx_2 + \ldots
\]

would have to be an exact differential in \((x_1, \ldots, x_{n-1})\); hence, it is necessary that the cross-relation hold

\[
(4)'' \quad B^*_{ij} + B^k_nG_j = B^k_i - B^k_nB^i - B^i_nB^k = B^i_k + B^i_nG_k.
\]

If these integrability conditions were not satisfied, we should have a contradiction to our assumption of an exact differential defining a surface yielding only closed cycles.

6. The Georgescu-Roegen demonstration that the contours on the budget-line planes of 3b and 3c are ellipses in the integrable case and spirals in the non-integrable case proceeds as follows. Setting all the \( x^i \)'s constant but \( x_1, x_2, \) and \( x_3 \) or more generally \( x_n \), then our budget plane is defined by

\[
B_1^0x_1 + B_2^0x_2 + x_n = \text{constant} = C
\]

\[
B_1^0dx_1 + B_2^0dx_2 + dx_n = 0
\]

where \( B_k^0 \) stands for the given constant prices; at the optimum point \((x_1^0, x_2^0, \ldots, x_n^0)\), of course, \( B_k^0 = B_k^k(x_1^0, \ldots, x_n^0)\). The "indifference" contours on the budget plane must, after substitution, satisfy

\[
B^1(x_1, \ldots, C - B_1^0x_1 - B_2^0x_2)dx_1 + B^2(x_1, \ldots, C - B_1^0x_1 - B_2^0x_2) - B_1^0dx_1 - B_2^0dx_2 = 0
\]

or

\[
W^1(X_1, X_2)dx_1 + W^2(X_1, X_2)dx_2 = 0
\]

where \( X_k = x_k - x_k^0, W_k(x_1 - x_1^0, x_2 - x_2^0) = B^k(x_1, \ldots, C - B_1^0x_1 - B_2^0x_2) \)

\[
W^k_j = \frac{\partial W^k(0, 0)}{\partial x_j} = B^k_j - B^k_nB^i_j = B^k_j - B^k_nB^i_j
\]

\( (j, k = 1, 2) \)
where it is understood that all partial derivatives are to be evaluated at the equilibrium point.

In the neighbourhood of the equilibrium point, the slope of the contour is determined by

\[(W^1_1 X_1 + W^1_2 X_2 + \ldots) dX_1 + (W^2_1 X_1 + W^2_2 X_2^2 + \ldots) dX_2 = 0\]

with higher powers of the \(X\)'s neglected. If \(W^1_2 = W^2_1\), which in terms of the \(B\)'s expresses our old integrability condition, this is an exact differential as it stands; we immediately verify that

\[W^1_1(X_1)^2 + 2W^1_2(X_1 X_2) + W^2_2(X_2)^2 = \text{constant}\]

represents the contours of the budget plane. If the 2 by 2 determinant of \([W^k_1]\), which is essentially a principal minor in equation (2), is positive, we have concentric closed ellipses surrounding an extremum or vortex point as in Figure 3b. If the determinant is negative, we have a family of hyperbolae surrounding a saddle point. If ordinal utility is at a maximum, the last case is ruled out.

If \(W^1_2 = W^2_1\), then we shall have an infinite number of contours running into the singular point. If the above determinant is positive and the asymmetry limited in amount so as to satisfy the Allen (1937) generalised law of diminishing marginal rate of substitution, or my weak axiom, then

\[\sigma > \sum_{k=1}^{n-1} dB^k dx_k + \sum_{k=1}^{n-1} B^k dx_k = 0\]

and we shall have a spiral point as shown in Figure 3c. Another less interesting possibility will be a nodal point in which all the contours, orthogonal to the above case, run into the singular point from one or two directions. All of these cases are determined by the nature of the two roots of the characteristic equation of the \(W\) determinant, \(m^2 + (W^1_2 - W^2_1)m + |W| = 0\). The economically interesting cases are when the middle term vanishes so as to yield pure imaginary roots, or does not vanish so as to yield complex numbers; the real parts of the latter serve to damp down the closed elliptical sinusoidal motions into converging spirals. See L. Ford: Differential Equations, pp. 48–52.

7. There remains the troublesome question of what are the appropriate curvature (often miscalled "stability") conditions to be assumed in the non-integrable case. In the integrable case where the consumer is at an interior
maximum point with the $x$'s continuous and reversible functions of relative prices and income, the matrix

$$-\left[ W^k_j \right] = -\left[ B^k_j - B^k_n B^i \right] = -\left[ \frac{\partial^2 x_n}{\partial x_j \partial x_k} \right] \text{ comp.}$$

must be positive definite, which in terms of (2) implies

$$-\left| \begin{array}{cc} B^1_1 - B^1_3 \\ -B^1 - I \end{array} \right| > 0, \quad -\left| \begin{array}{ccc} B^1_1 - B^1_2 - B^1_3 \\ -B^2_1 - B^2_2 - B^2_3 \\ -B^1 - B^2 - I \end{array} \right| > 0, \text{ etc.}$$

But when $W^k_j \neq W^i_k$, these last conditions are not equivalent to $-\left[ W^k_2 \right]$ being the coefficients of a positive definite quadratic form with

$$-W^1_1 > 0, \quad \left| \begin{array}{ccc} -W^1_1 \\ -W^2_1 + W^1_2 \\ 2 \end{array} \right| > 0, \quad \left| \begin{array}{ccc} -W^2_2 \\ -W^1_1 + W^1_2 \\ 2 \end{array} \right| > 0, \text{ etc.}$$

This was shown by Georgescu-Roegen, and in 1937 Allen pointed out how his 1934 conditions (8) should be strengthened to make (9) hold in the asymmetric case. Allen obviously prefers (9) to (8), because he postulates the generalised law of diminishing marginal rate of substitution indicated a few paragraphs ago.

But why should such a law hold for a Gustav who has no consistent scale of tastes? Allen's argument seems to be postulational rather than persuasive, its only intuitive plausibility residing in the case of the integrable Jeremy. If all we demand is reversibility of demand—and I myself do not even demand that—then it is not clear but that the less "fully-developed" 1934 conditions are better than the 1937 conditions, even though the 1934 conditions are not invariant under linear commodity transformations.¹

I am also a little puzzled by Wold's Theorem X, op. cit., III, pp. 78–88. When this was cited in Mr. Houthakker's footnote as demonstrating that the assumptions of (1) homogeneous and (2) reversible demand functions were inconsistent with (3) non-integrability, I was aghast and ran to read Wold's paper—for such a result seemed miraculous. Interpreted literally, that theorem does seem to say that (1), (2), and (3) are "self-contradictory". But this reading may arise from the English wording of the Swedish, for it

is my interpretation from the proof that Wold really means to say that any consumer who obeys (1), (2), and (3) is behaving in a “self-contradictory” fashion from the standpoint of consistent ordinal preference. This is completely in line with what we should expect from non-integrability. When a scientist assumes (1), (2), and (3) it is not implied that he is being “self-contradictory”. I hope and trust that Wold and I would be in agreement that reversibility as such has nothing essential to do with integrability.

*Massachusetts Institute of Technology,*
*Cambridge, Massachusetts.*