The law of perceptual stability: Well-definedness and validity*
(perception/vision/behavior/cognition/grammar)

LEON G. SHIMAN

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Communicated by Richard Held, January 9, 1978

ABSTRACT Axiomatic foundations for a theory of perception have been given in a prior communication [Shiman, L. G. (1978) Proc. Natl. Acad. Sci. USA 75, 2040–2053]. The theory is based on mathematically characterized stable images. Stable images model subjective (psychological) states of observers. We showed there is a universal procedure by which such stable images can be assigned to objects of perception (icons), employing observer-dependent perceptual judgements.

We construct an abstract model for an icon from the mathematical structure of a stable image that has been assigned to it. We identify the model with the icon. By identifying the model with the object of perception, we can analyze perception of the icon in terms of the mathematical properties of its associated stable images. Thus subjective phenomena, in this mathematical form, can be used to analyze the mechanisms of sensory processes.

Grounds for validity of the stability law are given. The procedures are demonstrably universal. A stable image, although subjectively derived from an icon, is a mathematical entity, and has therefore objective status. Stable images give a natural and potentially complete (mathematical) classification of icons, along nontraditional lines, which includes a classification and account of many well-known "visual illusions."

Extension of this (static) theory to include time-dependent (kinetic) cases is outlined. We consider grammar for vision, semantics, and the role of stability theory in constructing a unified theory of vision and natural languages.

4. Well-definedness

4.1. In a prior communication (1) we developed the abstract foundations for the law of perceptual stability. Here we will consider its well-definedness and validity. We have already seen that the theoretical form of this law gives a mathematical characterization of an individual’s judgment of stability of his field of vision. The mathematical and methodological framework in which the stability law is defined is stability theory. The validity of stability theory depends on the validity of the stability law. The form and scope of the theoretical setting, and its intimate relationship to the stability law and the judgments on which the law rests, led to the conjecture that the law of perceptual stability expresses a natural law that governs the sensory dependency of a biological system and its environment. The set of postulates for a theory of perception, in which the theory has been developed, reflects this conviction. Support for this conjecture is drawn from properties of the mathematics developed in 2 and their practical consequences, which we consider here.

We have shown that a stable image (F) is derived from an icon (I) by perceptual judgments. (A stable image is an abstract structure which we regard as a state of a biological system. A state of the sensory surround of that system is an icon) (see ref. 1.) The procedure we used to derive F from I depended only on the mathematical representations of three perceptual judgments. The procedure is self-evidently universal with respect to observers and icons and does not depend in any way whatsoever on prior descriptions of the icon. This means that F is a model for I that is dependent on exactly those characteristics of I to which the particular perceptual judgments rendered of I, by which F was derived, were sensitive.

Icons have traditionally been described in terms of their color, texture, reflectance, geometry, etc. Studies of the mechanisms of vision have traditionally been carried out relative to these formal characteristics of icons, and physiological accounts of these mechanisms have been sought in the same framework.

These methods are regularly employed in electrophysiology, psychophysics, and computation-theoretic studies of vision.

Such traditional methods for studying perception rely on a priori characterizations of icons; the descriptions of icons on which they are based are not necessarily natural to perception. In the framework of stability theory, we can use the intrinsic structures of images, relative to specific perceptual judgments, to analyze the perception of icons, and thus to discover exactly how properties of icons affect what an observer can perceive.

4.2. An abstract model for an icon (I) can be constructed from an image that I supports. This is done in the following way: If (i) I is an icon, (ii) E is the Euclidean plane, (iii) D is the domain of a stable image that I supports, and (iv) π is a stable homeomorphism of D into E, then π(D) is a model for I in E.

The stable homeomorphism π acts functorially on the set of stable decompositions of D, assigning to each stable decomposition S(D) of D a stable decomposition S(π(D)) of π(D). Thus, if (i) F is a stable state of I and (ii) S(D) is the stable decomposition of D by F, then S(π(D)) is a representation of F in the model π(D) of I.

For ease of exposition we abuse notation and write π(S(D)) for S(π(D)).

4.3. This abstract model for an icon can be identified with an object of perception. In the simple model we have constructed for I, E can be identified with the (plane of) the surface of I and π(D) can be identified with the spatial extension (domain) of I. (It thus becomes meaningful for us to say that, relative to π, "F breaks I into pieces."

* This is the second of two parts; the first part of this work, which contains sections 1–3, is ref. 1.
* For example, the search for systematic organic or gross behavioral response to pattern geometry, brightness, speed, and shape of pattern movement is now widespread, for many different species.
* As by Kuffler (2) and many others. We assess neither validity nor interpretation of their results.
* In work that follows the basic program of Fechner, part I (3).
* Such as the work of Ginsburg (4), which is, at least in part, motivated by possible neurophysiological correlates of a method for spatial frequency analysis of patterns.
This practice of identifying abstract mathematical structures with physical events is familiar in physics. We do not hesitate to describe the motion of stars and planets in the mathematics of space-time, the phenomena of electricity and magnetism in the mathematics of electromagnetic fields, or atomic phenomena in the mathematics of quantum mechanics.

In an analogous way an icon is identified with its mathematical model in order to consider how it is related to an image it supports. It is a fact of perception that, for any observer, there is a lower bound on the resolving power of the visual system, i.e., there are things that cannot be perceived by the unaided eye. It thus follows, under identification of an icon with its model \( \pi(D) \), that there are (infinitely) many distinct representations of a stable state \( F \) that are visibly indistinguishable in the plane. To be precise, if \( \pi(S(D)) \) is a model for \( F \) in \( E \), then there are (infinitely) many models \( \pi_1(S(D)) \) of \( F \) in \( E \) that are equivalent to it, relative to all perceptually significant properties of \( I \).

We can also consider distinct representations \( \pi_1 \) in the icon plane \( E \) for each of the components \( (e_i, u_i) \) of \( F \). So that the set of \( \pi_1(u_i) \) and their corresponding boundary structures are arbitrarily related in \( E \). There are (infinitely) many such arrays in \( E \) that are equivalent to \( (S(D)) \) relative to all perceptually significant properties of \( I \).

However, the topological relationship of the set of components \( u_i \) of \( F \), and the orientation of the boundary arcs associated to each \( u_i \), for all \( i \), are unambiguously interpretable in \( I \). This can be verified by each observer. It is verified by applying the procedure given in 3 of ref. 1 for constructing a stable image of an icon.

As we have now seen, the map \( \pi \) characterizes abstractly a relation between a biological system and its sensory surround. The simple model that we have used for \( I \) is mathematically and procedurally consistent with the 3-fold characterization of an icon given in 3.2. Although the model is abstract, it is not constructed by "idealization," in the now-popular sense of approximation, and so we have sacrificed neither accuracy of observation nor exactness of correspondence of mathematical structures with (physical) events. Therefore, it is always possible to use this model for \( I \) to study the mechanics of visual processes. On these grounds, we say that stability theory is well-defined relative to basic mechanical conditions of perception.

4.4. We have constructed a model for an icon (\( I \)) from an image that \( I \) supports. An image is derived from \( I \) by perceptual judgments rendered of \( I \), where \( I \) is treated as an object of perception. The derivation can be stated in terms of our model for \( I \).

Consider a simple case. Suppose we are given an icon (\( I \)), in which a "disc" is perceived to lie on a "surface." Such an identification breaks \( I \) into two pieces. The two pieces represent the regions that are perceived, respectively, as "disc" and "surface." Identification of a region as a disc implies that a judgment of convexity can be rendered uniformly along the boundary of that region, and thus, according to the adjunction rule, that the entire boundary of the region is adjoined to it, i.e., that the orientation of the boundary of that region is uniformly inward. Identification of the complementary region (in the open decomposition of \( D \)) as a "surface," shows, when evaluated along the shared boundary, that the identity is independent of boundary conditions, and thus that the boundary is oriented uniformly outward. The two orientation assignments are compatible, and so this decomposition of \( D \) defines a stable state \( F \) for \( I \). By the stability hypothesis 3.6, we now conclude that there is a stable image of \( I \), which corresponds to \( F \), and so we identify that image with \( F \).

We can easily see that these perceptual judgments have a well-defined and consistent interpretation in the model for \( I \) in \( E \) described in 4.2. We can represent: (i) \textit{Stability}, by a global property of \( \pi(D) \). (ii) \textit{Identification}, by a property of a subset \( \pi(u) \) relative to its open complement \( \pi(D-u) \). (iii) \textit{Adjunction}, by the global consistency of the identification of the collection of subsets \( \pi(u) \) defined by \( F \) in \( \pi(D) \), although the adjunction rule is actually applied locally along each of the boundary arcs of each of the separately identified regions of \( I \).

This establishes a direct relationship of image and icon. But if the relationship is expressed in terms of the biology of vision, then it is necessarily subject to the limiting conditions on perceivability discussed in 4.3, in addition to the restrictions of the static case that were stated in 1.3.

4.5. As a consequence of the correspondence we have established between image and icon, we are now free to use the abstract properties of subjective states (whose fundamental structure we take to be the stable state) to analyze mechanical processes of perception. To do this, it is necessary to enrich the model we are using for \( I \), so that all intrinsic properties of \( I \), which affect the perception of \( I \), are representable. \( I \) can then be analyzed relative to the structures of images that \( I \) supports.

An image is not a simple function of easily described properties of an icon. It is, for example, easy to observe that region boundaries in a stable decomposition do not, in general, follow obvious pattern discontinuities in the icon. But thorough analysis of an icon relative to all the images it both can and cannot support must, in principle, yield a complete characterization of that icon which is natural with respect to human perception.

In general, for an icon \( I \) it is not possible to predict what an observer will see. A central feature of the construction of stability theory is that it formally incorporates this fundamental fact of perception. The stability hypothesis asserts that an icon has a stable state if it can be stably perceived, but it neither prescribes the state nor tells how to find an appropriate decomposition. Stability theory allows for individual differences in human perception, and provides a framework to give them an exact accounting.

5. Validity

5.1. We have shown in the preceding discussion that two basic conditions necessary for the validity of stability theory are satisfied: the abstract setting is well-defined, and its practical interpretation is unambiguous. We have also assumed the truth of the stability hypothesis. This hypothesis is the logical keystone of the theory. The grounds for our assumption are given in the four concluding sections of this paper. They briefly treat universality, classification of icons, kinetic extension of static theory, and biological interpretation of the abstract structures derived from perceptual judgments.

5.2. \textit{Universality} is the assertion that stability theory applies without exception to all observers and all icons independent of cultural bias and the physical attributes of an observer, and, with the possible exception of cases of hallucination and dreams, that the structure of an image has objective, observer-independent, validity.

The claim of universality rests on several self-evident features of the theory: (i) The three judgments on which the stability hypothesis depends can be understood and applied without ambiguity by all observers. (ii) A stable state associated by an observer with an icon is a purely mathematical structure. As such, any two stable states (having an objective status) can always be compared. (iii) Two observers can determine whether
two icons share a common tactile and luminous source (a common reference object). Abusing language, we sometimes identify icons with their common reference object. (iv) Two observers can always compare the stable states that they have individually associated to a common reference object.

A stable state \( F \) will be considered to be observer-independent, if: (i) For each of two or more observers \( O_i \), there is an icon \( I_i \) that \( F \) is mathematically equivalent to a stable state \( F_i \), which observer \( O_i \) can verify is supported by icon \( I_i \) for all \( i \).

5.3. Icons can be classified in three distinct ways: (i) relative to intrinsic properties of their abstract (mathematical) representations [3, 2, (i)]; (ii) by physical characteristics of particular reference objects [3, 2, (ii)]; and (iii) by mathematical properties of images [3, 6]. The third alternative gives a classification of icons in accord with how they can be perceived. Following this, we will describe briefly how icons can be classified by the stable states they support.

We assume that an icon \( I \) has been mathematically defined, so that we can speak of the class \( J \) of all icons. The class \( J \) is just a (complete) collection of mathematical objects of the same kind. We suppose further that for each icon \( I \) a set \( \mathcal{F}(I) \) of stable states \( F(I) \) is associated to it, in which the symbol \( \mathcal{F} \) stands for the class of all abstract stable states. It is obvious that \( \mathcal{F} \) induces a classification on the subsets of \( J \), classifying icons according to properties of their associated stable states.

The practical content of this abstract picture is easy to see. For all icons \( I \), \( \mathcal{F}(I) \) is a set of stable states associated by an observer with the icon \( I \), and corresponding to \( I \) is a reference object. Typical reference objects are photographs, posters, paintings, projections, etc. We identify such reference objects with icons; although, more precisely, an icon defines an equivalence class of reference objects, any one of which can be chosen to represent the class. So in practice these common objects can be classified according to the stable states they can be said to support (disregarding all their familiar physical characteristics and their symbolic or ritualistic functions).

Stability theory can be applied to all such objects. Typical cases are discussed in three prior papers of the author (5–7). It is easy to see that the well-known and unusual perceptual aspects of graphic work of Escher and Dali, of graphic constructions by the psychologists Rubin (8) and Kanizsa (9), and of woodcuts of Kuniyoshi, to cite specific cases, can be given an explicit mathematical characterization using our techniques (5).

Two well-known cases demonstrate how closely the mathematical theory fits experience. Necker's "cube" (see 1,1 footnote) has three fundamental stable states which are described in a natural and exhaustive way by the corresponding mathematics. The illusion of the "three-pronged fork" can be explained by the fact that a stable state cannot be found that preserves both the three-pronged and two-pronged ends. This accounts for the characteristic ambivalent appearance of the figure.

5.4. The conditions cited in 1.3 delineate the confines of stability theory and give the special restrictions that define the static case. Those special restrictions bear repetition. We required that "all perceptual judgments rendered by an observer of his visible surround are assumed rendered of a reference object which, relative to the observer-frame, is a wholly contained, proper part of his field of vision, and is fixed, static, flat and has no discernible holes."

These restrictions can be easily removed by mathematical refinement of the structures we use to represent image and icon, and by a corresponding augmentation of the list of perceptual judgments that we model. The unrestricted theory can be called the "kinetic" or "dynamic" case. However, the static case can be easily recovered from the kinetic case whenever the stated restrictions hold.

The kinetic case can be constructed from the theory of stability we have given in the following way: (i) The mathematical representations of image and icon can both be parameterized by time, so that, relative to an augmented set of perceptual judgments, time-dependent aspects of perception can be modelled consistently with the case given. (ii) Judgments such as color relationship, geometrical relationship, and relative motion, position, and orientation of parts of the visual field can be defined in terms of stable states. (iii) The analytic model and the mathematics of stable states can be used without essential alteration to represent the perception of icons that encompass the full field of vision. (iv) By augmenting the mathematical structure of the icon in a time-dependent extension of the basic model we have given, the physical distance relative to an observer of an object of perception can be modelled. It should be noticed that from stable states, as here defined, relative-depth relations can always be assigned locally to two adjacent regions along any common boundary arc, and spatial continuity can also be expressed.

The claim that the stability principle governs all unrestricted cases is strongly supported by our experience. To cite one familiar case: in strange or ill-lighted surroundings, misconstruction of spatial relationships can be observed to behave in accord with the stability hypothesis.

5.5. In summary, let us consider where we stand. The static theory has been shown to yield a grammar for vision (5). Study of relations on identities, which can be defined in the mathematics of stability theory, appears to lead to a well-defined theory of semantics (6). We have evidence that a stability theory for natural languages can in fact be constructed, whose mathematics and method are compatible both with stability theory for vision and with standard linguistic theories. And the issue of the dependence of language and vision can be addressed using these techniques. [For discussion, see ref. 6 and the unpublished manuscript by L. G. Shiman and R. T. Oehrle "An Analytic Grammar for Vision and Natural Languages," (1977).]

We can now propose, sanctioned by the postulate of biological interpretability, that the explicit mathematical conditions discovered in analyzing visual perception be used to frame and test hypothetical accounts of the biology of perceptual processes.

Traditional questions concerning the structure and function of the visual system can in principle be interpreted in any sufficiently complete extension of stability theory (as provided by the dynamic extension of the static case) and thus they fall within the same proposal. But in that setting, the ancient question of how the cognitive and biological aspects of vision are related can be formally addressed.

The mathematics and procedures through which we have characterized the stability law offer access to a new conception of the study of vision. As with any natural principle, a formal realization of the law of stability has only one function: to provide a bridge from immediate experience to scientific understanding.

I would like to thank the Department of Mathematics, the Office of the Provost, and the Department of Psychology, Massachusetts Institute of Technology, for their sustained support. And I would like to acknowledge my personal debt to Kenneth Hoffman, Richard Held, and Walter Rosenblith; and to Lynn Hackett, Ross Shiman, and I. M.
Singer; and to Richard Oehrle, Allen Kantrowitz, Susan Taylor, and Kosta Tsipis for their critical assistance.

8. Rubin, E. (1921) *Visuell wahrgenommene Figuren* (Gyldendalske Boghandel, Copenhagen, Denmark).