On the Glitch Phenomenon

Leslie Lamport and Richard Palais [1]

ABSTRACT. The Principle of the Glitch states that for any device which makes a discrete decision based upon a continuous range of possible inputs, there are inputs for which it will take arbitrarily long to reach a decision. The appropriate mathematical setting for studying this principle is described. This involves defining the concept of continuity for mappings on sets of functions. It can then be shown that the glitch principle follows from the continuous behavior of the device.

There has recently been an increasing awareness of the synchronizer "glitch" phenomenon [5,6]. The usual description of this phenomenon states that any device for deciding which of two asynchronous events occurs first can hang up in a metastable state for an arbitrarily long time. The purpose of this brief paper is to indicate an appropriate mathematical setting for describing the glitch phenomenon and for "proving" its existence. The glitch phenomenon can be generalized to the following principle:

> For any device making a decision among a finite number of possible outcomes, based upon a continuum of possible inputs, there will be inputs for which the device takes arbitrarily long to reach its decision.

It is assumed that the decision making is non-trivial, i.e., that not all possible inputs to the device lead to the same outcome. A proof of this principle must be based upon some continuity assumption about the device, and we will attempt to clarify the continuity principles that are involved.

There seem to be three methods by which people attempt to show that glitches can be avoided:

1. Allowing the device to make an arbitrary decision when it has trouble deciding. This simply introduces additional "don't care" decision outcomes, and does not help.
2. Introducing noise to drive the device out of its metastable state. The noise can be considered to be just an unpredictable input. The introduction of noise cannot eliminate the possibility of the device

[1] Research supported in part by NSF Grant No. NPS 75-08555 and National Software Works contract number F30602-76-C-0094

Keywords: glitch,synchronization

1980 Mathematics Subject Classification.

hanging up for an arbitrarily long time, but can make it impossible to predict which inputs will cause it to do so.

3. Arguing that a decision making device, such as a flip-flop, introduces a discontinuity because it always reaches one of a discrete set of stable final states from any of a continous range of initial states. Although the mapping from initial to final states is discontinuous, we will see that this does not contradict the basic continuity assumption upon which a proof of the glitch phenomenon is based.

We now indicate the appropriate formalism for considering the glitch phenomenon. Let $\mathbf{R}$ denote the set of real numbers, let $\mathbf{I}$ and $\mathbf{O}$ be two sets, and let $\mathcal{I}$ and $\mathcal{O}$ be sets of mappings from $\mathbf{R}$ to $\mathbf{I}$ and from $\mathbf{R}$ to $\mathbf{O}$, respectively. An element of $\mathbf{I}$ represents a possible value for the input to the device at some instant. An element i of $\mathcal{I}$ represents a possible input to the device, where $i(t)$ is the value of the input at time t. Similarly, $\mathbf{O}$ is the set of possible output values, and $\mathcal{O}$ is the set of possible outputs. (For simplicity, we assume that the device operates for all times. We could also assume that it starts at some specific time.) The device defines a mapping $\Delta : \mathcal{I} \to \mathcal{O}$; namely $\Delta(i)$ is the output produced by the input i. In other words, if the input at any time t is $i(t)$, then $\Delta(i)(t)$ is the output at time t.

As an example, consider an electronic arbitration device with two input wires, labeled a and b, and one output wire, functioning as follows. A single positive voltage pulse will arrive at each of the two inputs at some times t_a and t_b after time $t = 0$ and before time $t = 1$. If these two input pulses arrive sufficiently far apart, then the device is to produce a well-defined positive output pulse should the pulse on wire a arrive first, and the negative of that pulse if the pulse on wire b arrives first. If the two input pulses arrive closer together than the time resolution of the device, then the device may produce either the positive or the negative pulse; however, it may *not* produce any type of output other than the specified positive or negative pulse. Assume for convenience that all the pulses have the same shape and height, and let $p_s : \mathbf{R} \to \mathbf{R}$ be the continuous function that represents a positive pulse starting at time s. i.e., $p_s(t)$ is the voltage at time t for such a pulse. To be specific, we assume for the example that $p_{s+\epsilon}(t) = p_s(t - \epsilon)$. Then $\mathbf{I}$ is a subset of the set $\mathbf{R} \times \mathbf{R}$ of ordered pairs of numbers, and $\mathcal{I}$ is the set of functions of the form (p_r, p_s) with $0 < r, s < 1$, where $(p_r, p_s)(t) = (p_r(t), p_s(t))$. The function (p_r, p_s) represents pulses arriving on wire a at time r and on wire b at time s. The set $\mathbf{O}$ is a subset of $\mathbf{R}$, and $\mathcal{O}$ is some set of functions from $\mathbf{R}$ into $\mathbf{R}$ containing $\pm p_s$ for some s; we will see that realistic "continuous" behavior of the device implies that $\mathcal{O}$ must also contain the zero function, which represents the possibility of the device hanging up forever. For a more realistic example, we can consider a family of pulses that are acceptable approximationsto p_s. This yields obvious modifications to the sets $\mathcal{I}$ and $\mathcal{O}$.

The above description is purposely made explicit and concrete for purposes of exposition, but our argument will apply equally to much more and abstract devices provided only that they satisfy some basic properties to be made precise below.

The existence of the glitch is deduced from the *continuity* of the mapping $\Delta : \mathcal{I} \to \mathcal{O}$. Continuity is defined for mappings between topological spaces, and this is the mathematically natural degree of generality in which to derive the results we shall consider. However, for simplicity we will restrict our attention to the special case of *metric* spaces, and we present next a very brief, self-contained account of the few basic facts about metric spaces that we shall need, referring the reader to [2] or any elementary topology text for more details.

A metric space is a set X of "points" in which we have a well-defined notion of the distance $d(x, y)$ between any ordered pair of points (x, y) in $X \times X$. The distance function d is assumed to satisfy the following three natural properties:

(i) $d(x, y) = d(y, x)$;
(ii) $d(x, y) \geq 0$, with equality only when $x = y$; and
(iii) $d(x, z) \leq d(x, y) + d(y, z)$ (the "triangle inequality").

The most obvious example of a metric space is the set $\mathbf{R}$ of real numbers with its "usual" distance function $d(x, y) = |x-y|$. A sequence of points $\{x_n\}$ in X is said to *converge* to the point x in X if the sequence $d(x_n, x)$ of real numbers converges to zero, and in this case we write $x_n \to x$. From (ii) it is easy to see that a given sequence $\{x_n\}$ can converge to at most one point x, so it makes sense to say that x is *the* limit of the sequence $\{x_n\}$, written $x = \lim_{n\to\infty} x_n$. If X and Y are both metric spaces and $f : X \to Y$ is a mapping between them, then f is defined to be *continuous* if whenever $x_n \to x$ in X it follows that $f(x_n) \to f(x)$ in Y. Note that this condition can also be written as $\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$.

Continuity is one of the most important concepts in topology and, as our definition shows, it is based on the the notion of convergence. We shall assume that our sets $\mathbf{O}$ and $\mathbf{I}$ are metric spaces, so the notion of continuity is defined for the mappings $i : \mathbf{R} \to I$ and $\Delta(i) : \mathbf{R} \to O$. Our sets $\mathcal{I}$ and $\mathcal{O}$ are assumed to be sets of *continuous* functions, and we now want to define what it means for the mapping $\Delta : \mathcal{I} \to \mathcal{O}$ to be continuous. This of course requires defining the appropriate notion of convergence in $\mathcal{I}$ and $\mathcal{O}$.

Let us assume more generally that we are given a metric space $\mathbf{S}$ with a distance function d and also a set $\mathcal{S}$ of continuous functions from $\mathbf{R}$ into $\mathbf{S}$. To simplify notation, given a positive number r let us define

$d^r(f,g) = \max_{-r \le t \le r} d(f(t), g(t))$ (i.e., the maximum distance between corresponding points on the graphs of f and g in the interval $[-r, r]$). We will say that a sequence $\{f_n\}$ in $\mathcal{S}$ converges to an element f in $\mathcal{S}$. if for each positive r, f_n converge to f *uniformly* on the interval $[-r, r]$—this means that $d^r(f_n(t), f(t))$ should converge to 0 for each r. Following standard mathematical terminology, we shall also refer to this mode of convergence as "convergence in the compact-open topology". It is not hard to check that $d^*(f,g) = \sum_{r=1}^{\infty} 2^{-r} d^r(f,g)/(1 + d^r(f,g))$ defines an explicit distance function d^* for $\mathcal{S}$ such that $d^*(f_n, f) \to 0$ if and only if f_n converges to f in the compact-open topology.

(To get some feeling for this mode of convergence, let us take $\mathbf{S} = \mathbf{R}$, and take for $\mathcal{S}$ all continuous functions $f : \mathbf{R} \to \mathbf{R}$. Let f_0 denote the identically zero function. Define $f_n(t)$ to be zero if t is less than 0 or greater than $2/n$, $f_n(t) = nt$ for $0 \le t \le 1/n$, and $f_n(t) = 2 - nt$ for $1/n \le t \le 2/n$. Since $f_n(t) = 0$ when $n > 2/t$ it is clear that f_n converges to f_0 pointwise, but $d^1(f_n, f_0) = 1$ so the convergence is not uniform on $[-1, 1]$. However, if we take $f_n(t) = e^{x-n}$, then since each f_n is monitonically increasing, it follows that $d^r(f_n, f) = f_n(r) = e^r e^{-n}$, which tends to zero as $n \to \infty$, so f_n *does* converge to f_0 in the compact-open topology, even though each $f_n(t)$ tends rapidly to infinty with t.)

As an example, let $\mathbf{S}$ be the set $\mathbf{R}$ of real numbers, and let p_s be the element of $\mathcal{S}$ of our exampe above. We assume that $p_s(t)$ equals zero except when t is in some finite interval (depending on s). Then the continuity of the function p_s together with the relation $p_{s+\epsilon}(t) = p_s(t-\epsilon)$ implies the expected result that if s_n converges to s in $\mathbf{R}$, then $p_{s_n} \to p_s$ in the compact-open topology of $\mathcal{S}$. That is, the mapping $F : \mathbf{R} \to \mathcal{S}$ defined by $F(s) = p_s$ is continuous. A more surprising result of our definition is that if $s_n \to \infty$, then $p_{s_n} \to 0$ (where 0 denotes here the identically zero function). This follows from the fact that if $p_s(t) = 0$ for all $t < r$, then $d^r(p_s, O) = 0$.

We will need one more concept from topology before we are ready to discuss the glitch phenomenon. Let U be a subset of a metric space S, and let $[0,1]$ as usual denote the interval of real numbers t with $0 \le t \le 1$. If u_0 and u_1 are points in U, then a path in U from u_0 to u_1 is a continuous mapping $\pi : [0,1] \to S$ such that $\pi(0) = u_0$, $\pi(1) = u_1$, and $\pi(t)$ is in U for all t in $[0,1]$. We say that U is a *pathwise connected* subset of S if such a π can be found for each choice of u_0 and u_1 in U. If F is a continuous mapping of U into a metric space T and $\pi : [0,1] \to \mathcal{U}$ is as above, then the composition $F \circ \pi : [0,1] \to T$ is a path in T from $F(u_0)$ to $F(u_1)$. It follows that if U is a pathwise connected subset of T, then $F(U)$ is pathwise connected subset of T, where $F(U)$, the image of U under F, is the set of all points in T of the form $F(u)$ for some point u in U.

Let $\mathcal{S}$ again be the space of all continuous functions from $\mathbf{R}$ into $\mathbf{R}$, and let $\mathcal{U}_r$ be the set of all functions $\pm p_s$ with $0 < s \leq r$. If we set $p_\infty = 0$, this also defines $\mathcal{U}_\infty$. Now one can show that $\mathcal{U}_r$ is *not* pathwise connected for any r with $< r < \infty$, e.g., there is no continuous path $\pi : [0,1] \rightarrow \mathcal{U}_{100}$ such that $\pi(0) = p_1$ and $\pi(1) = -p_1$. However,$\mathcal{U}_\infty$ *is* pathwise connected. For example, we can define a continuous path $\pi : [0,1] \rightarrow \mathcal{U}_\infty$ with$\pi(0) = p_1$ and $\pi(1) = p_1$ as follows:

$$\pi(s) = \begin{cases} p_{1/(1-2s)}, & \text{if } 0 \leq s < 1/2; \\ 0, & \text{if } s = 1/2; \\ p_{1/(2s-1)}, & \text{if } 1/2 \leq s \leq 1. \end{cases}$$

If we let $\mathcal{U}_r$ be the set of all pulses that are sufficiently close to $\pm p_s$ for some $0 < s \leq r$, then we again find that $\mathcal{U}_r$ not pathwise connected unless $r = \infty$. (This assumes that no pulse is close to both p_s and $-p_{s'}$, unless $s = s' = \infty$.)

To establish the Principle of the Glitch, one proves three things:

1) The mapping $\Delta : \mathcal{I} \rightarrow \mathcal{O}$ is continuous, using the compact-open topologies on $\mathbf{I}$ and 0.
2) The space $\mathcal{I}$ is pathwise connected.
3) The set of outputs in $\Delta(I)$ for which the decision is made before some fixed time r is not pathwise connected.

Since 1) and 2) imply that $\Delta(\mathcal{I})$ is pathwise connected, 3) shows that for any finite time r there must be inputs in $\mathcal{I}$ for which the device does *not* reach a decision by time r.

Parts 2) and 3) of the proof are usually easy. In our example, to prove that $\mathcal{I}$ is pathwise connected, we must construct a continuous path $\pi : [0,1] \rightarrow \mathcal{I}$ with $\pi(0) = (p_{r_0}, p_{s_0})$ and $\pi(0) = (p_{r_1}, p_{s_1})$, for any r_1 and r_0 between 0 and 1. This can easily be done by taking $\pi(t) = (p_{(1-t)\, r_0 + t\, r_1}, p_{(1-t)\, s_0 + t\, s_1})$. Part 3 is proved by showing the $\mathcal{U}_r$ is not pathwise connected for $0 < r < \infty$.

To prove that Δ is continuous with respect to the compact-open topologies, one must make some assumption about the nature of the dynamical equations that govern the behavior of the device. If one assumes that $\mathbf{I}$ and $\mathbf{O}$ are finite-dimensional manifolds, and that the behavior of the device is described by a system of ordinary differential equations (possibly involving delays) then one need only make fairly natural assumptions about these equations in order to deduce the continuity of Δ. The reader is referred to [1,3,4] for the appropriate theorems.

The mathematical situation is not so satisfactory if $\mathbf{I}$ and $\mathbf{O}$ are infinite dimensional and the device is described by a system of partial differential equations. We know of no general result that can be applied in this case.

Any proof of the existence of the glitch phenomenon raises the question of the extent to which a mathematical result can be applied to a physical situation. This is a metaphysical question that is beyond the scope of this paper. We merely observe that science is based upon the assumption that an approximately correct theory will describe the approximate behavior of a system. If the proof is based upon a theory that closely approximates the physical device, then we may safely conclude that the device must occasionally take very much longer than usual to make a decision. Whether is must “really” take an unbounded length of time to decide cannot be determined from any approximate theory.

References

[1] E. A. Coddington and N. Levinson. Theory of Ordinary Differential Equations. McGraw Hill, New York, 1955.

[2] J. Dieudonné. Foundations of Modern Analysis. Volume 1, Academic Press, New York, 1960.

[3] J. Hale. Functional Differential Equations. Springer-Verlag, New York, 1971.

[4] M. Hirsch and S. Smale. Differential Equations, Dynamical Systems and Linear Algebra. Academic Press, New York, 1974.

[5] M. Pěchouček. Anomalous response times of input synchronizers. IEEE Transactions on Computers, 25:133-139, February 1976.

[6] T. J. Chaney and C. E. Molnor. Anomalous behavior of synchronizer and arbiter circuits. IEEE Transactions on Computers, 22:421-422, April 1973.