

EXISTENCE THEOREM FOR THE DIFFERENCE EQUATION

$$Y_{n+1} - 2Y_n + Y_{n-1} = h^2 f(y_n)$$

F. WEIL

Department of Physics-Mathematics
Université de Moncton
Moncton, N. B., Canada

(Received November 27, 1978)

ABSTRACT. For the difference equation $\frac{(Y_{n+1} - 2Y_n + Y_{n-1})}{h^2} = f(Y_n)$ sufficient conditions are shown such that for a given Y_0 there is either a unique value of Y_1 for which the sequence Y_n strictly monotonically approaches a constant as n approaches infinity or a continuum interval of such values. It has been shown previously that the first alternative is related to the existence of a Peierls barrier energy in the dislocation model of Frenkel and Kontorova.

KEY WORDS AND PHRASES. Existence Theorem, Difference equations.

1980 MATHEMATICS SUBJECT CLASSIFICATION CODES. A3910

1. INTRODUCTION.

In this paper we discuss the conditions for the existence and uniqueness

of the solution to the nonlinear difference equation

$$\frac{Y_{n+1} - 2Y_n + Y_{n-1}}{h^2} = f(Y_n) \quad (1.1)$$

introduced in § 3 of Hobart (1965). As stated there, the boundary conditions are

$$n \geq 0 \quad (1.2a)$$

$$0 \leq Y_1 < \pi \quad (1.2b)$$

$$Y_n \xrightarrow{\text{s.m.}} \pi \text{ as } n \rightarrow \infty \quad (\text{s.m.: strictly monotonically}) \quad (1.2c)$$

The function f is odd, twice differentiable, negative on the interval $0 < Y < \pi$ and zero at the ends of this interval. We assume also that f has been standardized according to § 3 of Hobart (1965). This assumption gives us reason not to incorporate h^2 in f .

2. DEFINITIONS.

Consider first the difference equation (1.1) together with the conditions (1.2a) and (1.2b) only. For a given Y_0 we can choose arbitrarily a value of $0 \leq Y_1 < \pi$ and then, step by step calculate Y_n . Only three cases occur

$$Y_0 < Y_1 < \dots < Y_N \leq \pi < Y_{N+1} \quad (2.1a)$$

$$Y_0 < Y_1 < \dots < Y_n \geq Y_{n+1} \text{ with } Y_n < \pi, \quad (2.1b)$$

or $Y_0 \geq Y_1$

$$Y_0 < Y_1 < \dots < Y_n < Y_{n+1} < \dots < \pi \quad (2.1c)$$

We shall call a value of Y_1 for which (2.1a) [(2.1b) or (2.1c)] holds, a "large" ["small" or "correct"] Y_1 . Furthermore a "too large" ["too small"] Y_1 is a large [small] Y_1 for which any larger [smaller] Y_1 is large [small].

We notice that the Y_n 's corresponding to a correct Y_1 form a strictly monotonically increasing sequence bounded by π . Thus the Y_n 's have a limit ℓ as n approaches infinity. The limit ℓ must satisfy $\ell - 2\ell + \ell = h^2 f(\ell)$ or $f(\ell) = 0$ for which, according to the restrictions stated in the introduction, the only suitable root is π .

3. THEOREM I.

If $Y_{n+1} - 2Y_n + Y_{n-1} = h^2 f(Y_n)$,

$$\frac{d^2 f(Y)}{dY^2} \equiv \ddot{f}(Y) \quad \text{exists,}$$

Y_0 is constant,

$$1 + h^2 \dot{f}(Y) > 0,$$

$4 \leq n \leq m$ implies $0 < Y_{n-2} < Y_{n-1} < \pi$, and

$$\ddot{f}(Y) > 0 \quad \text{on} \quad 0 < Y < \pi;$$

then $\frac{dY_n}{dY_{n-1}} > 0$ for each $2 \leq n \leq m$.

PROOF. We first show $\frac{dY_n}{dY_{n-1}} > 0$ for $n = 2$ and $n = 3$:

$$\frac{dY_2}{dY_1} = 2 + h^2 \dot{f}(Y_1) > 1 > 0, \quad (3.1)$$

and
$$\frac{dY_3}{dY_2} = 2 - \frac{dY_1}{dY_2} + h^2 \dot{f}(Y_2), \quad (3.2)$$

so
$$\frac{dY_3}{dY_2} > 1 + h^2 \dot{f}(Y_2) > 0. \quad (3.3)$$

Now we show $\frac{dY_n}{dY_{n-1}} > 0$ for $4 \leq n \leq m$ assuming $\frac{dY_{n-2}}{dY_{n-3}} > 0$
and $\frac{dY_{n-1}}{dY_{n-2}} > 0$:

$$\frac{dY_n}{dY_{n-1}} = 2 - \frac{dY_{n-2}}{dY_{n-1}} + h^2 \dot{f}(Y_{n-1}). \quad (3.4)$$

Thus $\frac{dY_n}{dY_{n-1}} > 0$ if

$$h^2 \dot{f}(Y_{n-1}) > -2 + \frac{1}{\left(2 - \frac{dY_{n-3}}{dY_{n-2}} + h^2 \dot{f}(Y_{n-2})\right)} \quad (3.5)$$

Since $\frac{dY_{n-1}}{dY_{n-2}} > 0$ has been assumed,

$$\frac{dY_{n-3}}{dY_{n-2}} < 2 + h^2 \dot{f}(Y_{n-2}) \quad (3.6)$$

But $4 \leq n \leq m$ so $0 < Y_{n-2} < Y_{n-1} \leq \pi$ and $\ddot{f}(Y) > 0$ on

$0 < Y_{n-2} < Y < \pi$.

Thus $\dot{f}(Y_{n-2}) < \dot{f}(\pi)$, (3.7)

resulting in $\frac{dY_{n-3}}{dY_{n-2}} < 2 + h^2 \dot{f}(\pi)$. (3.8)

And also
$$\dot{f}(Y_{n-2}) < \dot{f}(Y_{n-1}) . \quad (3.9)$$

Using (3.8) and (3.9) we can modify (3.5) to $\frac{dY_n}{dY_{n-1}} > 0$

if
$$h^2 \dot{f}(Y_{n-2}) > -2 + \frac{1}{(h^2 \{\dot{f}(Y_{n-2}) - \dot{f}(\pi)\})} . \quad (3.10)$$

Noting (3.7) we obtain $\frac{dY_n}{dY_{n-1}} > 0$ if

$$\begin{aligned} & [h^2 \dot{f}(Y_{n-2})]^2 + (2 - h^2 \dot{f}(\pi)) [h^2 \dot{f}(Y_{n-2})] \\ & - (1 + 2h^2 \dot{f}(\pi)) < 0 . \end{aligned} \quad (3.11)$$

The two roots of this polynomial in $h^2 \dot{f}(Y_{n-2})$ are such that $r_- < -2$ and $h^2 \dot{f}(\pi) < r_+$. With (3.7) we obtain

$$r_- < h^2 \dot{f}(Y_{n-2}) < r_+ \quad (3.12)$$

on which range the polynomial is negative so (3.11) is satisfied.

Thus $\frac{dY_n}{dY_{n-1}} > 0$ for each $2 \leq n \leq m$.

4. THEOREM II.

Under the assumptions of Theorem I and the additional assumption that $f(Y) < 0$ on $0 < Y < \pi$, we show that large implies too large and small implies too small.

PROOF. Assume a constant value for Y_0 and an initial value of $Y_1 = Y_1^L$ that is large. This means there is an N such that Y_0 and Y_1^L through $Y_{N+1} - 2Y_N + Y_{N-1} = h^2 f(Y_N)$ give $Y_0 < Y_1^L < \dots < Y_N^L \leq \pi < Y_{N+1}^L$. By Theorem I (with $m = N + 1$) if Y_1 is increased, Y_{N+1} must also increase until $Y_N = \pi$. By

Theorem I (with $m = N$) if Y_1 is further increased, now Y_N must also increase until $Y_{N-1} = \pi$. Since N is finite, repetition of this process can be continued giving finally that Y_2 must also increase until $Y_1 = \pi$. During each step Y_1 is large. Thus if $Y_1 = Y_1^L$ is large, all $Y_1 > Y_1^L$ are large.

Assume a constant value for Y_0 and an initial value of $Y_1 = Y_1^S$ that is small. Now suppose there is a smaller $Y_1 = Y_1^*$ that is either large or correct. If Y_1^* were large, this would contradict the argument that large implies too large. If Y_1^* were correct, we would also have a contradiction. Since Y_1^S is small, there is an M such that $Y_0 < Y_1^S < \dots < Y_n^S \geq Y_{n+1}^S$ with $0 \leq Y_n < \pi$. (If $Y_1^S \leq Y_0$, that small implies too small is trivial). The assumptions that $f(Y) < 0$ on $0 < Y < \pi$ and $Y_{n+1} - 2Y_n + Y_{n-1} = h^2 f(Y_n)$ give that there must be a first $Y_p^S < 0$ for some $0 < p < M + 1 + \left\{ \frac{Y_n^S}{-h^2 f(Y_n^S)} \right\}$ with no $Y_n^S \geq \pi$ for $0 < n < p$. If below Y_1^S there were a correct Y_1^* , then by Theorem I (valid for all n if Y_1 is correct and for all n to and including $N + 1$ if Y_1 is large) as Y_1 is increased from Y_1^* either all Y_n 's increase or at least one $Y_n \geq \pi$ with no prior $Y_n < 0$ ($n \neq 0$). For neither case can such a Y_1 be small. There is a contradiction in assuming a correct Y_1 below a small Y_1 . Thus if $Y_1 = Y_1^S$ is small, all $Y_1 < Y_1^S$ are small.

5. ALGORITHM.

If $f(Y) < 0$ on $0 < Y < \pi$ and zero for the end points, and if the assumptions of Theorem I are satisfied, we can by the algorithm described in § 2 of the paper by Hobart (1965) construct a correct value of Y_1 for a given Y_0 . This involves choosing an interval bounded above by a large Y_1 , and below by a small Y_1 (initially

$0 < Y_1 < \pi$), testing the midpoint for large or small, retaining the (half) interval bounded as the original, and repeating the process on this interval.

If the midpoint is at no step correct, this process leads to a unique limit point which we shall now argue must be a correct point and the only correct point. Certainly there are no correct points to be found on the discarded intervals for if the midpoint is large [small], the discarded interval contains only points which are large [small]. Since f is differentiable, it is continuous. Thus all points in an infinitesimal neighbourhood of a large [small] point must be large [small]. But the limit point has in its neighbourhood both large and small points, so it must therefore be a correct point and the only correct point.

If the midpoint is at some step a correct point, either it is the only correct point or there is a continuous interval of correct points. Two correct points cannot be separated by a large [small] point since above [below] a large [small] point there can be only large [small] points. A test can be made which distinguishes between an isolated correct point and a continuum interval of correct points: If a midpoint $Y_1 = Y_1^C$ is correct, apply the algorithm described above separately to the intervals $Y_1^C < Y_1 < \pi$ and $0 < Y_1 < Y_1^C$. If at no step for either the midpoint is correct, then Y_1^C is unique. If the midpoint for either is at some step correct, then there is a continuum interval.

6. THEOREM III.

Unless the result of the algorithm is a continuous interval of correct values of Y_1 , it defines a unique $Y_1 \equiv g(Y_0)$ for each Y_0

on $0 \leq Y_0 < \pi$ since it is easily verified that $Y_1 = \pi$ is large and $Y_1 = 0$ is small. For application to the Frenkel-Kontorova model, we need the domain extended. Define $\beta \equiv g(0)$ and note that necessarily $\beta < \pi$. We now extend the domain of definition g to include $-\beta < Y_0 < 0$ by showing that for Y_0 in this domain $Y_1 = 0$ is small. Relabel $Y_{n+1} = \bar{Y}_n$. If $\bar{Y}_1 = \beta$ is the unique correct \bar{Y}_1 for $\bar{Y}_0 = 0$, then all $0 < \bar{Y}_1 < \beta$ are small for $\bar{Y}_0 = 0$. Noting that $f(Y_1) = 0$ so that $Y_2 = -Y_0$, it follows that $Y_1 = 0$ is small for each $0 > Y_0 > -\beta$.

THEOREM III. If the assumptions of Theorem I and the additional assumptions that $f(Y) < 0$ on $0 < Y < \pi$ and $f(Y) = 0$ for $Y = 0$ or $Y = \pi$ are satisfied, then either for each $0 \leq Y_0 < \pi$ there is one and only one correct $Y_1 = g(Y_0)$ and for each $-g(0) < Y_0 < 0$ there is one and only one correct $Y_1 \equiv g(Y_0)$ or for some $-\pi < Y_0 < \pi$ there is a continuous interval of correct values of Y_1 .

7. APPLICATION.

Assuming the function f is odd and that for the function f chosen there is no continuum of correct Y_1 values, we can use the function g to define the path of configurations connecting II with I in the Frenkel-Kontorova (1938) model as generalized in Hobart (1965). For a given $-g(0) < Y_0 < g(0)$, Y_1 is chosen so that $Y_n \xrightarrow{\text{s.m.}} \pi$ as $n \rightarrow \infty$, that is $Y_1 = g(Y_0)$; and Y_{-1} is chosen so that $Y_n \xrightarrow{\text{s.m.}} -\pi$ as $n \rightarrow -\infty$, that is $Y_{-1} = -g(-Y_0)$. The difference equation (1.1) is satisfied for all n except zero for which

$$(Y_0) \equiv f(Y_0) - \frac{g(Y_0) - 2Y_0 - g(-Y_0)}{h^2} \quad (7.1)$$

is the nonzero external force only on the zeroth atom which is necessary to hold static a general intermediate configuration. The configurations I and II are given by the conditions that $Y_0 = Y_0^I = 0$ and $Y_0 = Y_0^{II} = -g(Y_0^{II}) < 0$ respectively. The connecting path is given by $Y_0^{II} \leq Y_0 \leq Y_0^I$. The barrier energy is

$$V(I) - V(II) = \int_{Y_0^{II}}^{Y_0^I} (\xi) d\xi \quad (7.2)$$

REFERENCES.

1. Frenkel, J. and Kontorova, T. On the theory of plastic deformation and twinning, Phys. Z. Sowjet 13 (1938) 1-10.
2. Frenkel, J. and Kontorova, T. Series of plastic deformation and twinning, J. Phys. (U.S.S.R.) 1 (1939) 137-149.
3. Hobart, R. Peierls stress dependence on dislocation width, J. Appl. Phys. 36 (1965) 1944-1948.