Title	A graphical example of policy improvement technique
Sub Title	
Author	柳井, 浩(Yanai, Hiroshi)
Publisher	慶応義塾大学藤原記念工学部
Publication year	1961
Jtitle	Proceedings of the Fujihara Memorial Faculty of Engineering Keio University Vol.14, No.53 (1961.) ,p.62(22)- 71(31)
JaLC DOI	
Abstract	This report provides a simple graphical procedure for obtaining the path of light, or in other words the shortest-time path between two given points, passing through n different media. (Fig. 1) Although this procedure may be useful in itself, it may be even more useful as a teaching aid in illustrating some properties of dynamic programming formulation and its solution by policy improvement technique. In I and II, the problem and its dynamic programming formulation are explained. In III and IV, an approach by policy improvement technique and a graphical procedure are proposed. And with some mathematical notes, an example worked out is given in V.
Notes	
Genre	Departmental Bulletin Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO50001004-00140053-0022

A Graphical Example of Policy Improvement Technique

(Received May. 2, 1962)

Hiroshi YANAI *

Abstract

This report provides a simple graphical procedure for obtaining the path of light, or in other words the shortest-time path between two given points, passing through n different media. (Fig. 1)

Although this procedure may be useful in itself, it may be even more useful as a teaching aid in illustrating some properties of dynamic programming formulation and its solution by policy improvement technique.

In I and II, the problem and its dynamic programming formulation are explained. In III and IV, an approach by policy improvement technique and a graphical procedure are proposed. And with some mathematical notes, an example worked out is given in V.

I. The path of light in inhomogeneous media

Consider now, n parallel glass plates, with equal thickness d, but with different indices of refraction closed tightly to each other. And assume that these plates are optically isotropic. The velocities of light in each of these n plates are given as $v_1, v_2, \ldots\ldots, v_n$ respectively. Given two points A_0 and A_n at the both ends of these n plates, it is asked to obtain the path of light between them. According to Fermat's principle, light takes the path of the shortest-time duration. It is then obvious that the path can not be a curve but a connected sequence of straight line segments, which has refracting points on the boundary lines of two neighbouring glass plates. Then, taking x-coordinate parallel to the faces of the glass plates, any admissible path can be completely characterised by a sequence of the values of x-coordinate at which the refractions take place: $(x_0=0)$, x_1, x_2, ..., (x_n=the coordinate of $A_n=H$). Then the time duration required by a path characterised by such a sequence is given by the representation

$$\sum_{i=1}^{n}(\sqrt{(x_i-x_{i-1})^2+d^2}\,/\,v_i)\ . \tag{1}$$

Now, since the path of light gives the minimum to this representation, the problem reduces to that of obtaining the sequence x_1, x_2, ..., x_{n-1} which gives the minimum to this representation (1)

*Hiroshi YANAI: Student of Graduate School of Keio Univ.

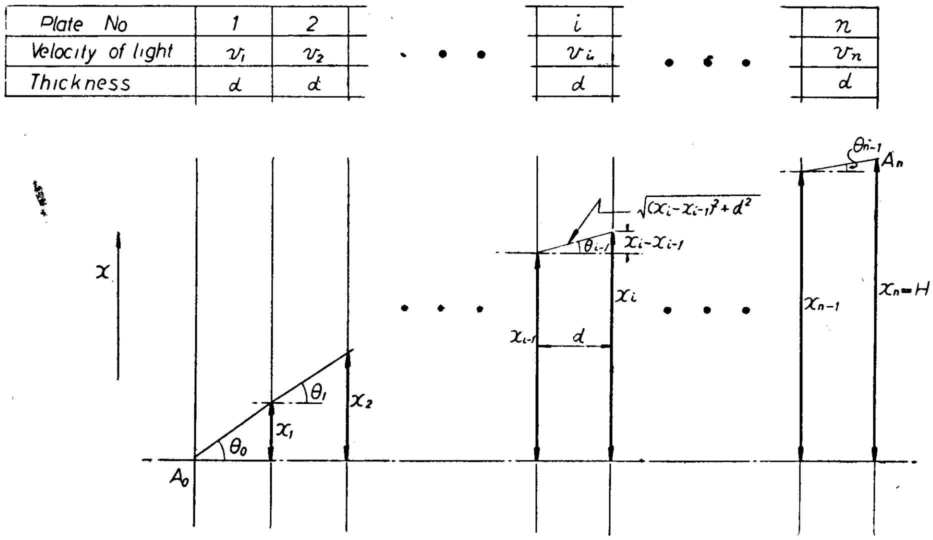

Fig. 1. n **parallel glass plates.**

$$\min_{x_1,\, x_2,\, \dots,\, x_{n-1}} \sum_{i=1}^{n} \frac{\sqrt{(x_i-x_{i-1})^2+d^2}}{v_i}. \tag{2}$$

Differentiating partially the representation (1) as to x_i ($i=1,\ 2,\ \dots\dots,\ n-1$) and equating them to zero, we get the following $n-1$ relations;

$$\left.\begin{aligned}
&\frac{x_1-x_0}{v_1\sqrt{(x_1-x_0)^2+d^2}}=\frac{x_2-x_1}{v_2\sqrt{(x_2-x_1)+d^2}}\,,\\
&\cdots\qquad\cdots\qquad\cdots\\
&\frac{x_i-x_{i-1}}{v_i\sqrt{(x_i-x_{i-1})^2+d^2}}=\frac{x_{i+1}-x_i}{v_{i+1}\sqrt{(x_{i+1}-x_i)^2+d^2}}\,,\\
&\cdots\qquad\cdots\qquad\cdots\\
&\frac{x_{n-1}-x_{n-2}}{v_{n-1}\sqrt{(x_{n-1}-x_{n-2})^2+d^2}}=\frac{x_n-x_{n-1}}{v_n\sqrt{(x_n-x_{n-1})^2+d^2}}\,.
\end{aligned}\right\} \tag{3}$$

Let the angles between the line segment (starting at the refracting point on the boundary between ith and $i+1$st plates) and the normal to the glass plates be θ_i. Then relations (3) reduce to

$$\left.\begin{aligned}
&\frac{\sin\theta_0}{v_1}=\frac{\sin\theta_1}{v_2}\,,\\
&\cdots\qquad\cdots\qquad\cdots\\
&\frac{\sin\theta_{i-1}}{v_i}=\frac{\sin\theta_i}{v_{i+1}}\,,\\
&\cdots\qquad\cdots\qquad\cdots\\
&\frac{\sin\theta_{n-2}}{v_{n-1}}=\frac{\sin\theta_{n-1}}{v_n}\,.
\end{aligned}\right\} \tag{4}$$

It is obvious that these are so-called Snell's law in optics. In any form neither (3) nor (4), it is not easy to solve these equations to obtain the numerical results because we must be confronted with either $n-1$ simultaneous equations of the 4 th order, or simultaneous trigonometrical equations with n variables.

II. Dynamic programming formulation

But, since from another point of view, this is a typical variational problem (cf. Eq. (2)) we may make use of dynamic programming method. Define $f_i(x)$ as

$f_i(x)$ = the time required by the sohrtest-time-path to A_n starting at a point on the boundary between ith and $(i-1)$ st glass plates which has the coordinate x. (5)

Then, according to the principle of optimality, we obtain the recurrence relations:

$$f_i(x)=\underset{y}{\mathrm{Min}}\left[\frac{\sqrt{(y-x)^2+d^2}}{v_i}+f_{i+1}(y)\right], \qquad i=1,\ 2,\ \ldots\ldots,\ n-1 \tag{6}$$

with the boundary condition,

$$f_n(x)=\frac{\sqrt{(x_n-x)^2+d^2}}{v_n}=\frac{\sqrt{(H-x)^2+d^2}}{v_n}, \tag{7}$$

where y is the coordinate of point of refraction on the boundary between ith and $(i+1)$ st media.

By solving back these functional equations from $f_{n-1}(x)$ to $f_1(x)$ numerically, we may obtain the path of light. But, as is readily seen, it requires tremendous amount of computation. To avoid this complexity, the policy improvement procedure may be proposed.

III. An approach by policy improvement technique

The idea of policy improvement technique is not new. Some examples of this idea may be seen in the solution of linear programming problem and of course in some solutions of functional equations of dynamic programming. In our case, this is a procedure of improving a conveniently taken path which connects the two points A_0 and A_n (an admissible or a feasible path), so that the time required becomes shorter and shorter with the replacement of the path in some part of it and the path finally reaching the required shortest-time-path in the limit.

In order to make the explanation definite, we set $n=3$, $v_1=6$, $v_2=2$, and $v_3=1$ respectively. Consider now, as a first feasible path, the path such as given by a heavy line in Fig. 2: $A_0(x_0=0)$, $A_1^0(x_1^0=0)$, $A_2^0(x_2^0=0)$, $A_3(x_3=H)$, where the superscrips attached to some of the notations denote the numbers of steps of improvements.

Plate No	1	2	3
Velocity of light	6	2	1
Thickness	d	d	d

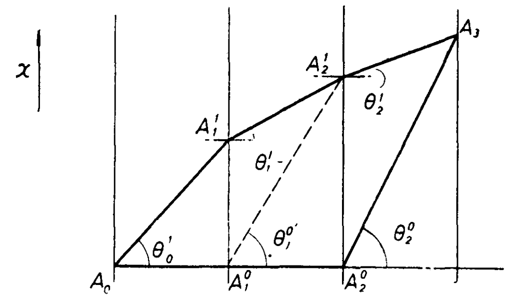

Fig. 2. The case of 3 parallel plates.

Now, let us start improving the first feasible path in the part of the path from A_1^0 to A_3. The improved path, if any, certainly requires shorter time than the first feasible path. Now, what is the shortest-time path between the two points A_0 and A_3? If we denote the improved path as A_0, A_1^0, A_2^1, A_3, the problem reduces to obtain the coordinate of the point A_2^1. This problem is fairly reduced in size because it corresponds to a problem to solve the recurrence relations (6) with $n=2$. And furthermore this problem can be easily solved using Snell's law numerically with the trigonometrical tables. That is, denoting the stages of incompleted improvement stage (explained later) by the prime to each corresponding notation, the improved angle between the normal to the plates and the new line segments, $\theta_1^{0\prime}$ and θ_2^1 must have the relation:

$$\frac{\sin\theta_1^{0\prime}}{v_2}=\frac{\sin\theta_2^1}{v_3}\ , \tag{8}$$

with the boundary condition

$$d\cdot(\tan\theta_1^{0\prime}+\tan\theta_2^1)=H\ , \tag{9}$$

if it takes the shortest-time route. And this corresponds to the relation as to y, the coordinate of the point A_2^1:

$$\frac{y}{v_2\cdot\sqrt{y^2+d^2}}=\frac{H-y}{v_3\cdot\sqrt{(H-y)^2+d^2}}\ , \tag{10}$$

which directly follows from equation (3).

IV. A graphical convention

Although it is not disappointedly difficult to solve equation (10) numerically, it may be quite a task to solve it for different values of H, which will be required later when we improve a path which is refracting at another point on this boundary. At this point, a graphical convention is to be proposed. That is, we start the calculation with some conveniently chosen values of $\theta_1^{0'}$ $(0 \leq \theta_1^{0'} < \pi/2)$ in equation (8). Then, calculate H of equation (9) through the value of θ_2^1. Plotting the results as points, having H as their abscissae and $d \cdot \tan\theta_1^{0'}$ as their ordinates correspondingly, and fitting a curve to them, we may make use of it to obtain the coordinate of the point A_2^1, the new refraction point on the boundary between the plate No. 2 and the plate No. 3 for different values of H. In Fig. 3 such a curve is drawn in a some what generalised form to utilize it for further steps of improvement, which may be easily understood by the following explanation.

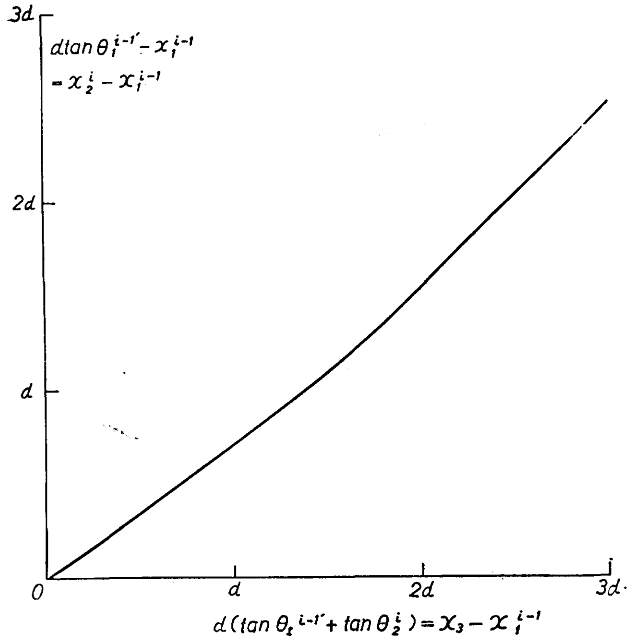

Fig. 3. The curve giving the local optimality.

The next procedure in improving the feasible path is to be considered in the part of the path between the points A_0 and A_1^2. Again, the problem is to find the shortest-time path between them; to obtain the coordinate of A_1^1. This can be obtained exactly in an analogous way we used for that of A_2^1. For this purpose, the curve of Fig. 4 may be utilized.

Now, we have completed the first step of improvement and we have the path A_0, A_1^1, A_2^1, A_3 instead of the first feasible path A_0, A_1^0, A_2^0, A_3. The next step of improvement may be done in the same way; first, we improve the part A_1^1, A_2^1, A_3. This can be done by entering into the abscissa of the graph of Fig. 3 with the difference between the ordinate of A_3 $(x_3=H)$ and the ordinate of A_1^1 (x_1^1); $x_3-x_1^1$

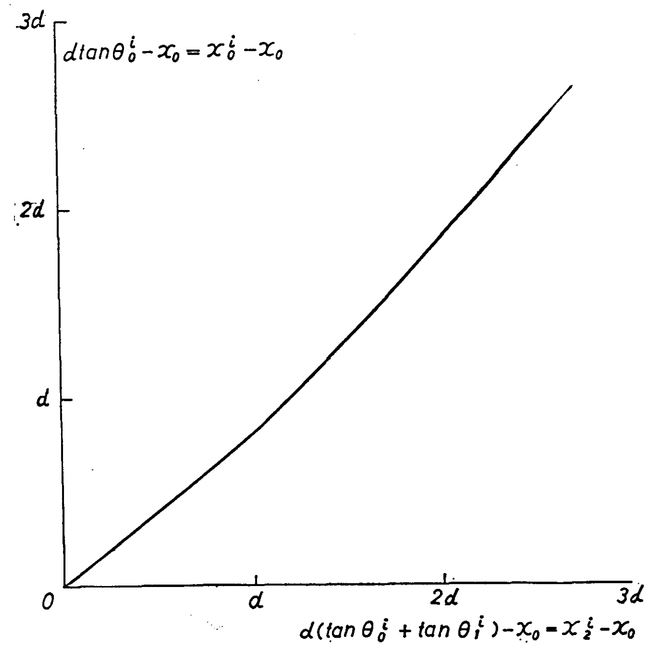

Fig. 4. The curve giving the local optimality.

and obtaining the difference between the coordinate of A_2^2 (x_2^2) and the ordinate of $A_1^1(x_1^1)$; $x_2^2 - x_1^1$. The next procedure, the procedure of obtaining the point A_1^2 is done by using the curve of Fig. 4. entering into whose abscissa with the coordinte of A_2^2 and obtaining the coordinate of the newly improved point A_1^2. In this way we complete the second step of the improvement procedure.

We may continue in this way the steps of the improvement procedure with the aid of the two curves in Fig. 3 and Fig. 4 in which the superscript i denotes the fact that the ith step of improvement is concerned.

V. Some mathematical notes on this procedure

Continuing the improvement in this way, we make the time required shorter and shorter, never becoming longer than the preceeding paths. In mathematical language, we may be sure of the monotone decreasing of the time associated with the paths in the process of the improvement.

We have another important mathematical question: "Does the improved path really converges to a path, and is this the path which offers the shortest time?" This question may be answered in the following way.

First, let us show the uniqueness of the path which satisfies the conditions (3) or equivalently (4). Suppose that we have two paths connecting the two points A_0 and A_n, each of them satisfing the conditions (4). The refraction angles (θ_0,

θ_1, ..., θ_{n-1}) and (φ_0, φ_1, ..., φ_{n-1}) of the two pahts, certainly can not be negative, and can not be equal nor greater than $\pi/2$

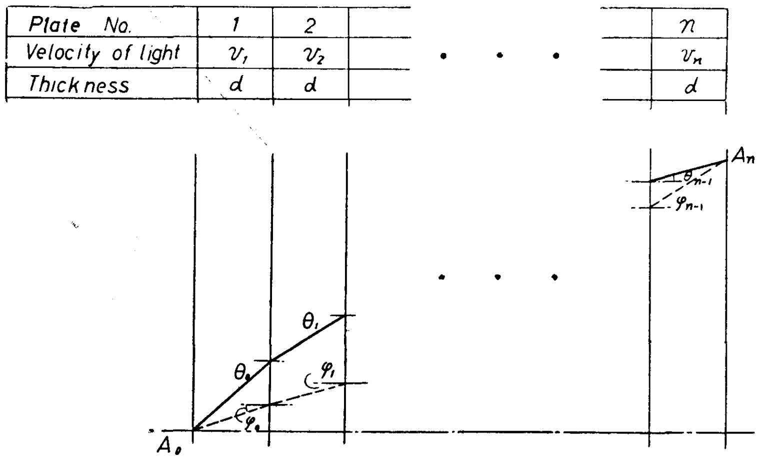

Fig. 5. The uniqueness of the path.

We may assume that $\theta_0 > \varphi_0$ without loss of generality. According to Snell's law,

$$\frac{\sin\theta_0}{\sin\theta_1}=\frac{\sin\varphi_0}{\sin\varphi_1}\ , \tag{11}$$

which leads to

$$\sin\theta_1=\sin\varphi_1\cdot\frac{\sin\theta_0}{\sin\varphi_0}\ . \tag{12}$$

Since the function "sine" is monotonely increasing between 0 and $\pi/2$,

$$\frac{\sin\theta_0}{\sin\varphi_0}>1\ , \tag{13}$$

and hence

$$\sin\theta_1 > \sin\varphi_1\ , \tag{14}$$

that is

$$\theta_1 > \varphi_1\ . \tag{15}$$

It is then obvious by induction that

$$\theta_i > \varphi_i\ , \qquad i=0,\ 1,\ \ldots\ldots,\ n-1 \tag{16}$$

which contradicts to the fact that both of the paths connect the two given points A_0 and A_n:

$$d\cdot\sum_{i=1}^{n-1}\tan\theta_i=d\cdot\sum_{i=0}^{n-1}\tan\varphi_i=H\ , \tag{17}$$

and the question stated above can be completely answered with these two characteristics—the monotone increasing improvement and the uniqueness.

Before going to the example worked out, which is to appear in Fig. 7, we wish to show another useful behavior of the path in the process of the improvement: The path in the procedure ot improvement, beginning with an initial feasible path which is to be improved upwards for the first step of improvement, will be improved upwards also for further steps of the improvement. Conversely, the path with an initial feasible path, which is to be improved downwards for the first step, will be improved downwards also for further steps of the improvement. ($i=2$, 3, ...) An example of the initial feasible path for the former case is $x_0=0$, $x_1^0=0$, $x_2^0=0$, ..., $x_{n-1}^0=0$, $x_n=H$. An example for the latter case is $x_0=0$, $x_1^0=H$, $x_2^0=H$, ..., $x_n=H$. This behavior can be easily shown by induction. Let us show the former case.

It is necessary to show that

$$x_i^m > x_i^{m-1}\ , \qquad \text{for} \quad i=1,\ 2,\ \ldots\ldots,\ n-1 \qquad m=2,\ 3,\ \ldots\ldots \tag{18}$$

under the assumption

$$x_i^1 > x_i^0\ , \qquad \text{for} \quad i=1,\ 2,\ 3,\ \ldots\ldots \tag{19}$$

Assume now up to $(m-1)$ st step that

$$x_i^j > x_i^{j-1}\ , \qquad \text{for} \quad i=1,\ 2,\ \ldots\ldots,\ n-1 \qquad j=1,\ 2,\ \ldots\ldots,\ m-1 \tag{20}$$

Then for $m=m$, it is seen that

$$x_{n-1}^m > x_{n-1}^{m-1}\ , \tag{21}$$

where x_{n-1}^{m-1} is the optimal point between x_{n-2}^{m-2} and the last point $x_n^0=H$ and the point x_{n-1}^m is the optimal point between x_{n-2}^{m-1} and the last point $x_n^0=H$. But since the curves giving the local optimality such as given in Figs. 3 and 4 have differential coefficients less than unity uniformly on the considering interval, it is obvious that the above inequality holds. Consider now the case of x_{n-2}^m. This is the point which is optimal between the two points x_{n-1}^m and x_{n-3}^{m-1}.

And the point x_{n-2}^{m-1} is the point which is optimal between the two points x_{n-1}^{m-1} and x_{n-3}^{m-2}. It is already shown that $x_{n-1}^m > x_{n-1}^{m-1}$ and $x_{n-3}^{m-1} > x_{n-3}^{m-2}$. Consider now a point y which is optimal between x_{n-1}^m and x_{n-3}^{m-2}. It is easily seen that the point y is above the point x_{n-2}^{m-1} by an analogous argument concerning the differential coefficients (of the curve giving the local optimality) and the fact $x_{n-1}^m > x_{n-1}^{m-1}$. Furthermore, it can be seen also that $x_{n-2}^m > y$ by the fact $x_{n-3}^{m-1} > x_{n-3}^{m-2}$. Then we can conclude that $x_{n-2}^m > x_{n-2}^{m-1}$. We can show the inequality (18) in this way inductively. The latter (downward) case can be shown exactly in the same way.

Fig. 7 is an example worked out starting from the two initial feasible paths: $x_0=0$, $x_1^0=0$, $x_2^0=0$, $x_3=H$ and $x_0=0$, $x_1=H$, $x_2=H$, $x_3=H$. By the behavior shown above, the path of light connecting the two points A_0 and A_3 must be somewhere in the hatched region.

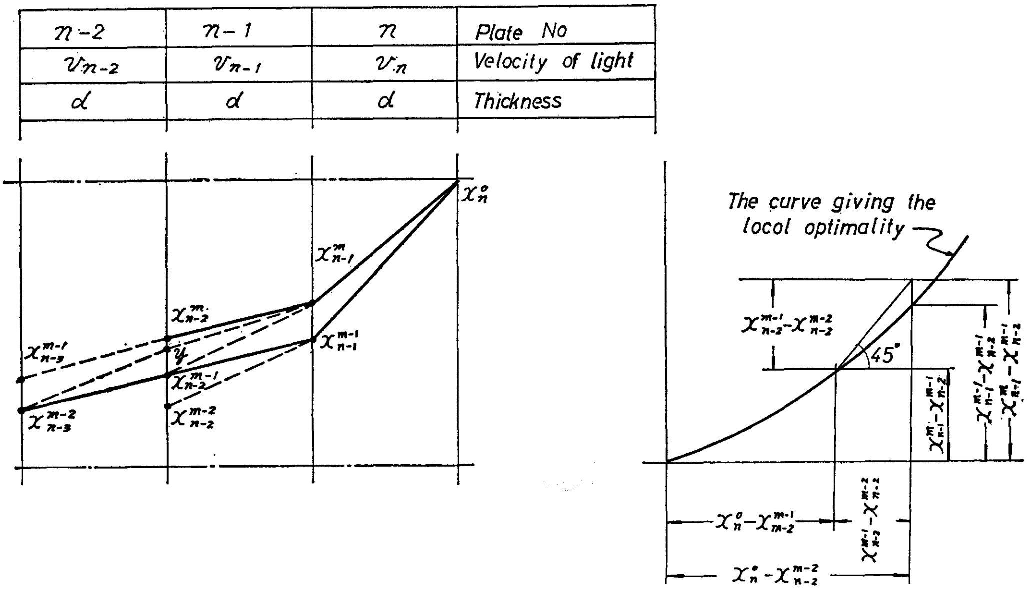

Fig. 6. *a*) Behavior of the paths in the process of the improvement. *b*) The curve giving the local optimality.

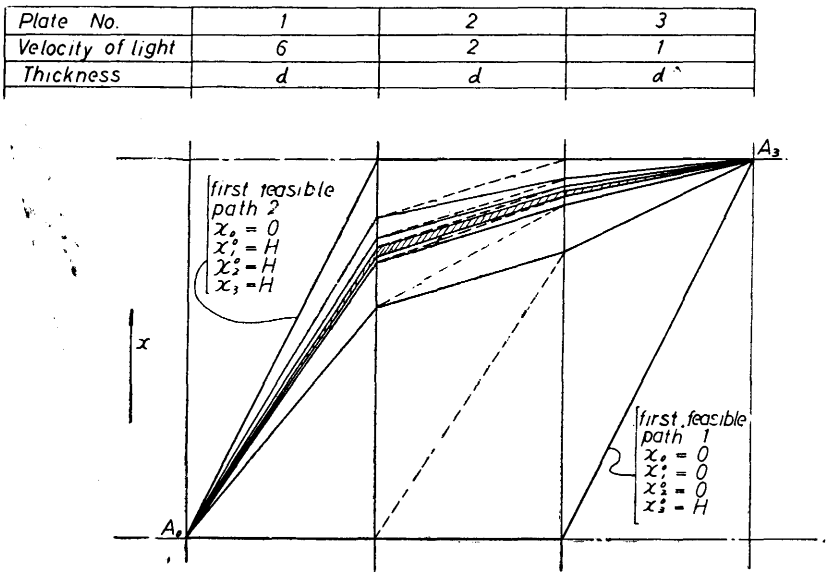

Fig. 7. Example worked out.

It should be mentioned here that although we have treated, in this article, the case of n glass plates of equal thickness, it is quite easy to extend the arguments to the case of n parallel plates with unequal thicknesses by a slight change of graphs giving the local optimality in Fig. 3 and Fig. 4. Furthermore this procedure can be used to solve a class of Brachistcrone problem approximately following the approach once used by J. Bernoulli.

Acknowledgement

The author wishes to show his acknowledgment to Prof. Ura of Keio Univ. who read through this article in detail and gave the author many profitable advices.