Spiral Curriculum: Towards Mathematical Foundations

Olga Kosheleva
Department of Teacher Education
University of Texas at El Paso
El Paso, TX 79968, USA
Email: olgak@utep.edu

***Abstract*—One of the fundamental ideas of modern education is the idea of spiral curriculum, when students repeatedly revisit the same sequence of topics at the increasing levels of depth, detail, and sophistication. In this paper, we show that under reasonable assumptions, the optimal sequence of presenting the material should indeed follow a spiral pattern.**

I. Introduction

A. What is a spiral curriculum

The concept of a spiral curriculum was developed by Harvard's professor Jerome S. Bruner, a one-time President of the American Psychological Association, and one of the major modern thinkers on education (see, e.g., [14]). This notion was first developed in his classic monograph [2].

The main idea of the spiral curriculum is that students repeatedly revisit the same sequence of topics at the increasing levels of depth, detail, and sophistication. This idea has been further developed in his following books [3], [4], [6], [7], [9], [10].

This approach has been successfully used at all levels of education, from the kindergarten level to the level of university education.

B. What we do in this paper

In this paper, we:

- describe the problem of selecting the optimal sequence of presenting the material,
- formalize this problem as a precise optimization problem, and
- prove that the resulting optimal sequence is indeed a spiral.

Thus, we provide an additional mathematical justification of the spiral curriculum approach.

C. What techniques we use

In this paper, we use techniques based on symmetries and symmetry groups.

These techniques have been successfully used in physics. In particular, in [12], [13], we used symmetry techniques to describe optimal trajectories and optimal shapes – e.g., including optimal shapes of celestial bodies; see also [16].

In [1], [15], we extended these techniques to education problems. In this paper, we use these techniques to explain the optimality of spiral curriculum.

II. Preliminary discussions

A. It is reasonable to restrict ourselves to finite-parametric descriptions

Let us start with an appropriate description of the set of all the material that needs to be learned.

The ultimate objective of our mathematical analysis is to help the educators with selecting the best sequence of presenting the material. For this purpose, we must be able to present different parts of the material in the computer.

Inside any given computer, we can only store finitely many bits, and therefore, we can represent the information only about finitely many parameters. So, it is reasonable to restrict ourselves to *finite-dimensional descriptions*.

In other words, different parts of the material can be described by the values finitely many parameters $x_1, \ldots, x_n$. So, each part of the material can be represented as a point in a n-dimensional space R^n.

B. We need a curve, not a trajectory

Ideally, we should describe which part of the material should be presented at different moments of time t. In other words, we need to know the optimal dependence $x(t)$ of the material presented at moment t on this moment of time. So, from the mathematical viewpoint, it is desirable to describe a *trajectory*, i.e., a mapping from the set of real numbers (representing time) to the n-dimensional space R^n.

From the practical viewpoint, however, there are two aspects to this choice:

- first, we must select the order in which we present the material;
- second, we must select the speed with which the material is presented.

Selecting the proper order is difficult. Once the order is selected, selecting the speed of presentation is much easier: we just need to keep track of how fast students learn, and make sure that they learn the material to the desired level before moving to the next part of the material.

In view of this fact, in the following text, we will concentrate on the most difficult part of the problem: selecting the proper order. From the mathematical viewpoint, the order is represented by the corresponding *curve*, i.e., by a 1-dimensional set describing the covered material. In these terms, we would be looking for the optimal curve.

C. We need a family of curves, not a single curve

The optimal sequence of presenting the material depends on what exactly we want: e.g., as we have shown in [1], [15], this sequence depends on

- whether we want the students to be able to flawlessly apply this knowledge right away or
- whether we allow for a follow-up training period (like internship for medical doctors) in which they work under proper supervision.

As a result, we do not expect to find a *single* optimal curve, we expect to find a *family* of curves which are optimal under different optimality criteria.

D. In the computer, we can only use finite-parametric families of curves

As we have mentioned earlier, the ultimate objective of our mathematical analysis is to help the educators with selecting the best sequence of presenting the material. For this purpose, we must be able to present different curves in the computer.

From the purely mathematical viewpoint, we can have families characterized by *infinite* number of parameters: e.g., the family of all possible curves. However, inside any given computer, we can only store finitely many bits, and therefore, we can represent the information only about *finitely many* parameters. So, it is reasonable to restrict ourselves to *finite-dimensional family of curves*.

E. Main problem: which families of curves should we choose

In principle, different families of curves can be used. Therefore, it is important to choose the right family.

Currently, this choice is mainly made *ad hoc*, at best, by testing a few possible families and choosing the one that performs the best on a few benchmarks. Since only a few families are analyzed, we are not sure that we did not miss the real good approximating family. (And since only a few benchmarks are used for comparison, we are not sure that the chosen family is indeed the best one.) It is, therefore, desirable to find the *optimal* family of curves.

III. What does "optimal" mean? Motivations for the following definitions

A. What is "optimality criterion"

Our objective is to find an optimal family of curves. When we say "optimal", we mean optimal w.r.t. to some *optimality criterion*. When we say that some *optimality criterion* is given, we mean that, given two different families of approximating sets, we can decide whether the first one is better, or that the second one is better, or that these families are of the same quality w.r.t. the given criterion. In mathematical terms, this means that we have a *pre-ordering relation* $\preceq$ on the set of all possible finite-dimensional families of sets.

B. We want to solve an ambitious problem: enumerate all finite-dimensional families of curves that are optimal relative to some natural criteria

One way to approach the problem of choosing the "best" family of curves is to select *one* optimality criterion, and to find a family of curves that is the best with respect to this criterion. The main drawback of this approach is that there can be different optimality criteria, and they can lead to different optimal solutions.

It is, therefore, desirable not only to describe a family of curves that is optimal relative to *some* criterion, but to describe *all* families of curves that can be optimal relative to different natural criteria. In this paper, we are planning to undertake exactly this more ambitious task.

C. Numerical optimality criteria

Pre-ordering is the general formulation of optimization problems in general, not only of the problem of choosing a family of sets. In general optimization theory, in which we are comparing arbitrary *alternatives* A, B, $\ldots$, from a given set $\mathcal{A}$, the most frequent case of such a pre-ordering is when a *numerical criterion* is used, i.e., when a function $J : \mathcal{A} \to R$ is given for which $A \preceq B$ if and only if $J(A) \le J(B)$.

For example, we may want to select a family for which the speed of learning the given material is the largest, or the percentage of material retained under a certain period of time is the largest. For both criteria, the speed and the percentage depend on the student body. We can therefore consider worst-case criteria – by selecting the curve under which the guaranteed learning time if the smallest or the guaranteed retention percentage is the largest. Alternative, we can consider average-case criteria – by selecting the curve under which the average learning time is the smallest or the average retention percentage is the largest.

D. Non-numerical optimality criteria naturally appear

For "worst-case" optimality criteria, it often happens that there are several different alternatives that perform equally well in the worst case, but whose performance differ drastically in the average cases. In this case, it makes sense, among all the alternatives with the optimal *worst-case* behavior, to choose the one for which the *average* behavior is the best possible. This very natural idea leads to the optimality criterion that is *not* described by a numerical optimality criterion $J(A)$: in this case, we need *two* functions: $J_1(A)$ describes the worst-case behavior, $J_2(A)$ describes the average-case behavior, and $A \preceq B$ if and only if:

- either $J_1(A) < J_2(B)$,
- or $J_1(A) = J_1(B)$ and $J_2(A) \le J_2(B)$.

We could further specify the described optimality criterion and end up with *a* natural criterion. However, as we have already mentioned, the goal of this paper is not to find *a* family of curves that is optimal relative to some criterion, but to describe *all* families of curves that are optimal relative to some natural optimality criteria. In view of this goal, in the following text, we will not specify the criterion, but, vice versa, we will describe a very general class of *natural* optimality criteria.

So, let us formulate what "natural" means.

E. The optimality criterion must be invariant with respect to reasonable symmetries

Problems related to geometric sets often have natural *symmetries*. For example, in our presentation of knowledge, we may want to rotate the original coordinates $x_1, \dots, x_d$ and consider new ones. The choice of a *rotated* coordinate system is equivalent to rotating all the points: $x \to R(x)$. As a result, in the new coordinates, each curve $X \in A$ from a family of curves A will be described by a "rotated" curve $R(X) = \{R(x) \mid x \in X\}$, and the original family A turns into a "rotated" family $R(A) = \{R(X) \mid X \in A\}$. It is reasonable to require that the relative quality of the two families of sets do not change under this rotation, if A is better than B, then $R(A)$ is better than $R(B)$.

Another reasonable symmetry is re-scaling. Usually, the choice of units to describe the parameters x_i is rather arbitrary. If we replace the original unit with a new unit which is λ times smaller, then the new values of the parameter are increased by a factor of λ, i.e., $x \to \lambda \cdot x$. It is, therefore, natural to require that the desired optimality criterion be invariant with respect to such rescalings.

F. The criterion must be final

If the criterion does not select any family as an optimal one, i.e., if, according to this criterion, none of the families is better than the others, then this criterion is of no use in selection.

If the criterion considers several different families equally good, then we can always use some other criterion to help select between these "equally good" ones, thus designing a two-step criterion. If this new criterion still does not select a unique family, we can continue this process until we arrive at a combination multi-step criterion for which there is only one optimal family.

Therefore, we can always assume that our criterion is *final* in the sense that it selects one and only one optimal family.

IV. Definitions and the General Result

Our goal is to choose the best finite-parametric family of curves. To formulate this problem precisely, we must formalize what a finite-parametric family is and what it means for a family to be optimal. In accordance with our informal description, both formalizations will use natural symmetries.

So, we will first formulate how symmetries can be defined for families of sets, then what it means for a family of sets to be finite-dimensional, and finally, how to describe an optimality criterion. Curves will be then introduced as particular cases of sets.

Definition 1. *Let $g : M \to M$ be a 1-1-transformation of a set M, and let A be a family of subsets of M. For each set $X \in A$, we define the result $g(X)$ of applying this transformation g to the set X as $\{g(x) \mid x \in X\}$, and we define the result $g(A)$ of applying the transformation g to the family A as the family $\{g(X) \mid X \in A\}$.*

Definition 2. *Let M be a smooth manifold. A group G of transformations $M \to M$ is called a* Lie transformation group*, if G is endowed with a structure of a smooth manifold for which the mapping $g, a \to g(a)$ from $G \times M$ to M is smooth.*

We want to define r-parametric families sets in such a way that symmetries from G would be computable based on parameters. Formally:

Definition 3. *Let M and N be smooth manifolds.*

- *By a* multi-valued function *$F : M \to N$ we mean a function that maps each $m \in M$ into a discrete set $F(m) \subseteq N$.*
- *We say that a multi-valued function is* smooth *if for every point $m_0 \in M$ and for every value $f_0 \in F(m)$, there exists an open neighborhood U of m_0 and a smooth function $f : U \to N$ for which $f(m_0) = f_0$ and for every $m \in U$, $f(m) \subseteq F(m)$.*

Definition 4. *Let G be a Lie transformation group on a smooth manifold M.*

- *We say that a class A of closed subsets of M is G-*invariant *if for every set $X \in A$, and for every transformation $g \in G$, the set $g(X)$ also belongs to the class.*
- *If A is a G-invariant class, then we say that A is a* finitely parametric family of sets *if there exist:*
  - *a (finite-dimensional) smooth manifold V;*
  - *a mapping s that maps each element $v \in V$ into a set $s(v) \subseteq M$; and*
  - *a smooth multi-valued function $\Pi : G \times V \to V$*

  such that:
  - *the class of all sets $s(v)$ that corresponds to different $v \in V$ coincides with A, and*
  - *for every $v \in V$, for every transformation $g \in G$, and for every $\pi \in \Pi(g, v)$, the set $s(\pi)$ (that corresponds to π) is equal to the result $g(s(v))$ of applying the transformation g to the set $s(v)$ (that corresponds to v).*
- *Let $r > 0$ be an integer. We say that a class of sets B is a* r-parametric class *of sets if there exists a finite-dimensional family of sets A defined by a triple (V, s, Π) for which B consists of all the sets $s(v)$ with v from some r-dimensional sub-manifold $W \subseteq V$.*

Definition 5. *Let $\mathcal{A}$ be a set, and let G be a group of transformations defined on $\mathcal{A}$.*

- *By an* optimality criterion*, we mean a* pre-ordering *(i.e., a transitive reflexive relation) $\preceq$ on the set $\mathcal{A}$.*
- *An optimality criterion is called G-*invariant *if for all $g \in G$, and for all $A, B \in \mathcal{A}$, $A \preceq B$ implies $g(A) \preceq g(B)$.*
- *An optimality criterion is called* final *if there exists one and only one element $A \in \mathcal{A}$ that is preferable to all the others, i.e., for which $B \preceq A$ for all $B \neq A$.*

- *An optimality criterion is called G-*natural *if it is G-invariant and final.*

Definition 6. *Let M be a smooth manifold. A set $X \subset M$ is called a* curve *if it is an image of a smooth mapping from a real line to M.*

Theorem 1. *Let M be a manifold, let G be a d-dimensional Lie transformation group on M, and let $r < d$ be a positive integer. Let $\mathcal{A}$ denote the class of all r-parametric families of sets from M, and let $\preceq$ be a G-natural optimality criterion on the class $\mathcal{A}$. Then:*

- *the $\preceq$-optimal family A_{opt} is G-invariant; and*
- *each set X from the $\preceq$-optimal family A_{opt} is a union of orbits of $\geq (d-r)$-dimensional subgroups of the group G.*

Proof. Since the criterion $\preceq$ is final, there exists one and only one optimal family of sets. Let us denote this family by A_{opt}.

1. Let us first show that this family A_{opt} is indeed G-invariant, i.e., that $g(A_{\text{opt}}) = A_{\text{opt}}$ for every transformation $g \in G$.

Indeed, let $g \in G$. From the optimality of A_{opt}, we conclude that for every $B \in \mathcal{A}$, $g^{-1}(B) \preceq A_{\text{opt}}$. From the G-invariance of the optimality criterion, we can now conclude that $B \preceq g(A_{\text{opt}})$. This is true for all $B \in \mathcal{A}$ and therefore, the family $g(A_{\text{opt}})$ is optimal. But since the criterion is final, there is only one optimal family; hence, $g(A_{\text{opt}}) = A_{\text{opt}}$. So, A_{opt} is indeed invariant.

2. Let us now show an arbitrary set X_0 from the optimal family A_{opt} consists of orbits of $\geq (d-r)$-dimensional subgroups of the group G.

Indeed, the fact that A_{opt} is G-invariant means, in particular, that for every $g \in G$, the set $g(X_0)$ also belongs to A_{opt}. Thus, we have a (smooth) mapping $g \to g(X_0)$ from the d-dimensional manifold G into the $\leq r$-dimensional set $G(X_0) = \{g(X_0) \mid g \in G\} \subseteq A_{\text{opt}}$. In the following, we will denote this mapping by g_0.

Since $r < d$, this mapping cannot be 1-1, i.e., for some sets $X = g'(X_0) \in G(X_0)$, the pre-image $g_0^{-1}(X) = \{g \mid g(X_0) = g'(X_0)\}$ consists of one than one point. By definition of $g(X)$, we can conclude that $g(X_0) = g'(X_0)$ iff $(g')^{-1}g(X_0) = X_0$. Thus, this pre-image is equal to $\{g \mid (g')^{-1}g(X_0) = X_0\}$. If we denote $(g')^{-1}g$ by $\tilde{g}$, we conclude that $g = g'\tilde{g}$ and that the pre-image $g_0^{-1}(X) = g_0^{-1}(g'(X_0))$ is equal to $\{g'\tilde{g} \mid \tilde{g}(X_0) = X_0\}$, i.e., to the result of applying g' to $\{\tilde{g} \mid \tilde{g}(X_0) = X_0\} = g_0^{-1}(X_0)$. Thus, each pre-image $(g_0^{-1}(X) = g_0^{-1}(g'(X_0)))$ can be obtained from one of these pre-images (namely, from $g_0^{-1}(X_0)$) by a smooth invertible transformation g'. Thus, all pre-images have the same dimension D.

We thus have a *stratification* (fiber bundle) of a d-dimensional manifold G into D-dimensional strata, with the dimension D_f of the factor-space being $\leq r$. Thus, $d = D + D_f$, and from $D_f \leq r$, we conclude that $D = d - D_f \geq n - r$.

So, for every set $X_0 \in A_{\text{opt}}$, we have a $D \geq (n-r)$-dimensional subset $G_0 \subseteq G$ that leaves X_0 invariant (i.e., for which $g(X_0) = X_0$ for all $g \in G_0$). It is easy to check that if $g, g' \in G_0$, then $gg' \in G_0$ and $g^{-1} \in G_0$, i.e., that G_0 is a *subgroup* of the group G. From the definition of G_0 as $\{g \mid g(X_0) = X_0\}$ and the fact that $g(X_0)$ is defined by a smooth transformation, we conclude that G_0 is a smooth sub-manifold of G, i.e., a $\geq (n-r)$-dimensional subgroup of G.

To complete our proof, we must show that the set X_0 is a union of orbits of the group G_0. Indeed, the fact that $g(X_0) = X_0$ means that for every $x \in X_0$, and for every $g \in G_0$, the element $g(x)$ also belongs to X_0. Thus, for every element x of the set X_0, its entire orbit $\{g(x) \mid g \in G_0\}$ is contained in X_0. Thus, X_0 is indeed the union of orbits of G_0. Q.E.D.

V. Towards Applying the General Result to the Education Problem: Appropriate Use of Symmetries

A. We need to describe appropriate symmetries

The above general result was formulated in terms of general symmetry groups. So to apply this general result to our education problem, it is necessary to find appropriate symmetries.

In geometric situations described in [12], [13], we have natural geometric symmetries like rotation. In the education problems, there are no natural symmetries, but we must find the appropriate symmetries in some indirect way.

B. To be able to describe symmetries, we first describe closeness

To describe symmetries, we will describe "closeness" between different parts of the material, and define symmetries as transformations which preserves this closeness – so that if x and y are ρ-close they either remain ρ-close after this transformation, or at least become $f(\rho)$-close for an appropriate function $f(\rho)$.

Closeness means that for each value of a distance ρ and for each point x, we have a set of all the points whose distance to x does not exceed x. We have different sets for different closeness criteria, for different points x, and for different value ρ. Let us describe the family of all such sets.

C. Linear transformations

It is reasonable to restrict ourselves to *linear* transformations $R^n \to R^n$, i.e., of all transformations of the type

$$x_i \to a_i + \sum_j a_{ij} x_j$$

with an invertible matrix a_{ij}.

D. Main result

We will show that the ellipsoids are the *simplest* optimal family, i.e., that of all possible optimal finite-parametric families that correspond to different G_e-invariant optimality criteria, ellipsoids have the smallest number of parameters.

Definition 7. *By a* closed domain*, we mean a closed set that is equal to the closure of the set of its interior points.*

Theorem 2. *Let $n > 0$ be an integer, $M = R^n$, G_e be the group of all affine transformations, and $\preceq$ be a natural* *(i.e., G_e-invariant and final) optimality criterion on the class $\mathcal{A}$ of all r-parametric families of connected bounded closed domains from R^n. Then:*

- *$r \geq n(n+3)/2$;*
- *if $r = n(n+3)/2$, then the optimal family coincides either with the family of all ellipsoids, or, for some $\lambda \in (0,1)$, with the family of all regions obtained from ellipsoids by subtracting λ times smaller homothetic ellipsoids.*

Comment. If we restrict ourselves to *convex* sets (or only to simply connected sets), we get ellipsoids only.

Proof. Due to Theorem 1, the optimal family $A_{\rm opt}$ is affine invariant, i.e., for every $X \in A_{\rm opt}$, and for every transformation $g \in G_e$, the set $g(X)$ also belongs to $A_{\rm opt}$.

1. Let us first show that $r \geq n(n+3)/2$. Indeed, it is known (see, e.g., [11]) that for every open bounded set X, among all ellipsoids that contain X, there exists a unique ellipsoid E of the smallest volume. We will say that this ellipsoid E *corresponds* to the set X. Let us consider the set of ellipsoids $\mathcal{E}_c$ that correspond (in this sense) to all possible sets $X \in A_{\rm opt}$.

Let us fix a set $X_0 \in A_{\rm opt}$, and let E_0 denote an ellipsoid that corresponds to X_0.

An arbitrary ellipsoid E can be obtained from any other ellipsoid (in particular, from E_0) by an appropriate affine transformation g: $E = g(E_0)$. The ratio of volumes is preserved under arbitrary linear transformations g; hence, since the ellipsoid E_0 is the smallest volume ellipsoid that contains X_0, the ellipsoid $E = g(E_0)$ is the smallest volume ellipsoid that contains $g(X_0) = X$.

Hence, an arbitrary ellipsoid $E = g(E_0)$ corresponds to some set $g(X_0) \in A_{\rm opt}$. Thus, the family $\mathcal{E}_c$ of all ellipsoids that correspond to sets from $A_{\rm opt}$ is simply equal to the set $\mathcal{E}$ of all ellipsoids. Thus, we have a (locally smooth) mapping from an r-dimensional set $A_{\rm opt}$ onto the $\dfrac{n(n+3)}{2}$-dimensional set of all ellipsoids. Hence, $r \geq n(n+3)/2$.

2. Let us now show that for $r = n(n+3)/2$, the only G_e-invariant families A are ellipsoids and "ellipsoid layers" (described in Theorem 2).

Indeed, let X_0 be an arbitrary set from the invariant family, and let E_0 be the corresponding ellipsoid. Let $g_0 \in G_e$ be an affine transformation that transform E_0 into a ball $E_1 = g(E_0)$. This ball then contains the set $X_1 = g_0(E_0) \in A_{\rm opt}$.

Let us show, by reduction to a contradiction, that the set X_1 is invariant w.r.t. arbitrary rotations around the center of the ball E_1. Indeed, if it is not invariant, then the set R of all rotations that leave X_1 invariant is different from the set of all rotations $SO(n)$. Hence, R is a proper closed subgroup of $SO(n)$. From the structure of $SO(n)$, it follows that there exists a 1-parametric subgroup R_1 of $SO(n)$ that intersects with R only in the identity transformation 1. This means that if $g \in R_1$ and $g \neq 1$, we have $g \notin R$, i.e., $g(X_1) \neq X_1$.

If $g(X_1) = g'(X_1)$ for some $g, g' \in R_1$, then we have $g^{-1}g'(X_1) = X_1$, where $g^{-1}g' \in R_1$. But such an equality is only possible for $g^{-1}g' = 1$, i.e., for $g = g'$. Thus, if $g, g' \in R_1$ and $g \neq g'$, then the sets $g(X_1)$ and $g'(X_1)$ are different. In other words, all the sets $g(X_1)$, $g \in R_1$, are different.

Since the family A is G_e-invariant, all the sets $g(X_1)$ for all $g \in R_1 \subseteq G_e$ also belong to A. For all these sets, the corresponding ellipsoid is $g(E_1)$, the result of rotating the ball E_1, i.e., the same ball $g(E_1) = E_1$. Hence, we have a 1-parametric family of sets contained in the ball E_1.

By applying appropriate affine transformations, we will get 1-parametric families of sets from A in an arbitrary ellipsoid. So, we have an $n(n+3)/2$-dimensional family of ellipsoids, and inside each ellipsoid, we have a 1-dimensional family of sets from A. Thus, A would contain a $\left(\dfrac{n(n+3)}{2}+1\right)$-parametric family of sets, which contradicts to our assumption that the dimension r of the family A is exactly $n(n+3)/2$.

This contradiction shows that our initial assumption was false, and for $r = n(n+3)/2$, the set X_1 is invariant w.r.t. rotations. Hence, with an arbitrary point x, the set X_1 contains all the points that can be obtained from x by arbitrary rotations, i.e., the entire sphere that contains x. Since X_1 is connected, X_1 is either a ball, or a ball from which a smaller ball was deleted.

The original set $X_0 = g_0^{-1}(X_1)$ is an affine image of this set X_1, and therefore, X_0 is either an ellipsoid, or an ellipsoid with an ellipsoidal hole inside. Q.E.D.

E. Resulting symmetries

The reason we decided to consider closeness is that we wanted to describe symmetries as (linear) transformations which preserve closeness.

From the previous section, we know that closeness is described by ellipsoids. By an appropriate selection of coordinates, every ellipsoid can be described as a sphere (it is sufficient to take eigenvectors of the corresponding quadratic form as coordinates).

Thus, in the appropriate coordinates, closeness is described by Euclidean distance

$$\rho(x,y) = \sqrt{(x_1-y_1)^2 + \ldots + (x_n-y_n)^2}.$$

Transformations preserving Euclidean distance are rotations and shifts. Since we are also allowing transformations which change the distance – but keep equal distances equal – we must also allow dilations $x \to \lambda \cdot x$.

So, the natural group of symmetries G_a is generated by shifts, rotations, and dilations.

VI. Main result: Mathematical justification of the spiral curriculum method

A. Back to the main problem

We have already argued that for the above "optimal education" problem, the natural group of symmetries G_a is generated by shifts, rotations, and dilations. In accordance with Theorem 1, to find the optimal curves, we must therefore describe all 1-dimensional orbits of subgroups of this group G_a.

B. How we can solve the corresponding mathematical problem

A 1-D orbit is an orbit of a 1-D subgroup. This subgroup is uniquely determined by its "infinitesimal" element, i.e., by the corresponding element of the Lie algebra of the group G. This Lie algebra if easy to describe. For each of its elements, the corresponding differential equation (that describes the orbit) is reasonably straightforward to solve.

C. General Result

One can see that in general, the resulting curve is indeed a spiral – although some limit cases are also possible, and a limit of a spiral may have as different shape.

D. 3-D case

Let us illustrate all the possibilities in the 3-D case. A generic 1-dimensional orbit of the corresponding group G_a is a *conic spiral* that is described (in cylindrical coordinates) by the equations $z = k\rho$ and $\rho = R_0 \exp(c\varphi)$. Its limit cases are:

- a *logarithmic* (Archimedean) *spiral*: a planar curve ($z = 0$) that is described (in polar coordinates) by the equation $\rho = R_0 \exp(c\varphi)$.
- a *cylindrical spiral*, that is described (in appropriate coordinates) by the equations $z = k\phi$, $\rho = R_0$.
- a *circle* ($z = 0$, $\rho = R_0$);
- a *semi-line* (*ray*); and
- a *straight line*.

E. 2-D case

In the 2-D case, the general case if a logarithmic spiral, and the limit cases are a circle, a semi-line, and a straight line.

F. Comment about the 3-D case

In the 3-D case, there is an alternative (slightly more geometric) way of describing 1-D orbits: by taking into consideration that an orbit, just like any other curve in a 3-D space, is uniquely determined by its curvature $\kappa_1(s)$ and torsion $\kappa_2(s)$, where s is the arc length measured from some fixed point.

The fact that this curve is an orbit of a 1-D group means that for every two points x and x' on this curve, there exists a transformation $g \in G$ that maps x into x'. Shifts and rotations do not change κ_i, they may only shift s (to $s + s_0$); dilations also change s to $s \to \lambda \cdot s$ and change the numerical values of κ_i. So, for every s, there exist $\lambda(s)$ and $s_0(s)$ such that the corresponding transformation turns a point corresponding. to $s = 0$ into a point corresponding to s.

As a result, we get functional equations that combine the two functions $\kappa_i(s)$ and these two functions $\lambda(s)$ and $s_0(s)$. Taking an infinitesimal value s in these functional equations, we get differential equations, whose solution leads to the desired 1-D orbits.

VII. Conclusion

We start with the problem of describing an optimal curve, i.e., an optimal order of presenting the material. As a result of our analysis, we show that for every reasonable optimality criterion, all the curves from the optimal family of curves are spirals (or limits of spirals).

Thus, we have indeed provided a mathematical justification for the spiral curriculum.

Acknowledgments

The author is thankful to the anonymous referees for valuable suggestions.

References

[1] R. Aló and O. Kosheleva, "Optimization Techniques under Uncertain Criteria, and Their Possible Use in Computerized Education", *Proceedings of the 25th International Conference of the North American Fuzzy Information Processing Society NAFIPS'2006*, Montreal, Quebec, Canada, June 3–6, 2006.

[2] J. Bruner, *The Process of Education*, Harvard University Press, Cambridge, Massachusetts, 1960.

[3] J. S. Bruner, *Toward a Theory of Instruction*, Belkapp Press, Cambridge, Massachusetts, 1966.

[4] J. S. Bruner, *The Relevance of Education*, Norton, New York, 1971.

[5] J. Bruner, *Going Beyond the Information Given*, Norton, New York, 1973.

[6] J. Bruner, *Child's Talk: Learning to Use Language*, Norton, New York, 1983.

[7] J. Bruner, *Actual Minds, Possible Worlds*, Harvard University Press, Cambridge, Massachusetts, 1986.

[8] J. Bruner, *Acts of Meaning*, Harvard University Press, Cambridge, Massachusetts, 1990.

[9] J. Bruner, *The Culture of Education*, Harvard University Press, Cambridge, Massachusetts, 1996.

[10] J. Bruner, J. Goodnow, and A. Austin, *A Study of Thinking*, Wiley, New York, 1956.

[11] H. Busemann, *The geometry of geodesics*, Academic Press, New York, 1955.

[12] A. Finkelstein, O. Kosheleva, and V. Kreinovich, "Astrogeometry, error estimation, and other applications of set-valued analysis", *ACM SIGNUM Newsletter*, 1996, Vol. 31, No. 4, pp. 3–25.

[13] A. Finkelstein, O. Kosheleva, and V. Kreinovich, "Astrogeometry: towards mathematical foundations", *International Journal of Theoretical Physics*, 1997, Vol. 36, No. 4, pp. 1009–1020.

[14] H. Gardner, "Jerome S. Bruner", In: J. A. Palmer (ed.) *Fifty Modern Thinkers on Education. From Piaget to the present*, Routledge, London, 2001.

[15] O. Kosheleva and R. Aló, "Towards Economics of Education: Optimization under Uncertainty", *Proceedings of the Second International Conference on Fuzzy Sets and Soft Computing in Economics and Finance FSSCEF'2006*, St. Petersburg, Russia, June 28 – July 1, 2006, pp. 63–70.

[16] S. Li, Y. Ogura, and V. Kreinovich, *Limit Theorems and Applications of Set Valued and Fuzzy Valued Random Variables*, Kluwer Academic Publishers, Dordrecht, 2002.