PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON
DYNAMICAL SYSTEMS AND DIFFERENTIAL EQUATIONS
May 18 – 21, 2000, Atlanta, USA pp. **224–228**

P-GROUPS APPLICATIONS IN GENETICS

BORIS KHOTS AND DMITRIY KHOTS

Des Moines, Iowa, USA

Abstract. This paper describes mathematical models for application of infinite-dimensional P-groups to Genetics. A monomorphic transformation is built from the Gene expression data to the infinite-dimensional group G of all homeomorphisms of the segment $[0,1]$ of real numbers into itself in a compact-open topology. Topological-algebraic properties of G connected with set of all Gene-Gene and Gene-Trait dependences are investigated.

1. **Introduction.** This paper describes mathematical models for application of infinite-dimensional P-groups ([1]) to Genetics. Initial data is Gene expression data. It is possible to use various forms of this data. The simpliest example of DNA/RNA data is a huge set of pairs (length, number). It is important that this data is sufficient. We have the set of seeds (in agriculture) or human (in medicine) samples. And for all of them we have the initial data - Gene expression data. Using our new mathematical models (based on infinite-dimensional topological - algebraic objects idea) we build set of all "effective" dependences between different subsets of Gene expression data for any one sample, any subset of samples, and, in particular, for full set of samples. The set of all "effective" dependences between different subsets of Gene expression data is a characterization of given set of samples. If there is, for example, actual Gene which is responsible for desired trait (for example, yield - in agriculture or disease - in medicine), model finds it. But for general situation this one Gene doesn't exist - and only the set of all "effective" dependences between different subsets of Gene expression data is responsible for desired trait. With using another set of our new mathematical models we are able to tell how we can change this characterization for approaching to desired trait. With using the third set of our new mathematical models we are able to predict and control the future desired trait.

2. **Infinite-dimensional groups G, H, H_0, $\widetilde{H_0}$.** A significant role in infinite-dimensional group concepts applications to Genetics play groups G, H, H_0 and $\widetilde{H_0}$. G is a group of all homeomorphisms of the segment $[0,1]$ onto itself. H is a group of all increasing homeomorphisms of the segment $[0,1]$ onto itself. H is a subgroup of G. H_0 is a group of all local analytical homeomorphisms t of the real line into itself with a stationary point 0: $t = t(x) = \sum_{k=1}^{\infty} t_k x^k$ with condition $t_1 > 0$. Group operation is superposition: $t_1 t_2 = (t_1 t_2)(x) = t_1(t_2(x))$. $\widetilde{H_0}$ is a group of all formal power series q with real coefficients: $q = q(x) = \sum_{k=1}^{\infty} q_k x^k$ with condition $q_1 > 0$. Group operation is superposition: $q_1 q_2 = (q_1 q_2)(x) = q_1(q_2(x))$. H_0 is a

1991 *Mathematics Subject Classification.* Primary: 22, 92; Secondary: 51.

subgroup of $\widetilde{H_0}$. Groups G, H, H_0 and $\widetilde{H_0}$ with compact-open topology are locally homeomorphic to some P-space, groups H_0 and $\widetilde{H_0}$ are P - groups, groups G and H are not P-groups. Now we will prove this statement, for example, for group H. First of all we remember the definitions of P-space and P-group.

Definition 1. Linear topological space R over a field of real (complex) numbers is called P - space, if the following conditions are satisfied:

$R \supset \bigcup_{i=1}^{\infty} R_i$, R_i are Banach spaces, R_i closed subspace in R_{i+1}, $i = 1, 2, \ldots$.

There is the projector $P_i : R \to R_i$ for each $i = 1, 2, \ldots$.

$x_n \to x$ $(x_n, x \in R, n \to \infty)$ if and only if $P_i(x_n) \to P_i(x)$ for each $i = 1, 2, \ldots (n \to \infty)$

In particular, Hilbert and Banach spaces are P-spaces.

Definition 2. Local topological group G is called P - group, if the following conditions are satisfied:

$G \subset R$ with $0 = e$, R is P - space , 0 is zero in R, e is unity element of G .

$G \cap R_i$ contains neighborhood of zero in R_i for each $i = 1, 2, \ldots$

$P_i(ab^{-1}) = P_i(P_i(a)P_i(b)^{-1})$ for each $i = 1, 2, \ldots$, $a, b \in G$

$|P_i(xy) - P_i(x) - P_i(y)||P_i(x)|^{-1} \to 0, |P_i(xy) - P_i(x) - P_i(y)||P_i(y)|^{-1} \to 0$, if $|P_i(x)| \neq 0; |P_i(y)| \neq 0; |P_i(x)| + |P_i(y)| \to 0$;

$|.|$ - norm in R_i for each $i = 1, 2, \ldots$, $x, y \in G$

In particular, classical finite - dimensional groups Lie and Birkhoff groups are P - groups. Groups of homeomorphisms of some classes of the manifolds are P - groups. Some groups of the formal power series are P - groups. General topological-algebraic properties of P-spaces and P-groups are described in [2].

Theorem 1. *The group H of all increasing homeomorphisms of the segment* $[0, 1]$ *onto itself with compact-open topology is locally homeomorphic to some P-space.*

Proof. The group operation $(. * .)$ in H is a superposition, i.e. for any $(f * g)(x) = f(g(x))$ for all $f, g \in H$. The compact-open topology in this case is the same as induced by the metric

$$d(f, g) = sup[d(f(x), g(x)) : x \in [0, 1]].$$

We define s as the space of all sequences $(x_{n,i}), i = 1, 2, ..., 2^n, n = 0, 1, 2, \ldots, (x_{n,i}) \in (0, 1)$, with the usual product topology. Thus

$$s = \prod_{n=0}^{\infty} \prod_{i=1}^{2^n} (0, 1)_{n,i}.$$

A basis for this topology of s is the collection of sets of the form

$$\prod_{n=0}^{\infty} \prod_{i=1}^{2^n} U_{n,i},$$

where $U_{n,i}$ is an open subset of $(0, 1)$ and for all but a finite number of the (n, i)'s, $U_{n,i}$ is $(0, 1)$. s is homeomorphic to the real Hilbert space l_2 ([3]). The real Hilbert space l_2 is a particular case of P - space. Let $(x_{n,i}) \subset s$. We will define $h \subset H$ associated with $(x_{n,i})$. Suppose that we have defined points

$$A_n = [0 = a_{n,0} < a_{n,1} < \cdots < a_{n,2^n} = 1],$$

$$B_n = [0 = b_{n,0} < b_{n,1} < \cdots < b_{n,2^n} = 1]$$

and define h such that $h(a_{n,i}) = b_{n,i}$ for $i = 0, 1, \ldots, 2^n$.

We extend the definition of h to a set of points $A_{n+1} \supset A_n$ onto $B_{n+1} \supset B_n$ with each of A_{n+1} and B_{n+1} having 2^{n+1} points. If n is odd, then let z_i be the midpoint of segment $[a_{n,i-1}, a_{n,i}]$ for $i = 1, 2, \ldots, 2^n$ and let $y_i = h(z_i) = x_{n,i}(b_{n,i} - b_{n,i-1}) + b_{n,i-1}$. If n is even, then let y_i be the midpoint of the segment $[b_{n,i-1}, b_{n,i}]$ and let $z_i = h^{-1}(y_i) = x_{n,i}(a_{n,i} - a_{n,i-1}) + a_{n,i-1}$. Then let

$$A_{n+1} = A_n \bigcup z_i, i = 1, 2, \ldots, 2^n$$

and

$$B_{n+1} = B_n \bigcup y_i, i = 1, 2, \ldots, 2^n.$$

Then we will have

$$A_{n+1} = [0 = a_{n+1,0} < a_{n+1,1} < \cdots < a_{n+1,2^n} = 1]$$

and

$$B_{n+1} = [0 = b_{n+1,0} < b_{n+1,1} < \cdots < b_{n+1,2^n} = 1]$$

with $h(a_{n+1,i}) = b_{n+1,i}$.

Proceeding in this way and letting $A = \bigcup_{n=1}^{\infty} A_n$ and $B = \bigcup_{n=1}^{\infty} B_n$, A and B will be dense in $[0,1]$ and $h[A] = B$ will be order preserving. Thus h will have a continuous extension to an homeomorphism g_h from H. If we let $F : s \to H$ be defined by $F((x_{n,i})) = g_h$, then F is the homeomorphism of s onto H. So, H is the group, which homeomorphic (and locally homeomorphic in particular) to P-space s. Moreover, H is locally Banach group. □

Theorem 2. *The group H of all increase homeomorphisms of the segment $[0,1]$ onto itself with compact-open topology is not P-group.*

Proof. Let $K_{g,d}$ is the set of increase homeomorphisms f of the segment $[0,1]$ onto itself with condition $f(x) = x$ for $x \in [0,g] \bigcup [d,1], 0 < g < d < 1$. $K_{g,d}$ is subgroup from H. If U is any open neighborhood of the unity element e in H (which is a $f(x) = x$ for all $x \in [0,1]$), there are g and d with sufficiently small $d - g$ such that $K_{g,d} \subset U$. The other words, H contains "small" subgroups. So, we can conclude, that H is not locally isomorphic to some P - group, because statement "Each P - group has a "canonical coordinates" ([4]) contradicts existing of "small" subgroups in H. □

3. The absence of "small" subgroups was the main part of the positive solution of the classical finite-dimensional Fifth Hilbert problem ([5]). So, we have negative solution of Fifth Hilbert problem for infinite-dimensional case ([6]).

Theorem 3. *Let L be a subset of H containing all elements without internal stationary points. For any $l \in L$ there exists a homeomorphism A_l of the interval $(0,1)$ onto a real line such that the isomorphism AA_l of the group G onto the group GG of all homeomorphisms of a line, generated by this homeomorphism, transfers the element l*

Proof. Let $l(x) > x$ for $x \in (0,1)$. If $x_0 \in (0,1)$, then let $x_n = l^n(x_0), n = 0, -1, +1, -2, +2, \ldots$. Let A_l^0 be a homeomorphism of the segment $[x_0, x_1]$ onto the segment $[0,1]$ such that $A_l^0(x_0) = 0$, $A_l^0(x_1) = 1$.

Let now

$A_l(x) = A_l^0(x)$, if $x \in [x_0, x_1]$ and
$A_l(x) = n + A_l^0(l^{-n}(x))$, if $x \in [x_n, x_{n+1}], n = 0, -1, +1, -2, +2, \ldots$.

Now we define

$$AA_l : G \to GG : AA_l(f) = A_l f A_l^{-1}, f \in G,$$

considering $f$first in $(0,1)$, then naturally extending it to $[0,1]$. We have $AA_l(l) = p_1$. In the same way, we can consider the case when for all $x \in (0,1) l(x) < x$. □

Corollary 1. *Each element $h \in H$ belongs to a one-parameter subgroup of elements of H.*

Corollary 2. *Each element $l \in L$ belongs to a one-parameter subgroup of elements of L.*

Theorem 4. *Let $\widetilde{H} = \{h = h(x) = \sum_{k=1}^{\infty} h_k x^k\}$ be a linear space of all formal power serieses of one real variable. Let $\widetilde{\Phi}$ be the embedding of $\widetilde{H_0}$ into $\widetilde{H}$: $\widetilde{\Phi}$: $h = h(x) \in H_0 \to h(x) - x \in \widetilde{H}$. Let $g(t) = g(t,x)$ be a curve in $\widetilde{H_0}$, and let $\widetilde{\Phi}(g(t)) = \sum_{k=1}^{\infty} g_k(t) x^k$. We will say that $g(t)$ is differentiable with respect to t, if $g_i(t)$ are differentiable functions, $i = 1, 2, \ldots$. An element $a = \sum_{k=1}^{\infty} a_k x^k \in \widetilde{H}$ will be called tangential vector to $g(t)$ at t_1, if $\frac{dg_i}{dt}(t = t_1) = a_i, i = 1, 2, \ldots$. For any $a \in \widetilde{H}$ and for any real t we have a one-parameter subgroup $g(t) = \sum_{k=1}^{\infty} g_k(t) x^k$ of $\widetilde{H_0}$ with the tangential at zero vector a.*

Proof. Let $g(t) = g(t,x)$ be a differentiable one-parameter subgroup in $\widetilde{H_0}$ with a tangential at zero vector a. As $g(t+s) = g(t)g(s)$, then the tangencial vector to $g(t)$ at t can be obtained by the formula:

$$\frac{d\widetilde{\Phi}(g(t))}{dt} = F_{g(t)}(a).$$

Thus a differentiable one-parameter subgroup $g(t)$ in $\widetilde{H_0}$ with a tangential at zero vector a satisfies the following differential equation:

$$\frac{d\widetilde{\Phi}(g(t))}{dt} = \left(\frac{\partial \widetilde{\Phi}(g(t))}{\partial x} + 1 \right) * a$$

($*$ denotes ordinary multiplication) with the initial condition

$$\widetilde{\Phi}(g(0)) = 0.$$

Inversely, a solution $g(t,x)$ of this equation with same initial condition defines a differentiable one-parameter subgroup in $\widetilde{H_0}$ with the tangential at zero vector a.

If

$$a = \sum_{k=1}^{\infty} a_k x^k, g(t,x) = \sum_{k=1}^{\infty} g_k(t) x^k,$$

then we have following system of differential equations:

$\dot{g}_1 = g_1 a_1$
$\dot{g}_2 = g_1 a_2 + 2 g_2 a_1$
$\dot{g}_3 = g_1 a_3 + 2 g_2 a_2 + 3 g_3 a_1$
...
...

$\dots$

$\dot{g}_k = g_1 a_k + 2g_2 a_{k-1} + \dots + k g_k a_1$

$\dots$

$\dots$

$\dots$

with the initial conditions: $g_1(0) = 1, g_2(0) = g_3(0) = \dots = 0$.

A solution of this system with these initial conditions can be represented as $g(t) = e^{At} g_0$, where g_0 is a column $(1, 0, 0, \dots)$ and infinite matrix A has a k-line $(a_k, 2a_{k-1}, \dots, ka_1, 0, 0, \dots)$. We have: $g_1(t) = e^{a_1 t}$ and

$$g_k(t) = a_k t + \sum_{n=2}^{\infty} t^n/n! \sum_{k \geq j_1 \geq j_2 \geq \dots \geq j_{n-1} \geq 1} j_1 j_2 \cdot \dots \cdot j_{n-1} \cdot a_{k-j_1+1} a_{j_1-j_2+1} \cdot \dots \cdot a_{j_{n-2}-j_{n-1}+1} a_{j_{n-1}}.$$

So, for any $a \in \widetilde{H}$ and for any real t we have a one-parameter subgroup $g(t) = \sum_{k=1}^{\infty} g_k(t) x^k$ of $\widetilde{H_0}$ with the tangential at zero vector a. □

Theorems 3 and 4 give us two of numerous topological-algebraic properties of the groups G, H, H_0 and $\widetilde{H_0}$ which are important for Genetic applications.

4. An example of monomorphism from Gene expression data to the group H. Let Gene expression data be: (m_i, n_i) for $i = 0, 1, 2, \dots, L$. Let $m_{i+1} - m_i = m$, so that $m_i = im$ for $i = 0, 1, 2, \dots, L$. Let $n_i > 0$ for $i = 1, 2, \dots, L$ and $n_0 = 0$. We now let $A(lm) = \sum_{i=0}^{l} n_i$, $l = 0, 1, 2, \dots, L$. and $B(l/L) = A(lm)$. Thus we have: $B(0) = 0$. Now let $f(l/L) = B(l/L)/B(1)$, so that $f(0) = 0$ and $f(1) = 1$. We have: $f(a) > f(b)$ if and only if $a > b$. Therefore: $f(x) = xL \frac{n_{l+1}}{\sum_{i=0}^{L} n_i} + \frac{\sum_{i=0}^{l} n_i - l n_{l+1}}{\sum_{i=0}^{L} n_i}$ with $x \in [l/L, (l+1)/L]$ and $l = 0, 1, 2, \dots, L-1$.

5. We'd like to thank Ilya Markevich for his help and Professors Shouchuan Hu and Joshua Du for their invitation to Year 2000 International Conference on Dynamical Systems and Differential Equations.

REFERENCES

[1] B. S. Khots, *Groups of local analytical homeomorphisms of line and P-groups*, Russian Math Surveys, v. XXXIII, # 3 (201), Moscow, London, 1978, 189–190.

[2] B. S. Khots, *P-spaces*, Lecture Notes in Pure and Applied Siences, Marcel Dekker, New York, July 2000, 227–232.

[3] R. D. Anderson and R. H. Bing, *A complete elementary proof that Hilbert space is homeomorphic to the countable infinite product of lines*, Bulletin of the American Mathematical Society, 74(1968), 771–792.

[4] B. S. Khots, *P-groups and P-algebras Lie*, Soviet Math (Iz VUZ), dep # 2265-B87, Kazan, Allerton, New York, 1988, 16 pp.

[5] D. Hilbert, *Mathematical Problems*, Speech presented at the II International Congress of Mathematicians, Paris, France, on August 8, 1900, Gessammelte Abhandlungen, v. III, 1935, 290–329.

[6] B. S. Khots, *Fifth Hilbert problem and infinite-dimensional P-groups*, Talk at the International Congress of Mathematicians, Berlin, Germany, August 1998, Proceedings.

E-mail address: bkhots@cccglobal.com; dkhots@blue.weeg.uiowa.edu