Parametric S-tree Method and Its Generalizations

Karen M. Bekarian
Department of Applied Mathematics,
Yerevan State University,
A. Manugian 1, Yerevan, Armenia

Anahit A. Melkonian
Theoretical Department,
Yerevan Physics Institute,
Alikhanyan Brothers str.2,
Yerevan, 375036, Armenia

A parametric approach is developed to the method of S-tree diagrams and its generalization for investigation of the hierarchical substructure of N-body nonlinearly interacting systems (e.g., clusters of galaxies). The introduction of a parametric function allows us to take into account the individual features of the particles which do not have a direct influence on the dynamics of the system but are crucial for analyzing the output data. An algorithm is proposed for the latter based on the construction of parametrical subgroups.

A generalization of the S-tree scheme is presented as well. This enables us to perform parallel substructure analysis with respect to various dependent or independent parameters, thus using fully the initial information of a system.

1. Introduction

Computers are powerful tools for the investigation and numerical simulation of complex systems, including information about their hierarchical structures [1]. A class of such problems, namely N-body gravitating configurations, has an essential role in the understanding of various structures in the universe such as star clusters, galaxies, clusters of galaxies, and so forth.

In the present paper we suggest an algorithm for a statistical method of investigating the hierarchical structure of gravitating N-body systems. The large-scale distribution of galaxies in the universe and the hierarchical structure of clusters of galaxies are two of the key problems of extragalactic astrophysics [2, 3]. This is important because the relaxation times of clusters of galaxies are of the order or exceed the age of the universe, those systems recall their initial conditions. Hence, the study of the hierarchical properties of systems may directly indicate the mechanisms of their formation and evolution.

Complex Systems, **11** (1997) 323–331;

Among the statistical methods used for investigating the hierarchical properties of galaxy clusters are the two-point correlation functions, Lee statistics, topological measures, wavelets and so on, for review see [4, 5].

The studies already performed enable us to conclude that many galaxy clusters may include subgroups with vivid dynamical peculiarities. The existence of the subgroups is also apparent from X-ray radiation data of the clusters. The development of more informative methods for the analysis of the problem is of particular interest, given the increase of the possibilities of observational technique.

The S-tree diagram method [5, 6] is based on the idea of introducing a *degree of boundness* between the particles and their sets and therefore self-consistently using both positional and kinematic information about the particles. The idea is therefore that the parameters of particles in a physically interacting system, such as the components of their 3-velocities and coordinates, have to correlate with each other. On the contrary, particles that do not satisfy the correlation have to be chance projections either in front of or behind the physical system. The problem is therefore to reveal the corresponding function of interaction, and to define its "degree," since obviously, various particles can have different degrees of interaction. This idea was realized *via* the S-tree technique, thus revealing the physically interacting system among the projected set of galaxies, and moreover establishing the subgroups with various degrees of interaction. This is achieved by considering the geometrical properties of the phase space of the system, namely, its two-dimensional curvature. The N-body system as a hamiltonian system is thus being transferred *via* the Maupertuis principle to the phase space and its dynamics are being described by the flow of geodesics on a riemannian manifold with its metric determined by the potential of interaction [7]. The S-tree method has already been used for the study of ESO Key Program data on the nearby Abell clusters [8] and many other observational data. In particular, the ESO data enabled the determination of a new class of galaxy configurations as dynamical entities given their peculiar properties.

The S-tree technique has been developed not only for point approximation, but also to account for the extension of particles [9], and to include certain generalized schemes [10].

The parametric approach of the S-tree method developed in the present paper allows us to consider individual properties of the particles, which are crucial for the analysis of the output information on the hierarchical structure of the system. For example, in the case of the galactic systems such a quantity can be the morphological class, which though does not influence the dynamics directly, has an essential impact on analyzing the output. The aim is therefore to have the algorithm include the parametric representation in the hierarchic structure intrinsically.

Then we extend the approach to generalized schemes that enable the algorithm to use different functions of boundness self-consistently, thus obtaining a more complete picture about the hierarchical structure of the system. We first introduce the necessary definitions and then move on to the algorithm.

2. S-tree and its generalization

We now briefly recall the main steps of the S-tree method. As mentioned previously, the key idea of this method is the introduction of the concept of the degree of boundness ρ.

Consider a set of N points:

$$X = \{x_1, \ldots, x_N\},$$

the function P

$$P : X \times X \to R_+ \quad \text{and} \quad \rho \in \mathrm{R}_+.$$

The basic definitions follow.

Definition 2.1. We say that $\forall x \in X$ and $\forall y \in X$ are *ρ-bounded*, if $P(x, y) \geq \rho$.

Definition 2.2. We say that $U \subset X (U \neq \emptyset)$ is a *ρ-bounded subgroup*, if:

1. $\forall x \in U$ and $\forall y \in \bar{U} \Rightarrow P(x, y) < \rho$;
2. $\forall x \in U$ and $\forall y \in U \quad \exists x = x_{i_1}, x_{i_2}, \ldots, x_{i_k} = y$,

and that $P(x_{i_l}, x_{i_{l+1}}) \geq \rho; \quad \forall l = 1, \ldots, k-1$.

Definition 2.3. We say that $U_1, \ldots, U_d$ is the *distribution* of the set X *via* ρ-bounded groups, if:

1. $\cup_{i=1}^{d} U_i = X$;
2. $i \neq j \quad (i, j = 1, \ldots, d) \Rightarrow U_i \cap U_j = \emptyset$;
3. $U_d (i = 1, \ldots, d)$ is a ρ-bounded group.

It is possible for the function P to consider different physical quantities. In [5], various quantities such as the energy, the potential, perturbations of potential, momentum, and so on, are listed and discussed. For example, the choice of the mutual distance of the particles to define the subgrouping, obviously, is equivalent to the corresponding correlation functions (spatial or angular, related with each other simply *via* the Limber equation), which is a rather incomplete characteristic for that aim. It can be shown that, at least for astrophysical problems, among

the most informative ones is the riemannian curvature of the configuration space [7], which determines the behavior of close geodesics, as known from basic courses on classical mechanics [11].

So, by the S-tree algorithm we obtained the distribution of the N-body system for any given function P and $\forall\rho$. This distribution will satisfy Definitions 2.1, 2.2, and 2.3. The final result can be represented either through tables or by graphs (S-tree).

Now we aim to use the S-tree scheme for different boundness criteria simultaneously.

We have t different functions $P_1, \ldots, P_t$, which satisfy the definitions introduced in the method of S-diagrams and that describe the degree of interaction of the system's members. To each P_α we correspond a matrix D_α, where $D_\alpha = (d_{ij}^\alpha), \alpha = 1, \ldots, t$ [10].

Consider the following matrix D:

$$D = (\bar{d}_{ij}), \quad i, j = 1, \ldots, N,$$

where

$$\bar{d}_{ij} = (d_{ij}^1, d_{ij}^2, \ldots, d_{ij}^t), \quad i, j = 1, \ldots, N,$$

which contains the entire information of the system, that is, it uses all P_t functions. For example, the previously mentioned riemannian curvature is a function of the coordinates, velocities of all the particles, their masses, potential of interaction, its derivatives, and so on.

Each function $P_\alpha, \alpha = 1, \ldots, t$ corresponds to a definite vector of boundness $\bar{\rho}_\alpha = (\rho_1^\alpha, \ldots, \rho_{Q_\alpha}^\alpha)$ and Q_α satisfies the conditions of the choice of the vector of boundness [5].

We construct the vector $\bar{\rho}_D = (\bar{\rho}_1^D, \ldots, \bar{\rho}_{Q_D}^D)$, where

$$\bar{\rho}_k^D = (\rho_{k_1}^1, \ldots, \rho_{k_t}^t); \quad k = 1, \ldots, Q_D; \quad k_\alpha = 1, \ldots, Q_\alpha, \quad \alpha = 1, \ldots, t.$$

The vector $\bar{\rho}_D$ is obtained with the help of all possible variations of the components of $\bar{\rho}_k^D (\alpha = 1, \ldots, t; k_\alpha = 1, \ldots, Q_\alpha)$, and it is obvious that $Q_D \le ((N^2 - N)/2 + 1)^t$.

Taking into account the fact that different P_α functions contain different information about the system, we introduce the concept of the "degree of influence" E_α for each of the functions under consideration. We will conditionally understand E_α as the number of physical quantities (distance, potential of the interaction and its derivatives, and so on), existing in the P_α formulae.

The next step of the algorithm is the transition from matrix D to matrix D_u with the help of a fixed current vector $\bar{\rho}_k^D$

$$D_u = (\bar{u}_{ij}); \quad i, j = 1, \ldots, N,$$

where

$$\bar{u}_{ij} = (u_{ij}^1, \ldots, u_{ij}^t); \quad u_{ij}^\alpha (\alpha = 1, \ldots, t)$$

is defined in the following way:

$$u_{ij}^{\alpha} = \begin{cases} E_{\alpha} & \text{if} \quad d_{ij}^{\alpha} \geq \rho_{k_{\alpha}}^{\alpha} \\ 0 & \text{if} \quad d_{ij}^{\alpha} < \rho_{k_{\alpha}}^{\alpha}. \end{cases}$$

Thus, the matrix D_u contains information on the degree of influence of each criterion too. Then we construct the matrix D_v in the following way:

$$D_v = (v_{ij}); \quad i,j = 1,\ldots,N,$$

where $v_{ij} = \sum_{\alpha=1}^{t} u_{ij}^{\alpha}$.

The matrix D_v characterizes the quantitative degree of boundness. At the same time we also take into account information on the ρ-boundness of particles, because matrix D_v is defined with the help of matrix D_u. According to the S-tree method we then construct the vector of boundness $\bar{\mu}$ for the matrix D_v:

$$\bar{\mu} = (\mu_1,\ldots,\mu_T).$$

It is obvious that $T \leq ((N^2 - N)/2 + 1)$.

Note, that while estimating Q_D in this particular case we are taking into account the symmetry of the initial matrices.

Thus the N-body system, according to the generalized S-tree method, is split into $\tilde{\rho}$-bounded subgroups, where $\tilde{\rho} = (\bar{\rho}_k^D;\mu_l); (l = 1,\ldots,T)$ in the following way. We obtained matrix D_u from matrix D with the help of some fixed $\bar{\rho}_k^D$, afterwards we get matrix D_v out of D_u using the described method. The last step of the algorithm is the distribution of the system into subgroups *via* matrix D_v for the fixed μ_l using the S-tree method.

To get complete information of the distribution it is necessary to apply the algorithm for all possible pairs $(\bar{\rho}_k^D;\mu_l)$, where $\bar{\rho}_k^D$ and μ_l accordingly are the components of the vectors $\bar{\rho}_D$ and $\bar{\mu}$.

The time L, which is needed for the realization of this algorithm, is $L \leq ((N^2 - N)/2 + 1)^{t+1}$ and for $t \ll N$ no difficulties can arise [12].

3. Parametrical approach

As mentioned previously, the S-tree method is splitting the system into subgroups according to the boundness parameter ρ.

The application of a parametrical approach to the S-tree method has two main goals.

1. To have a convenient way of representing output information of the S-tree algorithm with respect to individual properties of the particles.

2. To have more complete and precise information about the substructure in the output information.

Two possible versions of parametrical approach are considered.

1. After obtaining the distribution of the system by subgroups (using the S-tree method), we introduce a parameter describing certain properties of particles independent of the distribution.

   Consider the set of N discrete points

   $$X = \{x_1, \ldots, x_N\}$$

   and the function P, where

   $$P : X \times X \to R_+.$$

   Let us introduce the following function

   $$B : X \to X \times R; \quad B(x_i) = (x_i, b_i).$$

   We denote $(x_i, b_i) = x_i^{b_i}$ for convenience, where b_i is a parameter of P depending on the system under consideration.

2. The suggested method is applicable if these constraints are satisfied simultaneously:

   (a) $k \neq N$, where k is the number of subgroups;

   (b) $k \neq 1$ and $\exists l_i, l_j$, such that $t_{l_i} > 1$ and $t_{l_j} > 1$, where t_{l_c} is the number of particles in the l_cth subgroup, $l_c = 1, \ldots, k$.

   For any fixed x_i the following sum is considered:

   $$P_i^d = \Sigma P(x_i, x_q),$$

   where d is the number of $U_{j,\ldots,k}$ subgroups and the summation is performed for all $q = j, \ldots, k$ (accordingly, $x_j, \ldots, x_k$ forms the subgroup $U_{j,\ldots,k}$ and $x_i \not\in U_{j,\ldots,k}$).

   The sum

   $$P_i^0 = \Sigma P(x_i, x_q)$$

   (the summation is performed for all $q = j, \ldots, k$, where $x_i, x_j, \ldots, x_k$ form the subgroup $U_{i,\ldots,k}$) is considered as well. The concept of parametric boundness is introduced in Definition 3.1.

Definition 3.1. If P_i^0 is more than any other P_i^d, than x_i is called a *parametrically real* (p.r.) member of the subgroup; otherwise it is called *parametrically imaginary* (p.i.).

The final result of the output information in the parametric approximation is presented in the following way. If x_i is a p.r. member of subgroup U_{l_i}, we denote x_i as $x_i^{l_i}$; if x_i is a p.i. member of subgroup U_{l_i}, we denote x_i as $x_i^{l_j}$, where l_j is the number of the subgroup U_{l_j} with elements $x_j, \ldots, x_k$ on which the function $P_i^d (d = l_j)$ obtains its maximal value.

4. Generalized parametric scheme

Now we consider the corresponding generalizations for both versions of the parametrical approach.

I. Consider the set of N points: $X = \{x_1, \ldots, x_N\}$ and the function P, where

$$P : X \times X \to R_+.$$

For the obtained distribution (*via* the S-tree method) we introduce a multidimensional parameter, each component of which shows a definite peculiarity of the element, independent of the distribution.

Consider the following function

$$B : X \to X \times R^k, k = 1, 2, \ldots$$
$$B(x_i) = (x_i, \bar{b}_i),$$

where $\bar{b}_i = (b_i^1, \ldots, b_i^k)$.

We introduce the notation $(x_i, \bar{b}_i) = x_i^{\bar{b}_i}$ for convenience, where $b_i^j; j = 1, \ldots, k$ is not a variable of P.

In this particular case $\bar{b}_i$ is a multidimensional parameter, each component of which is normalized in accord with the physical content of the corresponding component and the system under investigation.

II. The initial limitations of the proposed method fully correspond to the limitations of the parametrical approach.

We fix some value x_i and consider all possible sums $P_i^d = \Sigma P(x_i, x_q)$, where d is the number of the subgroup $U_{j,\ldots,k}$ and the summation is taken for all the $q = j, \ldots, k$ (accordingly $x_j, \ldots, x_k$ forms the group $U_{j,\ldots,k}$ and $x_i / \in U_{j,\ldots,k}$).

It is obvious that the vector $\bar{P}_i^{\,d} = (P_i^{d_1}, \ldots, P_i^{d_{k-z}})$, where $k - z$ is the number of those subgroups for which the above mentioned limitations have been carried out.

We construct the vector $\tilde{P}_i^d$ from vector $\bar{P}_i^d$, where $\tilde{P}_i^d = (P_i^{d_{j_1}}, \ldots, P_i^{d_{j_{k-z}}})$, the components of which are the components of the vector $\bar{P}_i^d$ situated in decreasing order.

Let us consider the sum

$$P_i^0 = \sum P(x_i, x_q),$$

(the summation is taken by all $q = j, \ldots, k$), where $x_i, x_j, \ldots, x_k$ forms the group $U_{i,\ldots,k}$.

The concept of the generalized parametric boundness is introduced in Definition 4.1.

Definition 4.1. If P_i^0 is greater then all components of the vector $\bar{P}_i^d$, we say that x_i is a *parametrically real* (p.r.) member of the group $U_{i,\ldots,k}$, otherwise x_i is a *parametrically imaginary* (p.i.) member of the group $U_{i,\ldots,k}$.

The final representation of the output information in terms of the generalized parametric approach is made in the following way:

- if x_i is a p.r. member of the group U_{l_i}, denote x_i as $x_i^{l_i}$;
- if x_i is a p.i. member of the group U_{l_i}, denote x_i as $x_i\bar{l}_j$, where $\bar{l}_j = (l_{j_1}, l_{j_2}, \ldots)$. l_{j_r} are numbers of those subgroups, where $P_i^{d_{j_r}} > P_i^0$; $r = 1, \ldots, k - z$; $d_{j_r} = l_{j_r}$.

5. Conclusion

In this paper a parametric approach to the S-tree diagram method has been suggested. The proposed approach is useful for revealing the optimal distribution of the system under investigation [9].

The following developments were presented.

- A generalization of the discrete S-tree method.
- A generalization of the parametric approach to the S-tree method, in the case of multiparametric functions.

The methods described are more informative since they use all the information of the system in the input, while they enable the output to have a hierarchical structure depending not only on the kinematic and positional data, but also on the parameters having no direct influence on the dynamics of the system.

For example, in the case of clusters of galaxies the algorithm will use data from the catalogues not only on the coordinates and redshifts of galaxies but also the morphological, color, spectral index, and other information self-consistently in the output information on the hierarchic structure of the systems.

The hierarchical structure of N-body nonlinearly interacting particles is only one indication of their complex dynamics [13, 14].

Acknowledgments

M. A. is grateful to Prof. S. Randjbar-Daemi for hospitality in ICTP. We thank Prof. V. Gurzadyan for valuable discussions and A. Sanasarian for assistance.

References

[1] R. H. Landau and M. J. Paez, *Computational Physics—Problem Solving with Computers* (John Wiley and Sons, New York, 1997).

[2] P. Coles and F. Lucchin, *The Origin and Evolution of Cosmic Structure* (Wiley, Chichester, 1995).

[3] E. Escalera, A. Biviano, M. Girardi, *et al.*, "Structures in Galaxy Clusters," *Astrophysical Journal*, **423** (1994) 539–552.

[4] P. J. E. Peebles, *Principles of Physical Cosmology* (Princeton University Press, Princeton, 1993).

[5] V. G. Gurzadyan and A. A. Kocharyan, *Paradigms of the Large-Scale Universe* (Gordon and Breach, New York, 1994).

[6] V. G. Gurzadyan, V. V. Harutyunyan and A.A. Kocharyan, "A Method for Defining Structures in the Distribution of Galaxies," *Astronomy and Astrophysics*, **281** (1994) 964–966.

[7] V. G. Gurzadyan and G. K. Savvidy, "Collective Relaxation of Stellar Systems," *Astronomy and Astrophysics*, **160** (1986) 203–210.

[8] V. G. Gurzadyan and A. Mazure, "Galaxy Associations as Possible Common Features of Galaxy Clusters," *Monthly Notices of the Royal Astronomical Society*, **295** (1998) 177–187.

[9] K. M. Bekarian and A. A. Melkonian, "Investigation of the Substructure of Clusters of Galaxies by the Method of S-Tree Diagrams: Non-Point Approximation," *Astrophysics*, **40** (1997) 275–279.

[10] K. M. Bekarian and V. G. Gurzadyan, "Study of the Substructure in Galaxy Distribution: Further Development of S-Tree Techinique," Preprint ICTP-416 (1993) (ICTP, Trieste).

[11] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Nauka, Moscow, 1989).

[12] E. G. Coffman, *Computer and Job-Shop Scheduling* (Nauka, Moscow, 1984).

[13] H. J. De Vega, N. Sanchez and F. Combes, "Self-Gravity as an Explanation of the Fractal Structure of the Interstellar Medium," *Nature*, **383** (1996) 56–58.

[14] V. G. Gurzadyan, A. Melkonian, and K. Oganessian, "Feigenbaum Bifurcations and Stellar Dynamics," *Nuovo Cimento*, **112B** (1997) 309–311.