Topological Language for Classifying Linear Chord Diagrams via Intersection Graphs

Authors : Michela Quadrini

Abstract : Chord diagrams occur in mathematics, from the study of RNA to knot theory. They are widely used in theory of knots and links for studying the finite type invariants, whereas in molecular biology one important motivation to study chord diagrams is to deal with the problem of RNA structure prediction. An RNA molecule is a linear polymer, referred to as the backbone, that consists of four types of nucleotides. Each nucleotide is represented by a point, whereas each chord of the diagram stands for one interaction for Watson-Crick base pairs between two nonconsecutive nucleotides. A chord diagram is an oriented circle with a set of n pairs of distinct points, considered up to orientation preserving diffeomorphisms of the circle. A linear chord diagram (LCD) is a special kind of graph obtained cutting the oriented circle of a chord diagram. It consists of a line segment, called its backbone, to which are attached a number of chords with distinct endpoints. There is a natural fattening on any linear chord diagram; the backbone lies on the real axis, while all the chords are in the upper half-plane. Each linear chord diagram has a natural genus of its associated surface. To each chord diagram and linear chord diagram, it is possible to associate the intersection graph. It consists of a graph whose vertices correspond to the chords of the diagram, whereas the chord intersections are represented by a connection between the vertices. Such intersection graph carries a lot of information about the diagram. Our goal is to define an LCD equivalence class in terms of identity of intersection graphs, from which many chord diagram invariants depend. For studying these invariants, we introduce a new representation of Linear Chord Diagrams based on a set of appropriate topological operators that permits to model LCD in terms of the relations among chords. Such set is composed of: crossing, nesting, and concatenations. The crossing operator is able to generate the whole space of linear chord diagrams, and a multiple context free grammar able to uniquely generate each LDC starting from a linear chord diagram adding a chord for each production of the grammar is defined. In other words, it allows to associate a unique algebraic term to each linear chord diagram, while the remaining operators allow to rewrite the term throughout a set of appropriate rewriting rules. Such rules define an LCD equivalence class in terms of the identity of intersection graphs. Starting from a modelled RNA molecule and the linear chord, some authors proposed a topological classification and folding. Our LCD equivalence class could contribute to the RNA folding problem leading to the definition of an algorithm that calculates the free energy of the molecule more accurately respect to the existing ones. Such LCD equivalence class could be useful to obtain a more accurate estimate of link between the crossing number and the topological genus and to study the relation among other invariants.

Keywords : chord diagrams, linear chord diagram, equivalence class, topological language