Additive Combinatorics, Math 8803/4803

January 9, 2011

- **Instructor:** Ernie Croot
- **Office:** 103 Skiles.
- **Office Hours:** TBA
- **Meeting Time and Place:** MWF 1:05 - 1:55, Skiles 154.
- **Textbook:** I will be using the Tao-Vu book *Additive Combinatorics*, 2nd edition (ISBN-10 0521136563, ISBN-13 978-0521136563) to assign homework and to occasionally lecture from. I also plan to use several other sources, and I will post links or pdf files to notes when I do so.
- **Grade:** Basically, if you show up for most of the classes and turn in the homeworks (and at least make a good attempt) I will give you an A. There will be no in-class exams in this course. Near the end of the semester I *may* ask people to present lectures on a topic of their choice (within limits).
- **Content:** Mainly, I will focus on tools rather than big theorems. (In many cases if one understands the tools, and roughly how they are used in the proof of a big theorem, one can reproduce the proof oneself anyways.)

  In the beginning I plan to focus on elementary properties of sumsets, proving such results (tools) as the Cauchy-Davenport Theorem; the Balog-Szemeredi-Gowers Theorem; a structure theorem due to myself, Ruzsa and Schoen about sumsets; various properties of the "energy of sumsets"; the Brun Sieve and Rosser's Sieve; Schnirelmann's theorem and applications; the Plunnecke-Ruzsa Inequality; various "Ruzsa-type" inequalities; Behrend's construction; Szemeredi's elementary proof of Roth's Theorem; Solymosi's sum-product inequality (for the complex numbers); the Bourgain-Katz-Tao sum-product inequality (with applications); and possibly several more.

Then, I will switch gears somewhat and start talking about Fourier methods, and in particular will present: Roth's theorem on three-term progressions; the Szemeredi and Heath-Brown refinement; the Littlewood-Offord Inequality and various results about subset sums; and Ruzsa's proof of Freiman's theorem.

If time permits, I will discuss higher-order Fourier methods as developed by Gowers (and possibly ergodic methods), including an overview of his proof of Szemeredi's theorem for four-term arithmetic progressions.

Course Syllabus

Professor: Dr. Christine Heitsch **Office:** Skiles 226 **Phone:** (404) 894 - 4758

Email: heitsch@math.gatech.edu **Webpage:** http://www.math.gatech.edu/~heitsch

Office Hours: Wednesday 10AM – 11PM, and by appointment.

Lectures: Mon, Wed, Fri 9:05 – 9:55 in Skiles 270.

Website: http://www.math.gatech.edu/~heitsch/discmathbio.html

Textbook: There is no required text for this course; the interface of combinatorics with molecular biology is still a research frontier, and there is not yet a discrete mathematical biology textbook available. Instead, readings will be assigned from primary sources (research articles available online).

You are encouraged to consult the following recommended texts on reserve in the library. They are three computational molecular biology books that include significant mathematical content, as well as one discrete mathematics book, and will be helpful as background / supplemental reading.

- *Introduction to computational biology : maps, sequences and genomes* by Michael S. Waterman.
- *Computational molecular biology : an algorithmic approach* by Pavel A. Pevzner.
- *Algorithms on strings, trees, and sequences : computer science and computational biology* by Dan Gusfield.
- *Discrete mathematics and its applications* by Kenneth H. Rosen.

Course Description: This course will focus on the interaction between combinatorics and molecular biology, with particular emphasis on DNA and RNA sequences. This is a rapidly developing research area, accessible to advanced undergraduates and introductory graduate students, at the intersection of the mathematical, computational, and biological sciences.

Course Topics: We will cover recent combinatorial results, as well as current open problems needing new advances in discrete mathematics and its applications, for the following topics in computational molecular biology:

DNA — fragment assembly, sequence alignment and comparison, motif discovery, optical mapping, word design, and chromosome rearrangements.

RNA — comparative sequence analysis, secondary structure prediction, inverse folding, pseudoknots, and three-dimensional alignment and comparison.

If time permits, we will touch on other topics such as phylogenetics and biological networks.

Prerequisites: For undergraduates, Math 3012 or consent of instructor.

For graduate students, a reasonable background in discrete mathematics, equivalent to Math 3012, or the willingness to learn independently.

Familiarity with algorithms, at the level of CS 3510, will be very helpful but is not required.

Note that only a very general (high school) biology background is expected; we will cover the necessary biochemical topics as needed.

Grading Scheme: Grades will be based on the following class components:

60% Course project (30% quality of the work and 30% quality of the oral presentation)

30% Homework

10% Class participation and assigned readings

This distribution only approximate and is subject to modification. Any changes will be announced in class and posted on the course website.

Homework: Homework will be assigned approximately every two weeks. Late homework will not be accepted, except under serious unavoidable circumstances (significant illness, injury, death in the family). Collaboration is allowed when working on homework problems, but all resources (people, papers, websites, etc.) must be appropriately credited. Also, each student must write-up and submit an independent solution set in his/her own words.

Attendance: Regular attendance is expected. Exceptions will be accommodated only for valid, documented reasons including (1) official representation of the Institute and (2) medical emergencies.

Note: If you will not be able to meet the requirements of the class as stated, you must contact me within the first two weeks of class.

Academic Integrity: Students are reminded of the obligations and expectations associated with the Georgia Tech *Academic Honor Code* and *Student Code of Conduct*, available online at: `http://www.deanofstudents.gatech.edu/integrity/policies/honor_code.php` and `http://www.deanofstudents.gatech.edu/codeofconduct`.
Any violations must be reported to directly to the Dean of Students.

Additional Resources: See the Math 8803/4803: Discrete Mathematical Biology webpage at `http://www.math.gatech.edu/~heitsch/discmathbio.html`

Updates: This syllabus is subject to modification. Any changes will be announced in class and posted on the course website.