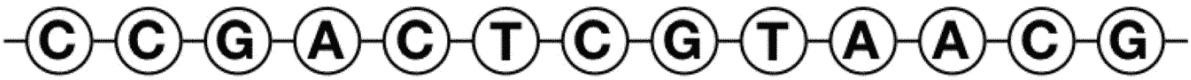
Field Theorists Fold RNA

An international collaboration with Henri Orland, Matt Pillsbury, Anthony Taylor, Graziano Vernizzi (started with a sabbatical year spent by Orland at KITP)

We exploit topological concepts embedded in large N matrix field theory to attack an important problem in biology—the folding of RNA

1

Our parents work hard to give us DNA, a message written in an alphabet with 4 letters: C, G, A, T



Think of beads on a string

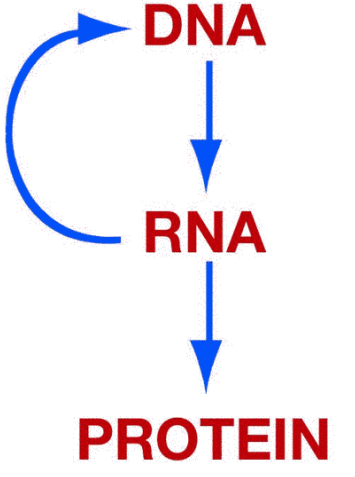
DNA consists of two complementary strands twisted into a double helix (see picture later)

Attraction between C and G & between A and T

2

"Central Dogma of Biology"

Watson & Crick



Information storage
~ hard disk

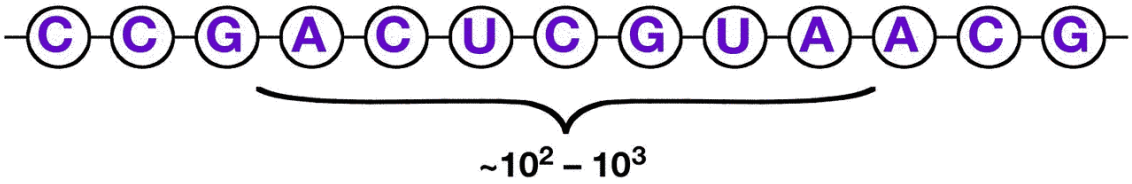
~ floppies to give to your friends

Function (enzyme, motor, muscle,…)

"You don't want to mess with your hard disk, but you can have as many copies of RNA as you want"

3

RNA

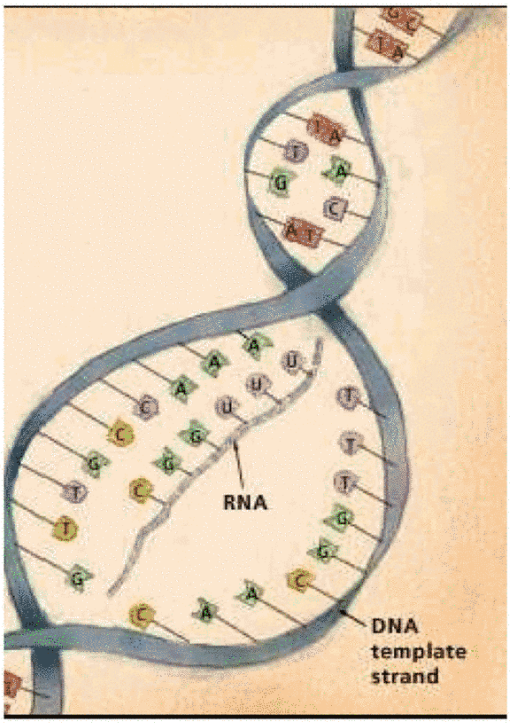


Much much much shorter than DNA

- A message written in a 4-letter alphabet, C, G, A, U
- Attraction between C and G & between A and U
- Hydrogen bond: saturates =) once a C paired with a G, it does not pair with another G
- Think of beads on a long chain connected by rigid rods. Saturation of hydrogen bond means we can only glue two beads together

4

In each one of our cells, stretches of DNA open up & information is copied onto RNA



5

RNA is single stranded, unlike DNA

C C G A C U C G U A A C G

Because of attraction between C and G & between A and U, once RNA floats free, it folds into a definite shape

Different message ⟶ different shape

RNA folding problem

Given the message, what is its shape?

Why is this important?

6

Watson & Crick thought that RNA is "merely" a "passive carrier of information" from DNA to protein ("messenger RNA")

Revenge of RNA

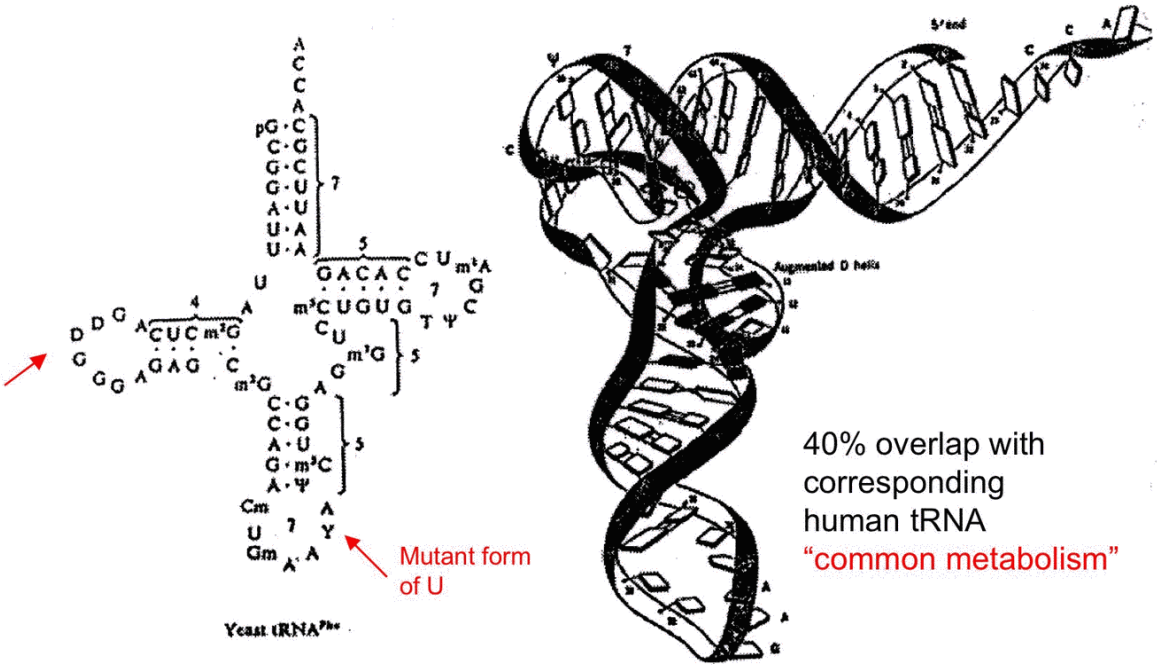
Important discovery of last 15 years or so: RNA plays a crucial enzymatic role

Biochemistry incredibly complicated, but to first approximation: shape of molecule ("lock and key")

By the way, biologists working on the origin of life now generally believe that it started with an "RNA world." DNA came later.

7

Yeast

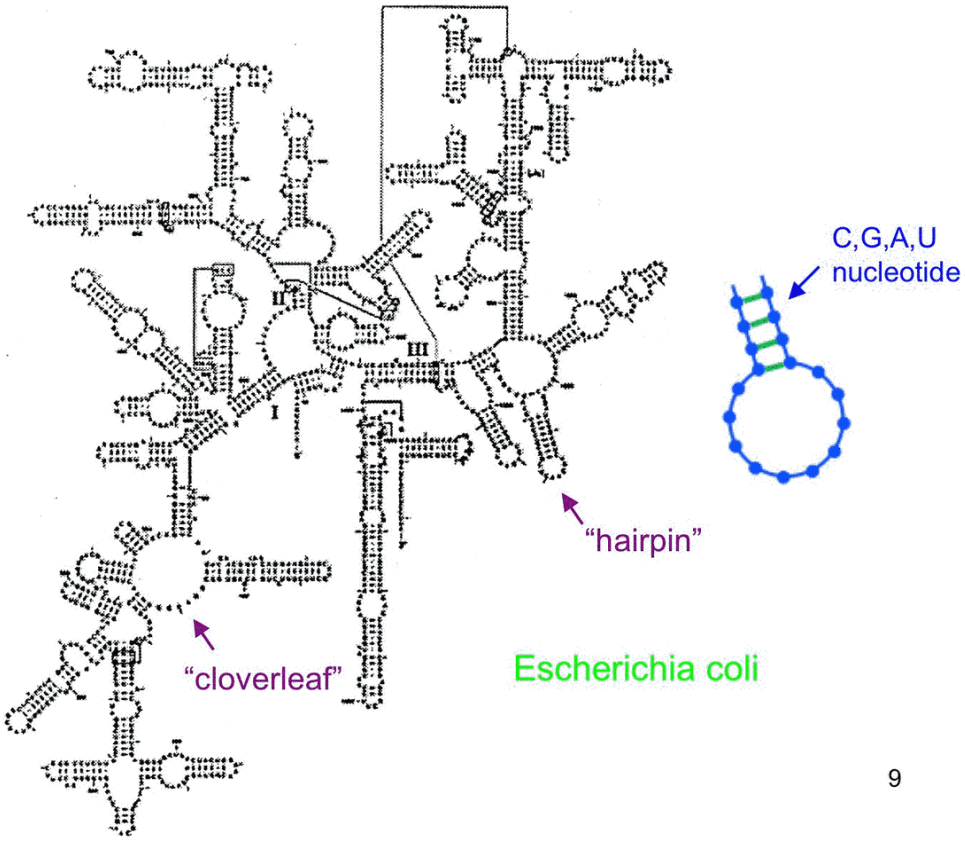


40% overlap with corresponding human tRNA "common metabolism"

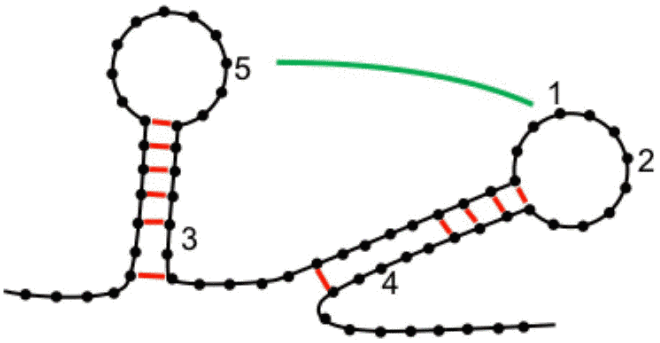
sequence space ⟶ shape space

8

Secondary Structure: small subunit ribosomal RNA



9



— glue

- planar as shown: two "hairpins"
- but if glue 4 to 3, no longer planar
- Rigidity constraint: chain not infinitely flexible, gluing 1 and 2 not allowed.
- If glue 5 and 1, gives "kissing hairpins"

10

$Z = 1$

$+ \Sigma V_{ij}$

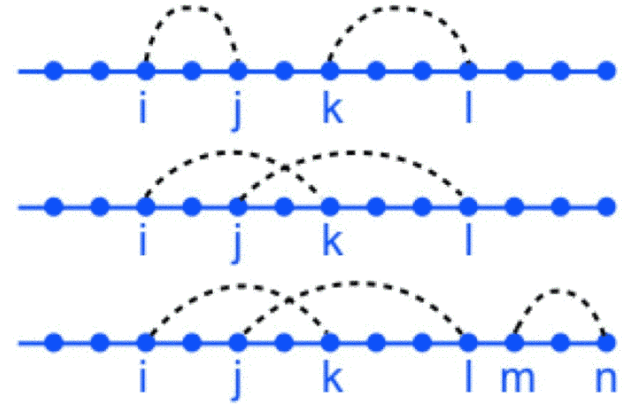
$+ \Sigma V_{ij} V_{kl}$

$+ \Sigma V_{ik} V_{jl}$

$+ \Sigma V_{ik} V_{jl} V_{mn}$

$+ \ldots$

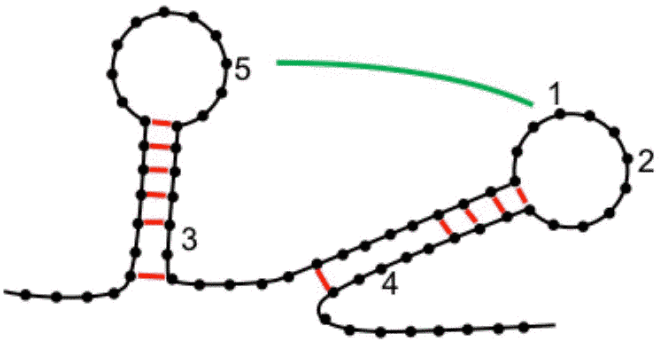
$+ \Sigma VVV \ldots V$



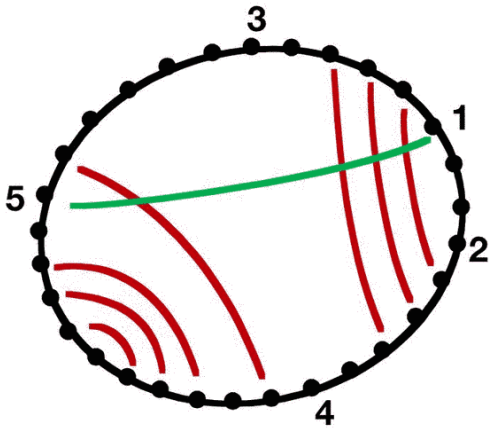
Just list all ~L! possibility

These look like Feynman diagrams: a quark propagating along, emitting and absorbing gluons

11

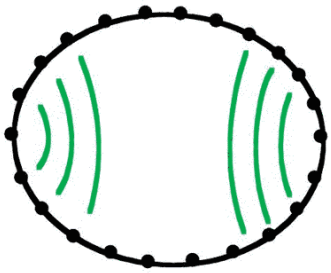


Mathematically, better to glue the two ends to form a circle



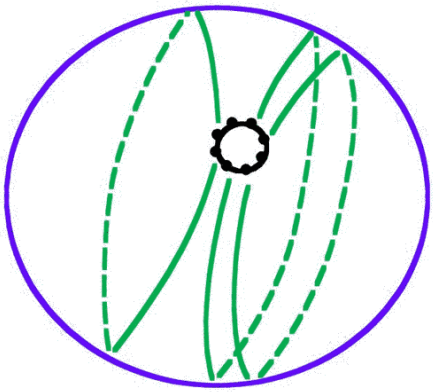
12

Mathematicians do not like to draw on a flat piece of paper; prefer manifolds without boundary, "compactify" plane into sphere so no boundary at infinity.



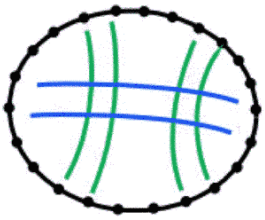
Put a hole in the sphere to represent RNA chain

"punctured sphere"



Sphere: genus g = o

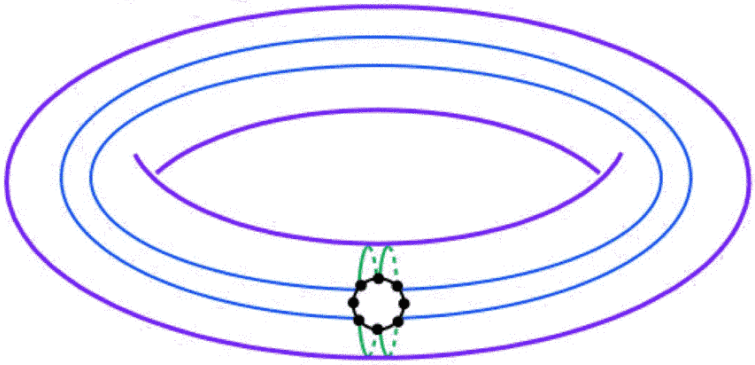
13



Physicists: lines cross

Mathematicians: on what kind of surface (what genus?) can we draw this without the lines crossing?

"Punctured Torus"
Torus (or doughnut, bagel) with a hole put into it



Punctured Torus
Torus: genus g=1

Different meaning in biology and mathematics!

14

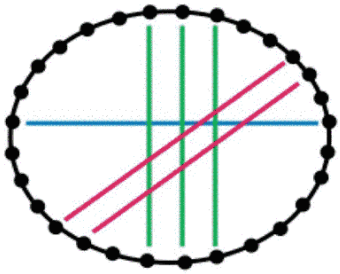
Genus

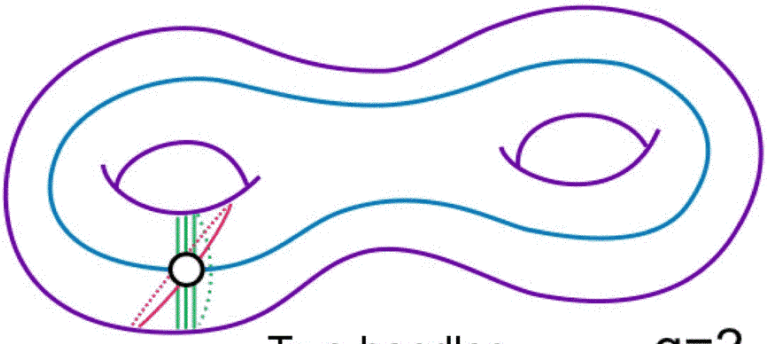
Mathematics (Topology):
How many handles?
Better word than holes

Biology (Taxonomy):
Group of species, e.g. Homo, Pan

A joke about a conference in mathematical biology

15





Two handles g=2

16

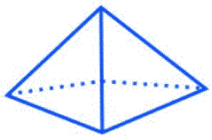
Biologists called planar diagrams, i.e. diagrams that could be drawn on a sphere without crossing secondary structures.

They call diagrams that could be drawn on a torus without crossing tertiary structures

Experimentally, the relative importance of tertiary structures to secondary structures can be dialed by varying Mg^{++} ion concentration (has to do with screening due to Mg^{++})

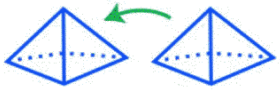
17

Euler's Theorem



F	V	E
4	4	6

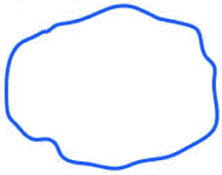
glue two together



ΔF	ΔV	ΔE
3-1 = 2	1	3

=

Topology of sphere



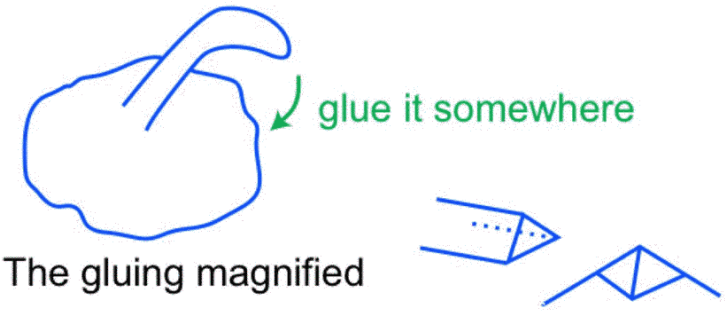
$\Delta F + \Delta V - \Delta E = 0$

$F + V - E = 2$

$= 4 + 4 - 6$

18

Changing the genus



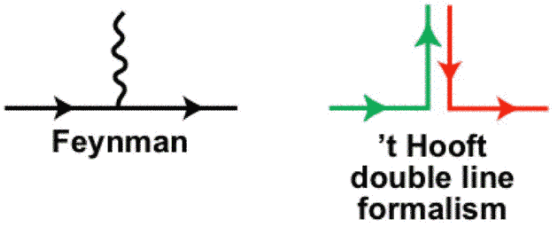
$\Delta F = -2$
$\Delta V = -3 \Rightarrow \Delta (F + V - E) = -2$
$\Delta E = -3$

for doughnut or torus
Euler characteristic number
F + V – E = O

Repeat
F + V – E = 2 (I – H) H=number of handles

19

Field theorists have a method ('t Hooft) inspired by large N quantum chromodynamics, quarks & gluons, with quarks in N different colors (real world: N=3)



quark goes into quark by emitting a gluon

N^2 different kinds of gluons described by double lines

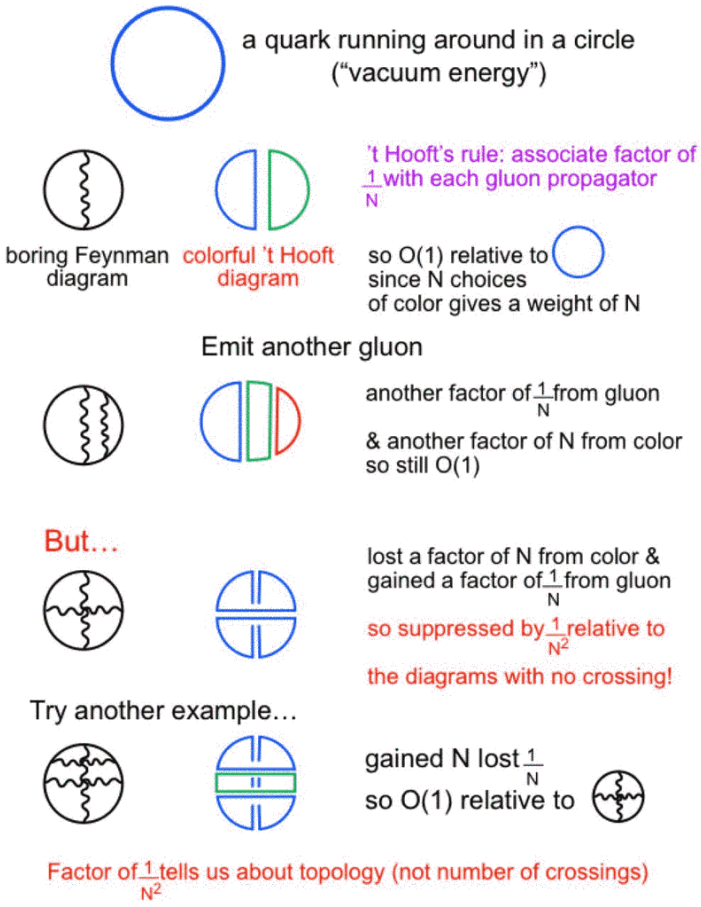
gluon field naturally a matrix field

A matrix is a mathematical object with two indices on it (just as a vector is a mathematical object with one index on it).

Remember the word "matrix"!

A^i_j ← color on the two lines
i, j = 1, 2, …N

20



Now theoretical physicists could exploit our knowledge of random matrix theory and large N quantum chromodynamics to formulate the RNA folding problem and have fun!

Random matrix theory was invented by Wigner (Nobel Laureate) in the mid-1950s to study nuclear physics, but its growth and ramifications over the last 50 years have been beyond the wildest dreams of Wigner and his friends, with implications and uses in diverse fields, e.g. disordered materials in condensed matter physics, pure mathematics, operator algebra, M-theory (supposed to contain string theory) & more recently econophysics