



The National Center for
Physics-Based Simulation
of Biological Structures
at Stanford

BioE 215

Constraints and Molecule Modeling

1 June 2007

Michael Sherman (Sherm)
Chief Software Architect, Simbios



NIH Roadmap
grant U54 GM072970

Topics

- Modeling matters
- Modeling matter
- Modeling molecules
- Modeling constraints in Simbody

2



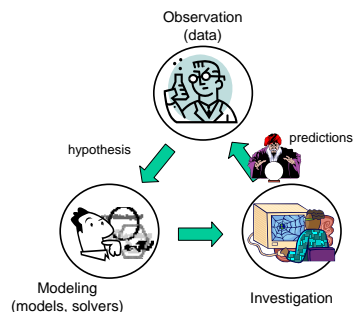
Next topic

- **Modeling matters**
 - Definition, purpose, simplification, validity
 - Modeling in Simbody
- Modeling matter
- Modeling molecules
- Modeling constraints in Simbody

3



Modeling is a process

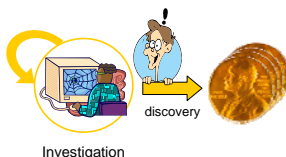


4



The purpose of modeling?

To create a *computational* system
more rewarding to investigate
than the original *physical* system.



5



Modeling is *valid simplification*

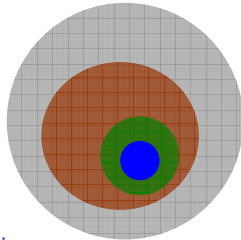
- Simplification is *desirable* for understanding
 - what's important?
- Simplification is *required*
 - for computational feasibility (i.e., speed)
 - to reduce modeling effort
 - because we don't know enough to build a better model
- **What is a valid model?**

6



Four levels of model validity

1. Face validity
 - "looks good"
 - visualization, animation, games
2. Replicative validity
 - "interpolates"
3. Predictive validity
 - "extrapolates"
4. Structural validity
 - componentwise prediction
 - "instructs"



7

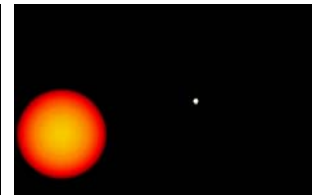
Structural validity example 1

- Stars follow predictable mass/color relationship
- Globular cluster observation: many stars break the rules
- Could collisions explain that?



Blue star is too massive.

8



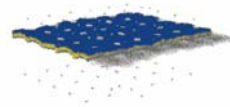
Neutron star shouldn't be red.



Structural validity example 2

(From last week's Nature!)

- Cell membranes naturally take on minimum curvature (spheres)
- How do they attain the convoluted states we see?
- Do membrane forces drive clumping of associated proteins and cause curvature?



36 "proteins" in 46,080 lipid molecules for 1ms
Reynwar, et al., Nature 447:461-4 (2007)



9

Model simplification summary

- For science & engineering use, build models which are
 - **simplified**, but
 - **structurally valid**
- Then use the models to learn about and manipulate the physical world



10

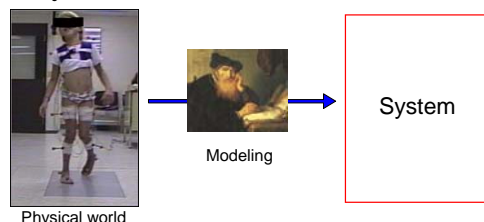
Next topic

- **Modeling matters**
 - Definition, purpose, simplification, validity
 - **Modeling in Simbody**
- Modeling matter
- Modeling molecules
- Modeling constraints in Simbody



11

Simbody terminology: modeling creates a "System"

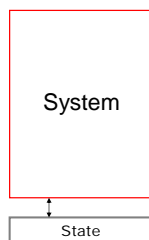


- A "System" is a computational embodiment of a mathematical model



12

Properties of a System



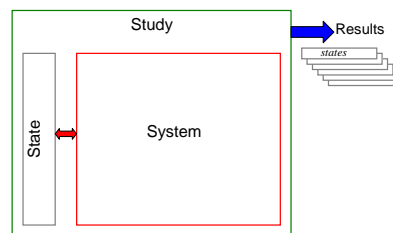
- Defines its parametrization
- But ... it is *stateless* (const).
- The fixed content of a system is called its *topology*
- Given a *State*, performs useful computations

13



Studying a system

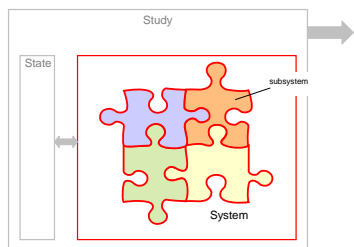
- System+State+Study → Simulation



14



Systems are composed from subsystems

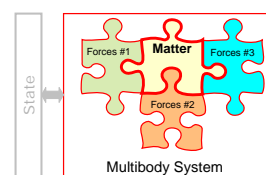


- Interlocking computations
- System provides the “edge pieces”

15



What’s in a multibody system?

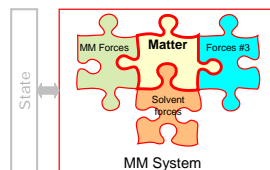


- Matter and forces
- Also:
 - Geometry (analytic & decorative)
 - Mass property calculation
 - Other properties, e.g. atom types

16



A molecular mechanics (MM) system is a kind of multibody system

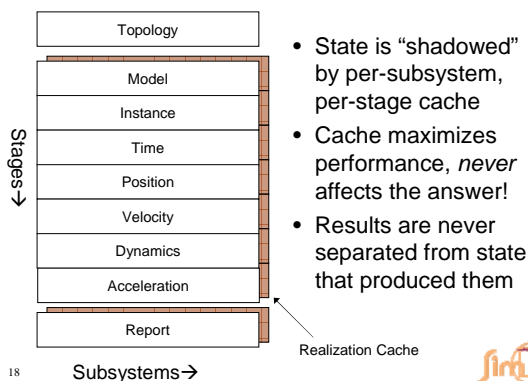


- Has matter and a molecular mechanics force subsystem
- Helpful to have a “modeler” for molecules of interest, to coordinate the matter & forces

17



States are “realized” in stages



18



Next topic

- Modeling matters
- **Modeling matter**
- Modeling molecules
- Modeling constraints in Simbody

19

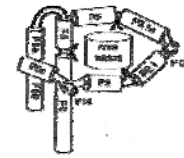


Similar models across multiple scales



Model of human

Huang, et al. Proc. 2nd Pacific Conf. Fundamentals Comp. Graphics, 1995



Model of RNA
(Tetrahymena group I intron)

Zheng, et al. PNAS 98(7), 2001

20



Matter and its properties

- Mass
- Spatial distribution
- Motion

21



Abstract matter

- The rigid body
- What is a rigid body?



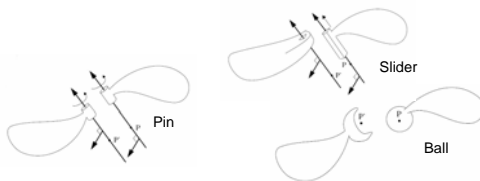
- Mass distribution: 10 *constants*
- Decorate w/geom & other props
- Ground is a (heavy) rigid body

22



Joints

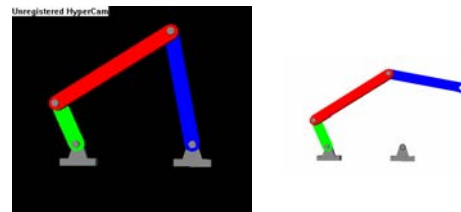
- Defines relative motion between bodies
- Examples



23



Joints can *permit* or *restrict* motion



- 4-bar linkage has only 1 DOF
- But has 3 if you remove any one joint

24



Mobilizers

(joints which *permit* motion)

- In Simbody, bodies do not have inherent mobility
- Mobilizers precisely define the allowable motion relative to parent
- A mobilizer *always* increases the system's mobility
- These define the generalized coordinates q and generalized speeds u
- Bodies + mobilizers form a tree



$$M\ddot{q} = f$$



25

Constraints

(joints which *restrict* motion)

- Trees can be a little too floppy ...
- *Constraints* introduce constraint equations (1 or more)
 - Restricts allowable motion – like negative mobility
 - E.g., ball constraint adds 3 constraint equations, -3 dofs
- Algebraic invariant relating q 's and u 's: $g(q, u) = 0$
- But ... might not be independent
- Must solve assembly problem before simulating
 - Find q such that $g(q) = 0$
- Constraints permit loops, generate additional unknown forces



$$M\ddot{q} = f - f_c$$

$$g(q) = 0$$



26

That's it for rigid matter

- Joints can be mobilizers, constraints or both
- Rigid bodies + mobility – constraints
- Write in “descriptor” form:

$$M\ddot{q} = f - f_c$$

$$g(q) = 0$$

- Solve assembly problem $g(q) = 0$
- Solve for \ddot{q} and f_c
- Integrate for trajectory
- Still need forces ...



27

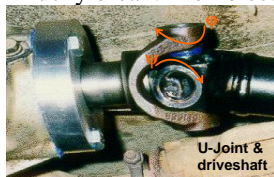
Next topic

- Modeling matters
- Modeling matter
- **Modeling molecules**
 - **Molecular mechanics**
 - Some sources of confusion
 - Modeling proteins and RNA
- Modeling constraints in Simbody

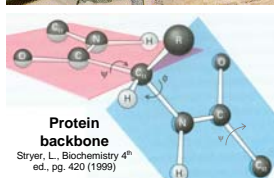


28

Lucky break: Biomolecules are machines too!

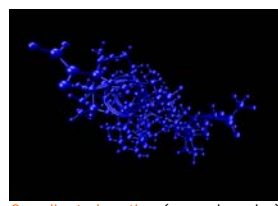


U-Joint & driveshaft



Protein backbone
Stryer, L., Biochemistry 4th ed., pg. 420 (1999)

Substructures (torsion space)



Coordinated motion (normal modes)
Courtesy Locus Pharmaceuticals



29

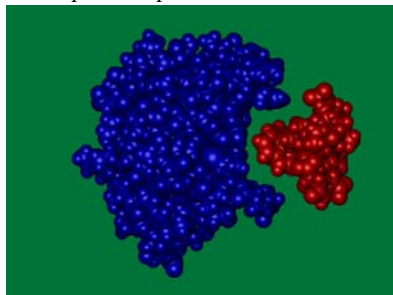
Molecular Mechanics (MM)

- Simulate a system composed of
 - Molecules of interest
 - Their environment
- Why? Understanding & insight; actionable info
 - Predict stable structure (e.g. protein folding)
 - Predict dynamic behavior (e.g. mode of operation)
 - Predict thermodynamic properties (e.g. binding affinity)
- Almost always interested in bulk (average) behavior
 - And often at equilibrium
- Domain is classical physics only
 - Classical mechanics ($F=ma$)
 - Statistical mechanics (entropy)
 - Electrostatics, continuum models



30

Protein/protein interaction MDM2/p53 Complex Formation



- Two very flexible molecules
- Substantial induced fit

Imagiro® Molecular Dynamics
Courtesy Locust Pharmaceuticals

31



Justification for MM

- “Ab initio” quantum calculations are accurate but too slow
- Quantum effects become classical in the aggregate
 - This is rigorously true
 - Statistical effects become deterministic
- Many biomolecular systems of interest are “big enough”
 - True for structural rearrangements, not for “chemistry”
- So MM postulates classical forms for system behavior
 - Springs, electrostatics, $F=ma$, etc.
 - Fit constants using averaged quantum mechanics and/or empirical measurements

32



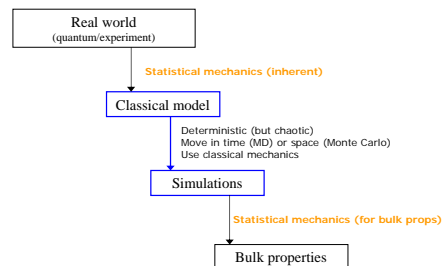
Next topic

- Modeling matters
- Modeling matter
- **Modeling molecules**
 - Molecular mechanics
 - **Some sources of confusion**
 - Modeling proteins and RNA
- Modeling constraints in Simbody

33



Statistical vs. deterministic mechanics

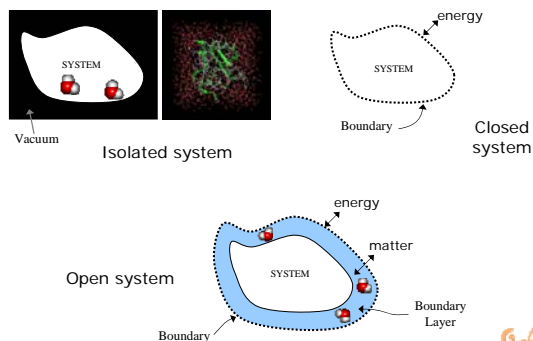


- Statistical behavior *emerges* from classical mechanics
- This is a source of great confusion

34



3 kinds of thermodynamic systems

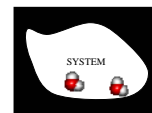


35



Isolated system (all atoms modeled)

- Energy should be conserved
- Forces associated with bonds
 - Stretch, bend, torsion
- Forces between every pair of atoms
 - Electrostatic (charge interactions)
 - van der Waals attraction (atoms are sticky)
 - “Steric” (atoms can’t overlap)
- That’s it
 - NO friction
 - NO damping
 - NO “entropic” forces

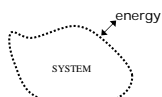


36



Closed system

- Environment not modeled atomistically
 - Energy is NOT conserved(!)
- Must model interaction between molecules & environment
 - But no exchange of matter
- Detailed interaction at the boundary
 - Van der Waals & electrostatics: unmodeled atoms
 - “Entropic” effects: unmodeled dofs
- Bulk environment is an infinite reservoir
 - Constant temperature: unmodeled kinetic energy (KE)
 - Constant pressure: unmodeled steric forces
 - Dielectric: unmodeled electrostatics

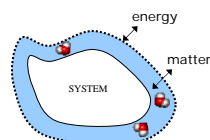


37



Open system

- Environment not modeled atomistically
- But treated as a reservoir of boundary molecules
 - And temperature, pressure, dielectric as before
- Neither mass nor energy conserved
- Most difficult to implement; most promising



38



Next topic

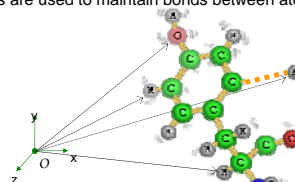
- Modeling matters
- Modeling matter
- **Modeling molecules**
 - Molecular mechanics
 - Some sources of confusion
 - **Modeling proteins and RNA**
- Modeling constraints in Simbody

39



Conventional mass & motion

- Every atom is a point mass
 - Located at nucleus
 - Electrons are not modeled separately
- Every atom has x,y,z Cartesian coordinates, velocities
- Forces are used to maintain bonds between atoms



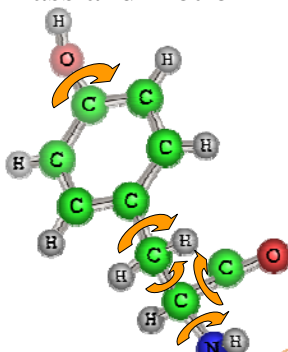
- This is not necessarily the best model! Just the easiest.

40



Generalized mass and motion

- Motion only where desired
- Performance is $O(n)$ in *mobility*, not *atom count*
- Simbody's generalized model can do **both** conventional and internal coordinate models



41



A Simbody molecular system



- 3 Ethanes
 - Rigid
 - 1 internal dof
 - Fully flexible
- 1 Oxygen (O2)

Simbody + MM subsystem: choose mobility(orange) as appropriate

42



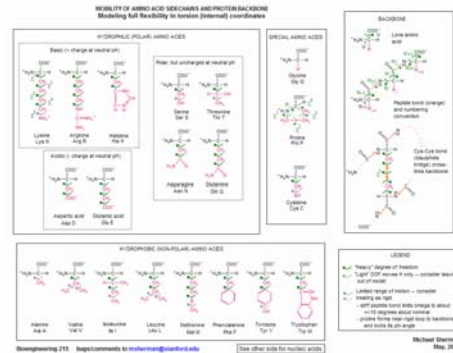
Proteins and nucleic acids

- Long chains
- All significant flexibility about torsions (single bonds)
- Bond stretch/bend generally ignorable
- Handout shows how to model full flexibility in torsion coordinates

43



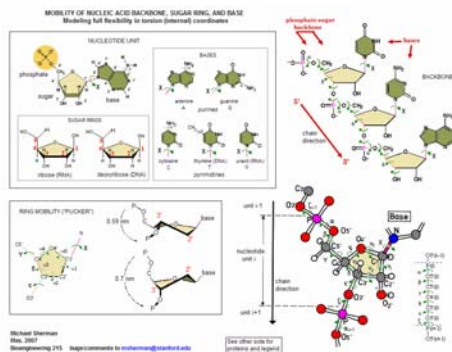
Handout: proteins



44



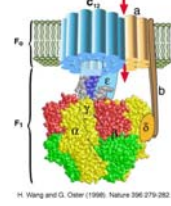
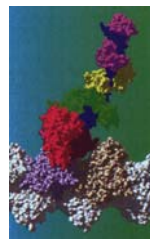
Handout: nucleic acids



45



But big molecules have big parts ...

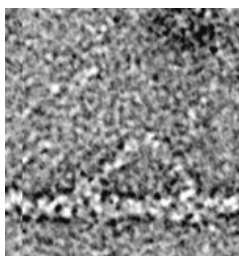


... can we model them accordingly?

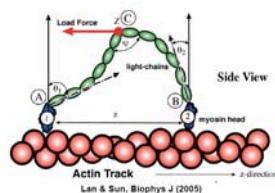
46



Mechanical models of molecules



Assembled from EM data
Walker, et al., Nature 405 (2000)



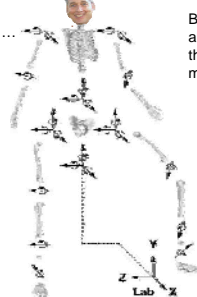
47



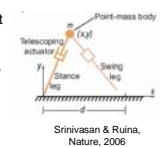
So how *should* you model a molecule?

- Start with an "easier" question:
How should you model a human skeleton?

Jeff says ...



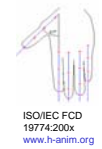
But what about these models?



Srinivasan & Ruina, Nature, 2006



Stammen, NHTSA 2001



ISO/IEC FCD 19774:200x
www.h-anim.org

48

Fregly, Reinbolt, et al. 2007



So how *should* you model a human skeleton?

- It depends on what you're trying to accomplish.
- Next question:
How should you model a molecule?

49



Next topic

- Modeling matters
- Modeling matter
- Modeling molecules
- **Modeling constraints in Simbody**

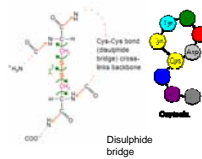
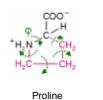
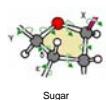
50



Molecule models that require constraints

- Fully-flexible polymers are *almost* trees

- Disulphide bridge & proline in proteins
- Sugar ring in nucleic acids



- Can model with forces or constraints

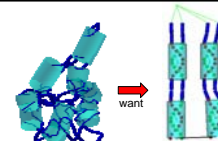
- Coarse-grained models usually heavily constrained

51



Example: coarse-grained RNA

First, build tree system (no constraints)



- 20 bodies, 20 ball mobilizers
- 60 unrestricted dofs
- Very floppy!

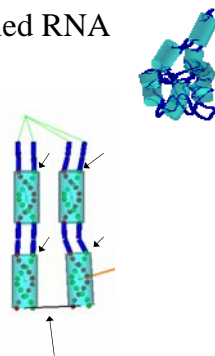
52



Example: coarse-grained RNA

Then add constraints

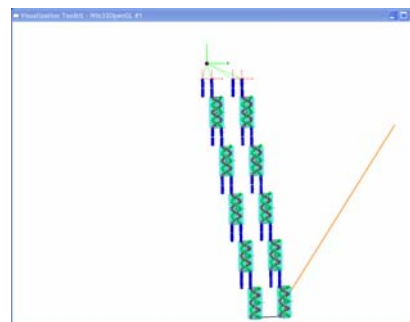
- 20 bodies
- Each is mobilized by a ball mobilizer, with 3 dofs (60 dofs total)
- Loops are closed with 4 ball constraints, -12 dofs
- One distance constraint, -1 dofs
- All constraints are independent
- So, 47 net dofs



53



Simbody simulation

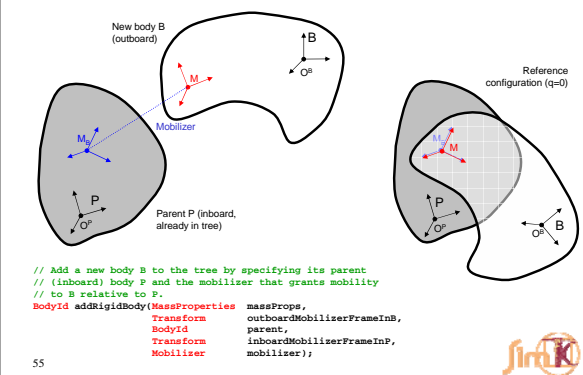


- RNA with rigid duplexes
- 50 bodies, 150 internal dofs
- 31 constraints
- Gravity & a spring??
- If you can imagine it, you should be able to try it
- Runs in a few minutes

54



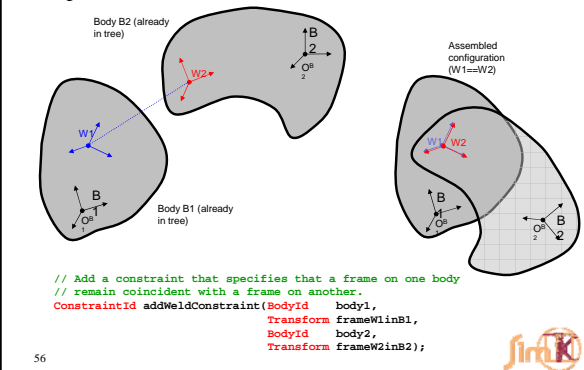
Mobilizers in Simbody (review)



55

Constraints in Simbody

Example: weld constraint



56

Constraints in Simbody

- Restrict the motion of one or more bodies
- Currently, three kinds available:
 1. Constant distance constraint (massless rod)
 - Relates **point** of body A to a **point** of body B: distance $|B-A|$ must have a constant value $d (>0)$ at all times
 - 1 constraint equation
 2. Coincident points constraint (ball joint)
 - Relates **point** of body A to a **point** of body B: distance $|B-A|$ must be zero at all times
 - 3 constraint equations
 3. Weld constraint
 - Relates **frame** on body A to a **frame** on body B: frames must be coincident at all times
 - 6 constraint equations
 - **BUG:** Current (bad) implementation requires a very good initial guess for the q 's
- **BUG:** resulting constraint equations must be *independent* (non-redundant)
 - Both among the constraints, and with respect to mobilizers
 - **That is, every constraint equation must reduce the mobility**

57

All currently available constraints

```
// From SimbodyMatterSubsystem.h
// Add a Constraint that specifies that a point on one body
// remains a fixed distance (>0) from a point on another (massless rod).
ConstraintId addConstantDistanceConstraint(BodyId body1,
                                          Vec3 pointOnBody1,
                                          BodyId body2,
                                          Vec3 pointOnBody2,
                                          Real distance);

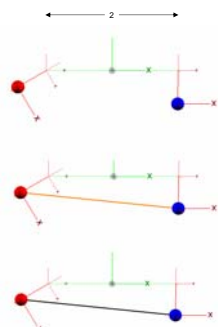
// Add a Constraint that specifies that a point on one body
// remains coincident with a point on another (ball joint).
ConstraintId addCoincidentPointsConstraint(BodyId body1,
                                          Vec3 pointOnBody1,
                                          BodyId body2,
                                          Vec3 pointOnBody2);

// Add a constraint that specifies that a frame on one body
// remain coincident with a frame on another (weld joint).
ConstraintId addWeldConstraint(BodyId body1,
                              Transform frameW1,
                              BodyId body2,
                              Transform frameW2);
```

58

Two pendulum example

See handout for Simbody code



- No connection
- Left pendulum has initial condition of -60 degrees
- Connected by a spring of natural length 2
- Spring is initially stretched
- Connected by a distance constraint $d=2$
- Constraint is initially violated; must be assembled

59

Discussion

msherman@stanford.edu

<https://simtk.org>

60