Investigating how graduate students connect microstates and macrostates with entropy

Nathan Crossette, Michael Vignal, and Bethany Wilcox
Department of Physics, University of Colorado - Boulder, Boulder, CO 80309

As a first step in a larger study of student reasoning in upper-division thermal physics, we conducted think-aloud interviews with 8 physics graduate students to probe their understanding of entropy. In this paper, we'll discuss results from a question which presented students with a novel system—a string in a bath of water—and asked students to rank the probabilities of particular arrangements of the string, define macrostates of the system, and discuss specifically what is meant by the entropy of the system. Exploring graduate students' understanding of entropy and their ability to solve problems and reason with entropic arguments will provide insights into how physicists develop a mature understanding of entropy as a physical quantity. We find a tendency for graduate students to project properties of macrostates onto constituent microstates, and discuss other observations. We identify connections to previous research and lay out the next steps for this project.

I. INTRODUCTION & BACKGROUND

Thermal physics concepts have a wide range of relevancy across the fields of biology, chemistry, physics, and engineering. Two core concepts are energy and entropy. Inspiration for this study came from the seemingly paradoxical fact that while lacking a concrete definition, the concept of energy causes little discomfort among students, yet the concept of entropy proves a subtle and difficult to understand concept despite having a more formal definition within the context of thermal physics.

In 2015, Dreyfus [1] thoroughly catalogued and summarized prior research on student learning in thermodynamics and statistical mechanics across the disciplines of physics, chemistry, and biology. Current PER literature on upperdivision thermal physics is limited, and work specifically related to entropy has mostly explored its thermodynamic and macroscopic contexts such as ideal gases [2], heat engines [3], thermal equilibrium [4], and the 2nd law [5, 6]. Other research from across the natural sciences also has addressed students' understanding of entropy [7–10].

This study directly addresses graduate students' understanding of entropy from a statistical mechanics perspective, specifically entropy as it relates to probability, which Dreyfus reports has received little coverage [1]. Though less research has focused on the statistical mechanics view of entropy making the results of prior research not directly applicable to this study, we do point out some interesting parallels between ours and prior work in Sec. III.

Additionally, graduate physics students remain an understudied population. Learning more about their conceptual difficulties will indicate truly persistent student struggles and provide insight into how graduate students reason and construct models. In addition to potentially improving graduate student learning, future work in this vein will provide more perspective on difficulties experienced by undergraduates and a better understanding of the transition from novice to expert physicist.

II. METHODOLOGY

Canonical examples used to explore entropy as the logarithm of the number of microstates include coin flips, small Einstein solids, and the two-state paramagnet. These systems demonstrate that even though entropy is a macroscopic quantity, a complete description of the quantity involves multiplicities of microstates. Furthermore, these examples have only one obvious way to classify macrostates.

The data presented in this paper were collected from a longer interview on entropy. Here, we will discuss a subset of the questions relating to a string jiggling in a bath of water. This question was crafted to probe students' understanding of the differences and connections between macrostates, microstates, and entropy in a novel context without simply recalling results from a familiar system.

A) Based on your intuition, rank the probabilities of finding the string in each of the three conformations shown above.
B) Is there a property of the string that can be used to define a set of distinct macrostates of the string? Are the conformations shown in the figure macrostates or microstates?
C) Based on your answer to part B, how would you rank the probabilities of finding the strand in each of the three conformations above?
D) Can you discuss what is meant by the "entropy" of the string, and how it relates to the possible conformations of the string?
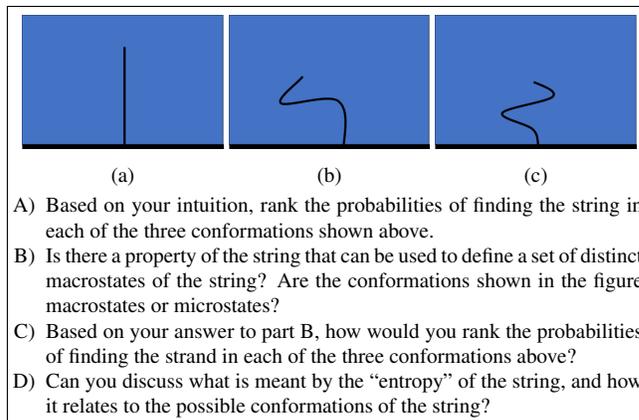
FIG. 1: Interview prompt provided to students. Space was provided for students to show their work.

We presented interviewees with 3 cartoons of a string sitting in a bath of water (Fig. 1). A preamble told students that the string and water had equal densities (*i.e.* gravity could be ignored) and that Brownian motion of the water could change the "'conformation' (meaning 'position' or 'arrangement') of the string." Additionally, interviewees were told to consider the water as a continuous medium, implying that the entropy of the bath of water should be ignored.

The images represent microstates of this system, so by the fundamental assumption of statistical mechanics, they are all equally probable conformations of the string. However, in part A, we expected a strong intuitive temptation to rank the string in Fig. 1a as the least likely conformation due to the fact that it is the only microstate in the macrostate of a completely straight string: the least likely macrostate.

In part B, we probe students' preferences for defining macrostates as well as the sophistication of their ability to do so. Many acceptable definitions could be applied to this system, and we sought to probe the range of studentgenerated ideas. By asking about whether the figures depicted macrostates or microstates, we hoped to learn to what extent students could recognize microstates in a novel situation and demonstrate an understanding of the difference between them and macrostates. By following up immediately with the prompt to re-rank the probabilities in part C, we could see whether students who gave an unequal ranking in the first part change their mind after explicitly prompted to think about microstates and macrostates.

Finally, part D intended to see whether students could consolidate their reasoning about microstates and macrostates, and articulate that the multiplicity of a macrostate—*i.e.* the number of microstates, W, within that macrostate—determines its entropy via Boltzmann's law: $S = k_B \ln W$.

Before interviews began, the questions were vetted by a physics faculty member familiar with block co-polymer models—the physical system on which this question was inspired—as well as thermodynamics instructors from Chemistry, Engineering, Biology, and Chemical Engineering. The questions were piloted in two interviews: one with a content

expert, and one with a PER graduate student. Interviewees were paid volunteers who responded to an email request to the physics graduate student population at the University of Colorado - Boulder. All students are referred to by pseudonyms.

The interviews took place in the lead up to and first several weeks of the Spring 2020 semester. A total of 8 graduate students participated. Most participants were in the second semester of their first year ($N = 5$). Of the rest, one was in the second semester of their second year, one was a transfer student with a Master's degree, and one was in their sixth year. Two were international students, and a total of six had previously taken a graduate course in Statistical Mechanics or were taking the course in the Spring of 2020. Seven had a physics (or engineering physics) undergraduate background and one had a background in computer science.

Interviews were conducted in a think-aloud setting [11] where participants were asked to verbalize their thoughts and reasoning as much as possible. The interview questions were printed on LiveScribe paper. In addition to synced written work and audio from a LiveScribe pen, interviews were audio and video recorded. An interviewer was present to answer questions, prompt the students to verbalize their thinking, and ask students to further explain their reasoning.

Written drafts of the interviews were obtained from Otter.ai transcription software. The drafts were then manually checked and edited against the recorded video. Interview transcripts were emergently coded based on student responses to question prompts. The coding methodology was inspired by Hammer's resources framework [12] for identifying the conceptual reasoning elements students employed in thinking through this problem. Responses were summarized and analyzed to identify the range of responses and the commonalities and differences between responses. After initial coding by the first author, all three authors collaboratively reviewed sections of interviews to verify code assignments and reach agreement on difficult-to-code passages.

III. RESULTS & DISCUSSION

We present student responses by topic. First, students' responses related to microstates and probability (which corresponds to parts A, B, and C in Fig. 1) are discussed, followed by reasoning about macrostates (part B), then students' reasoning about the entropy of the string and whether they connected the idea to the concept of multiplicity (part D).

A. Identification of Microstates and Probabilities

When first asked to rank the probabilities of the images in Fig. 1, 3 of the 8 students correctly responded that the probabilities were equal using either the fundamental assumption of statistical mechanics or an appeal to symmetry. A total of four students gave the ranking of $P(c) > P(b) > P(a)$, which from here on will be referred to as the 'intuitive ranking'. For a summary of student responses, see Tab. I.

Student reasoning behind the 'intuitive ranking' favored qualitative, vague metrics like how 'wavy,' 'wiggly,' or 'centered' (*i.e.* how much the string deviated from center) the string looked. In more quantitative forms, these metrics more appropriately describe macrostates (see Sec. III B). Usage of these metrics to rank microstates may suggest that students initially assumed the images represented macrostates. However, we see evidence that students consciously or unconsciously recognize the relative probabilities of overarching macrostates, then project these probabilities onto string conformations: microstates belonging to those macrostates.

After calling the images *microstates*, Chris employed this projection when considering whether to re-rank the states in part C:

Chris: *I guess going along the lines of the macrostates, you gotta think of which like, length of string is more likely. I would imagine the full length of the string is the least likely.*

This reasoning gets the causal link for macrostate probabilities backwards, and overlooks the equal probability between microstates. The graduate student in their sixth year, Harry, discussed the allure of this 'intuitive ranking':

Harry: *Um, I feel like if this was actually set up in the water, and I like looked into it and saw that the string was exactly straight I would think that was really weird. But then B and C? You know, I wouldn't think that was weird. I feel like that's just how it randomly went around.*

But I think the answer to this question is probably that each of the three microstates shown here is equally probable. And if I were to like bunch those into macrostates where I say like there's a string perfectly straight or string bunched up somehow, like, I would expect it to be more likely that it was bunched up somehow. But like having it bunched up in exactly the way shown is a very small section of that particular macrostate.

Of the four students initially giving the 'intuitive ranking', three (Beth, Chris, and Garth) later changed their ranking to equally probable after considering the question of whether the states in the figures were microstates or macrostates. Beth made this correction spontaneously after articulating that the images were microstates so their probabilities should be equal. Chris and Garth, however, recognized the images were microstates on their own, but struggled to settle on the equal ranking until the interviewer asked if they knew anything about the relative probabilities of microstates.

After fully considering the third prompt, a total of five students had decided the figures were microstates. Alex, Erik, and Fred stated the images represented *macrostates*. In terms of probability rankings, Erik vacillated between the 'intuitive ranking' and ranking the states equally. Interestingly, Alex and Fred both settled on the ranking $P(a) > P(c) > P(b)$ based on reasoning that it was most probable for the string to have less net deviation from the center. Their reasoning, illustrated in the following quotes, suggested a conflation between a macrostate and a macroscopic object.

TABLE I: Summary of graduate student responses. Letters in parentheses on title row refer to prompts in FIG. 1. Bold entries in the final column indicate rankings that changed after students considered macrostates and microstates.

Student	Initial Ranking (A)	Consideration in Initial Ranking (A)	Ideas for Macrostates (B)	Final Ranking (C)
Alex	$P(c) > P(b) > P(a)$	deviation from center	net deviation, conformations	$\boldsymbol{P(a) > P(c) > P(b)}$
Daana	$P(c) = P(b) = P(a)$	fundamental assumption of statistical mechanics	# of turns, net deviation, 'range' between points	$P(c) = P(b) = P(a)$
Beth	$P(c) > P(b) > P(a)$	'wavyness'	# of turns	$\boldsymbol{P(c) = P(b) = P(a)}$
Erik	$P(c) = P(b) = P(a)$	there is no preference for a particular conformation	conformations	$P(c) = P(b) = P(a)$
Chris	$P(c) > P(b) > P(a)$	'wiggliness', amount of deviation	distance between ends	$\boldsymbol{P(c) = P(b) = P(a)}$
Fred	$P(a) > P(c) > P(b)$	net 'closeness' of string to center	net deviation, conformation, # of turns	$P(a) > P(c) > P(b)$
Garth	$P(c) > P(b) > P(a)$	distance between string ends	distance between ends, position of end	$\boldsymbol{P(c) = P(b) = P(a)}$
Harry	$P(c) = P(b) = P(a)$	fundamental assumption of statistical mechanics	distance between ends, net deviation, # of turns	$P(c) = P(b) = P(a)$

Fred: *I look at that as a macrostate, because it's kind of like the combined behavior of the whole string... So macro, because macro means large.*

Interviewer: *Can you talk about how you would define a macrostate versus a microstate?*

Fred: *Well, I guess when I think of a macrostate, I think of like the collective behavior of a bunch of individual parts.... I don't know exactly what I could think of a microstate as, besides, like, the individual location of a piece of string and it's displacement from the center.*

Alex had similar reasoning, though not as explicitly.

Alex: *So we're talking about like positions in this string like each of these individual positions I would consider as a microstate. Whereas the macrostate is, like... the configuration itself.*

This choice of the straight string in Fig. 1a as the most likely configuration was unexpected and resembles findings from Loverude and Geller that some students have an unexpected association of equilibrium (and high entropy) with an 'ordered' state [4, 10]. While our interview question did not directly relate to equilibrium and students did not bring up ideas about equilibrium, it is possible students followed a similar line of reasoning. Though, instead of associating equilibrium with an 'ordered' state, students here ranked an 'ordered' state with a higher probability of occurring. It is important to note, however, that both Alex and Fred arrived at the conclusion that the straight string as a logical conclusion of their 'centered-ness' metric. Further study is warranted to determine whether the similarity between Loverude's finding and ours is more than superficial.

B. Definition of Macrostates

Defining macrostates for the string required some creativity, and all the students interviewed articulated appropriate classification systems (see Tab. I). A total of five students brought up multiple relevant ideas for defining macrostates, though not all represented fully formed definitions.

Macrostate classifications given by the interview students fell into one of five emergent groups: the net deviation of the string from center ($N = 4$); a displacement between two points on the string ($N = 4$); the number of 'turns,' 'kinks,' or 'bends' in the string ($N = 4$); the actual conformation of the string shown in the figures ($N = 3$); and there was one mention of the location of the string's end—a position as opposed to a displacement.

In the three cases where the actual conformation of the string was identified as a macrostate, two answers revealed conceptual errors in differentiating microstates and macrostates, as discussed in the previous section. The other response, from Erik, was accompanied by an appropriate argument that microstates were the precise arrangement of atoms and electrons in the string: that there are many ways to arrange the atoms and electrons in the string to form that specific conformation. While not incorrect, this does reveal an unproductive scale from which to reason about this system.

For the majority of students who initially gave an unequal probability ranking (Alex, Beth, Fred, and Garth), the initial ranking metric went on to form the basis for their macrostate definitions. For Alex and Fred, ideas about net deviation and net 'closeness' to center lead to macrostate definitions involving the net deviation of the string to one side or the other. Beth's metric of 'wavyness' preceded a definition based on the number of turning points on the string. Garth's reasoning about net expected displacement in a random walk foreshadowed a definition based on the distance between the two ends of the string.

The conversion of these metrics into macrostate definitions demonstrated a skill in turning intuition into models. It also demonstrated an expert-like ability to shift into more correct modes of reasoning. However, this also supports our interpretation that students, at least initially, project properties of macrostates onto constituent microstates.

C.   Reasoning about the Entropy of the String

Student responses to this question reflected the open-ended nature of the prompt. It also proved the most difficult task of those discussed here, which echoes findings from Loverude and Bucy that students struggle connecting entropy and multiplicity [4, 13]. Four students responded in some way that related multiplicity of states to entropy. Three students clearly articulated this connection. For example,

> **Daana**: *Entropy of the string [...] in some macrostate would be related to the number of microstates in which that system belongs to that macrostate.*

A fourth student came close to this answer, but didn't articulate the connection as concisely as the others:

> **Chris**: *Like so, basically what I'm thinking is there's like, there's, I guess theoretically infinite ways that the string could be at this length, but it has to lie on that semicircle. Whereas there's even more infinite ways that the string could be this length because it could snake in some other weird way. And I suppose this length is slightly shorter. I think that's what they were going for. There are slightly more ways that it could sort of snake into that position, I guess is what they're kind of asking.*

Two students, Erik and Fred, reasoned that the entropy of the string was a constant. Erik argued that since each of the conformations were "not structurally distinguishable" that there was some kind of equivalence between each conformation that meant entropy didn't change if the conformation of the string changed. Fred stated he couldn't think of a reason why the entropies of the states should be different. He recognized that this contradicted his unequal ranking of the states, but ultimately stuck with the ranking for the sake of consistency with the 'centered-ness' argument. These students neglected the deeper connection between entropy and multiplicity, but seemed to nearly activate the fundamental assumption of statistical mechanics.

The final two students, Alex and Beth, both reasoned about entropy by breaking the string up into chunks and considering which way each chunk bent (*i.e.*, did the chunk deviate left, right, or not at all). Alex reasoned that a "mixed" sequence of chunks was more favorable, and Beth did not expect most of the chunks in a sub-sequence to 'do the same thing'. For Alex, the idea came from an analogy with a two-state Ising magnet system with which he was familiar.

Given that other researchers have found a strong preference for the association of entropy with 'disorder' among undergraduates [2, 4], perhaps it is surprising that only two graduate students employed this particular association when asked directly about the entropy of the string. However, terms of similar ambiguity like 'wavy,' 'wiggly,' and 'chaotic' were used by students in this study, though mostly as a qualitative handle in developing more precise descriptions. The fact that half of the students initially ranked the more 'disordered' states as more likely configurations suggests at least some association between probability and 'disorder.'

IV.   CONCLUSIONS

In this study, graduate students were presented with images of microstates of a novel system and asked to rank the probabilities of finding the system in each of the states. They were then asked if the images represented microstates or macrostates, to define a set of macrostates of the system, and then to re-rank the states. Finally, students were asked to reflect on the meaning of entropy specific to this system.

When presented with microstates, many interviewees did not initially recognize them as such, and demonstrated a preference for an erroneous 'intuitive ranking' of microstate probabilities. Many possible explanations could account for this difficulty, such as an intuitive association between probability and 'disorder' similar to the association of entropy with 'disorder'. Furthermore, there was evidence that students projected macrostate likelihood on constituent microstates. Two students, however, settled on an unexpected ranking, which resembled other research where students associated equilibrium (the most likely state of an isolated system) with more order. Both these students also seemed to conflate macrostates with macroscopic objects.

Many students corrected their 'intuitive ranking' after a prompt requiring them to consider the macrostates and microstates of the system. However, many students did not spontaneously invoke the fundamental assumption of statistical mechanics. This could relate to the fact that the principle does not generalize to systems of microstates with different energies, causing students to be wary applying the idea in a new context, or simply that students struggle to activate the principle in new contexts.

When tasked with defining a set of macrostates in a novel context, our sample of students overwhelmingly succeeded in providing relevant classifications, with 5 out of the 8 students at least touching on multiple such classifications. However, this sample of students struggled connecting multiplicity of microstates to entropy in the novel context, echoing the findings of other studies that this connection presents a particularly troublesome difficulty. Finally, the imprecise identification of entropy as disorder did not explicitly appear as much as expected, though some students came up with their own imprecise phrases specific for the context.

This work is limited by the small sample size and the selection effect created by the solicitation of participants. In particular, this study occurred at a large, R1 university with a selective graduate physics program. Future work will compare the responses of undergraduate and graduate students to the same interview using the same coding scheme.

[1] B.W. Dreyfus, B.D. Geller, D.E. Meltzer, and V. Sawtelle, Resource Letter TTSM-1: Teaching Thermodynamics and Statistical Mechanics in Introductory Physics, Chemistry, and Biology, Am. J. Phys., **83**, 1 (2015)

[2] B.R. Bucy, J.R. Thompson, and D.B. Mountcastle, What Is Entropy? Advanced Undergraduate Performance Comparing Ideal Gas Processes, AIP Conference Proceedings, **818**, 1 (2006).

[3] T.I. Smith, W.M. Christensen, D.B. Mountcastle, and J.R. Thompson, Identifying student difficulties with entropy, heat engines, and the Carnot Cycle, Phys. Rev. ST Phys. Educ. Res., **11**, 2 (2015)

[4] M. Loverude, Identifying student resources in reasoning about entropy and the approach to thermal equilibrium, Phys. Rev. ST Phys. Educ. Res., **11**, 2 (2015)

[5] R. Leinonen, M.A. Asikainen, and P.E. Hirvonen, Grasping the second law of thermodynamics at university: The consistency of macroscopic and microscopic explanations, Phys. Rev. ST Phys. Educ. Res., **11**, 2 (2015)

[6] W.M. Christensen, D.E. Meltzer, and C.A. Ogilvie, Student ideas regarding entropy and the second law of thermodynamics in an introductory physics course. Am. J. Phys., **77**, 907 (2009)

[7] E.M. Carson and J.R. Watson, Undergraduate students' understandings of entropy and Gibbs free energy, U. Chem. Educ., **6**, (2002)

[8] J. M. Bennett and M. Sözbilir, A study of Turkish chemistry undergraduates' understanding of entropy, J. Chem. Educ. **84**, 1204 (2007).

[9] J. Haglund, S. Andersson, and M. Elmgren, Chemical engineering students' ideas of entropy, Chem. Educ. Res. Pract., **16**, 537 (2015)

[10] B. D. Geller, B.W. Dreyfus, J. Gouvea, V. Sawtelle, C. Turpen, and E. F. Redish, Entropy and spontaneity in an introductory physics course for life science students, Am. J. Phys. **82**, 394 (2014).

[11] M. E. Fonteyn, B. Kuipers, and S. J. Grobe, A description of think aloud method and protocol analysis. Qualitative health research **3**, 4, 430-441 (1993)

[12] D. Hammer, Student resources for learning introductory physics, Am. J. Phys. **68**, (2000)

[13] B. R. Bucy, Ph.D. thesis, University of Maine, 2007.