

Stellar Label Estimates through use of Machine Learning

G. Leigh Harwell

June 2026

• ■

Red Giant Branch vs Red Clump Stars

Best classification method through Asteroseismology

Red Giant Stars

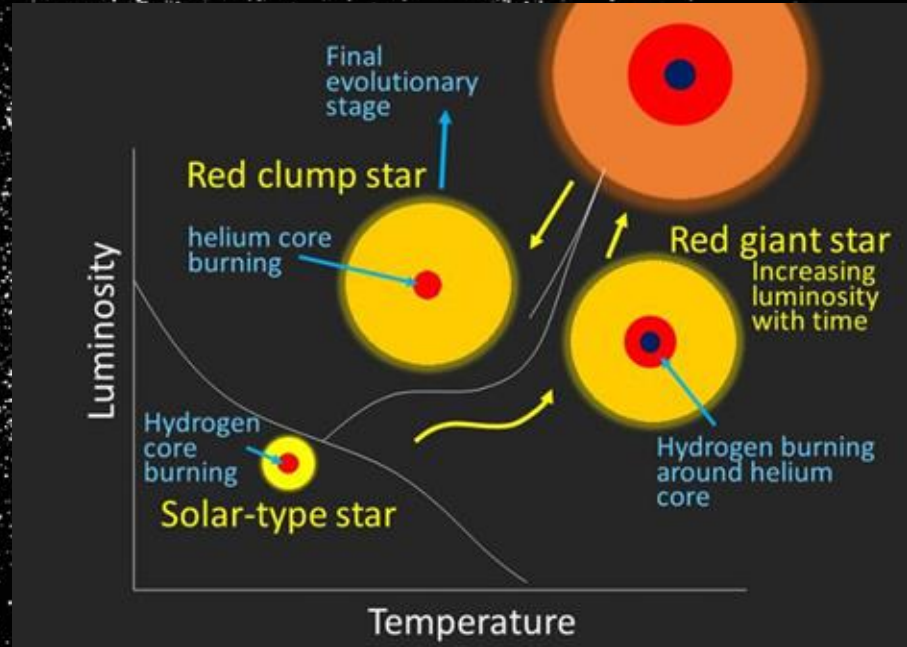
RGB

High mass Hydrogen burning shell
Low period spacing

Red Clump Stars

RC

High-Mass star that ignites helium fusion in their cores resulting in a "helium flash"
High period spacing



Red Giant Branch vs Red Clump Stars

Best classification method through Asteroseismology

Red Giant Stars

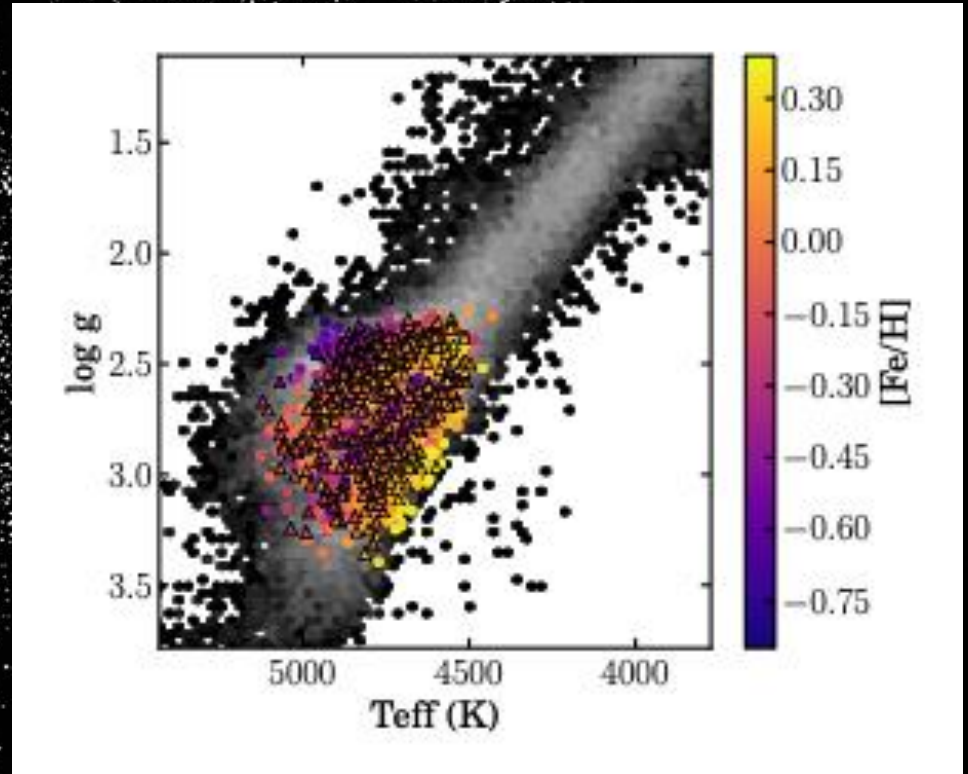
RGB

High mass Hydrogen burning shell
Low period spacing

Red Clump Stars

RC

High-Mass star that ignites helium fusion in their cores resulting in a "helium flash"
High period spacing

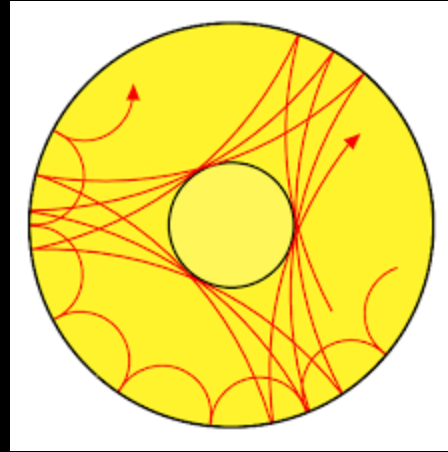


Asteroseismology

Solar-like oscillations in red giant stars arise from near-surface convection and can have either or both acoustic (p-mode) or gravity (g-mode) characteristics

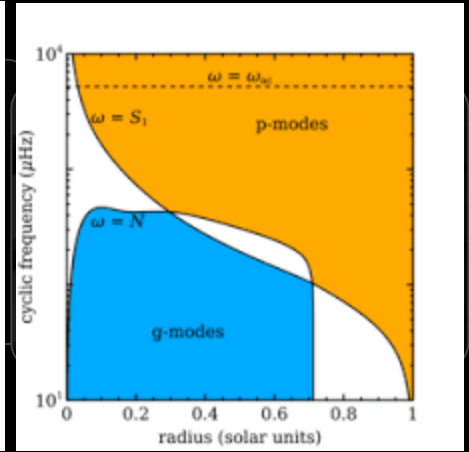
For RC stars the core density is lower than that of RGB stars of similar characteristics, which causes a significantly stronger coupling between g-mode and p-mode leading to appearance of observable 'mixed modes' in the oscillation spectrum.

This requires long space missions to acquire the data



P-mode

Associated with stellar envelope with pressure as a restoring force



G-mode

Probe its core with buoyancy as a restoring force.

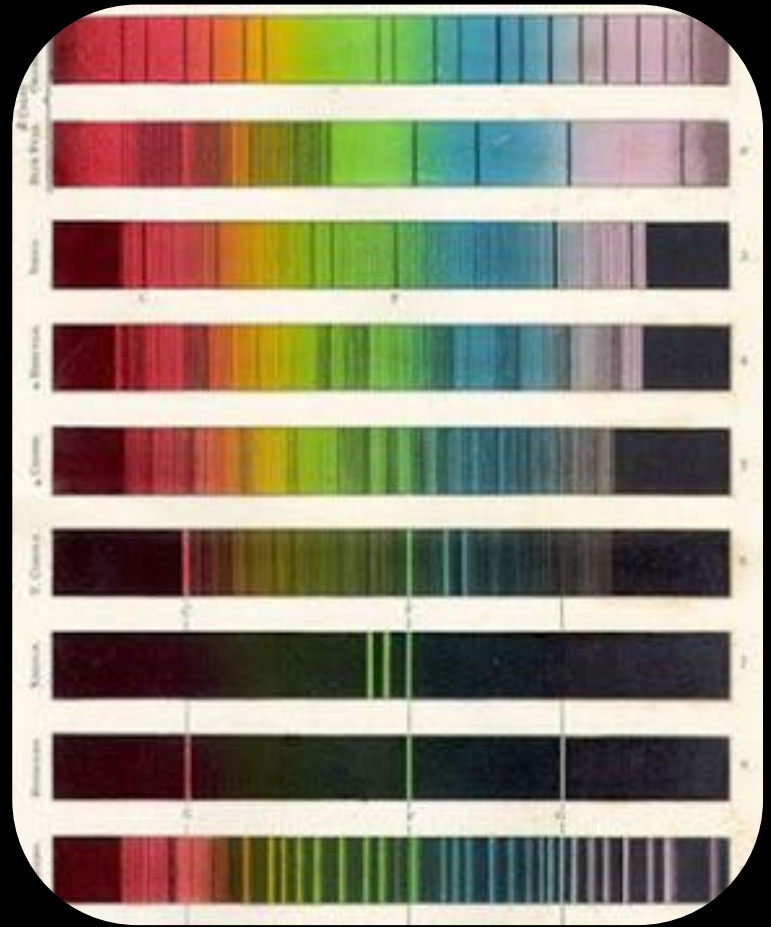
Bedding Tim, Huber Daniel. "ASTR 5970 Lecture 16-

Seismology." https://docs.google.com/presentation/d/1n4qnlVikQGskgZYAVGmYU1_9vY79EoGrD79NkOY4KYt/edit#slide=id.g2feb84dece6_0_122#slide=id.g2feb84dece6_0_122

Ness et al. 2016, "Spectroscopic Determination of Masses (and Implied Ages) for Red Giants." *The Astrophysical Journal*, 823:114 <https://opscience.idp.org/article/10.3847/0004-637X/823/2/114/pdf>

Main Problem

- Millions of spectra exist for known stars, with emission and absorption lines containing information about temperature, surface gravity, elemental abundances, and more
- This however this does not help us when it comes to differentiating between stars in the Red Giant Branch and in the Red Clump of the HR diagram
 - More luminous therefore more data
 - These are two different stages in a stars life cycle.
 - Important for: Stellar ages and Mass loss



The Cannon

Named after Dr. Annie Jump Cannon, American astronomer that logged spectra

Using a set of reference objects to train predictions on and then use to predict these labels for a new data set that does not include these labels in the survey.

Steps of use:

1. Get and format data
2. Use reference set to fit a spectral model (training)
3. Make sure the model is sensible, perform a leave-n-out cross-validation
4. Then use it to infer labels from all the other spectra in the dataset (testing)



Hoped results and current status:

Thus far we have a trained model of TheCannon. However, we are struggling to apply it to a new set of data.

We want TheCannon to be able to predict the period spacing accurately to be able to differentiate the two groups of red giants.

