# PHYS 142/242 Lab 01: DataHub, Jupyter, Python

Javier Duarte — January 9, 2023





### DataHub

- We will use DataHub for inclass hands-on portions
  - Recommend to use it for homework, final project, etc.
- Address: <u>datahub.ucsd.edu</u>
- Similar to public, free services Google Colab, but with access to better CPUs and GPUs and run by UCSD
- Provides a "Jupyter notebook" interface (Python-based but interactive coding like MATLAB/Mathematica)



### DATA SCIENCE / MACHINE LEARNING PLATFORM

### UC San Diego

Help - FAQ

Information Technology Services - Academic Technology Services



### UC San Diego Jupyterhub (Data Science) Platform

If you are unable to log in: Please try opening a private/incognito window in your browser | FAQ

### Student Resources

- Datahub/DSMLP Cluster Status
- Independent Study Access Request
- Data Science Resources
- Datahub/DSMLP Knowledge Base
  - Launching Containers from the Command Line
  - Configuring Your Container Launch
  - Building Your Own Custom Image

Transforring Eiles and Date

### Instructor Resources

- Request Datahub/DSMLP Instructional Technology Request (CINFO)
- Instructor Guidance for Datahub/DSMLP
- Educational Technology Services Instructional Github

Suno applicamento and gradoo with your Conve

- Blink Documentation
- Datahub Grading Tools
  - nbgrader



# Logging in

• After logging in, choose a course environment... PHYS 141 PHYS 241 - Comp Phys I: Prob Models & Simu - Duarte [SP23] jmduarte/phys141:latest (2 CPU, 4G RAM, 1 GPU)

📕 🔍 🔵 🔵 Jupyter⊢	lub-UCSD	×	· +			
$\leftarrow$ $\rightarrow$ G $\heartsuit$	datahub.uc	csd.edu/hub	o/spawn			
DAT	A SC	IENCE	/ MA	CHIN	E LEA	RNING F
Info	rmation T	echnology	Services	s - Acaden	nic Techno	logy Services
💭 Jupyter <mark>hub</mark>	Home	Token	DSML	Status	News	Services -
				S	elec	t Your
				0	Scientific ucsdets/s	Python + Machi cipy-ml-noteboo
				0	Scientific ucsdets/s	Python + Machi cipy-ml-noteboo
				0	RStudio:: I ucsdets/d	Data analysis w atascience-rstu
				0	DATAHUB Python 3,	: ucsdets/datas nbgrader (2 CPI
				0	DATAHUB Python 3,	: ucsdets/scipy nbgrader (2 CPI
				0	DATAHUB ucsdets/d	: RStudio: Data atascience-rstu
				0	scipy-ml: \$	Scientific Pytho



### Jupyter interface

• Spawns a "JupyterHub" (let's go step by step through all the buttons)

File Running Custers Curses Announcements Assignments     Select items to perform actions on them.     Upter Name + Last Modified File aize   c expstone-particle-physics-domain 6 months ago   c phys141 3 months ago	Control Panel     Files   Running   Clusters   Courses   Ansagrments     Select terms to perform actions on them.     Image: Clusters	C 合 合 datah	ub.ucsd.edu/user/jduar	rte/tree?						☆ □ (	👼 Incogi
File       Running       Clusters       Carres       Anouncements       Alore of Carres         Image: Clusters       Image: Clusters <th>Files       Runing       Clusters       Curses       Ansalgments         Sete tenes to perform actions on them.       Upload       Name 4       Last Modified       File size         C       capatone-particle-physice-domain       G       months ago         D       physi41       3       months ago</th> <th>💭 Jul</th> <th>oyter<mark>hub</mark></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th>Logout</th> <th>ontrol Panel</th> <th></th>	Files       Runing       Clusters       Curses       Ansalgments         Sete tenes to perform actions on them.       Upload       Name 4       Last Modified       File size         C       capatone-particle-physice-domain       G       months ago         D       physi41       3       months ago	💭 Jul	oyter <mark>hub</mark>						Logout	ontrol Panel	
Select items to perform actions on them.	Select items to perform actions on them.	Files	Running Clusters	Courses A	nnouncements	Assignments					
□ □ ↓       Name ↓       Last Modified       File size         □ □ capstone-particle-physics-domain       6 months ago         □ □ phys141       3 months ago	Image: Comparison of the size         Image: Comparison of the size </td <td>Select iter</td> <td>ns to perform actions on th</td> <td>nem.</td> <td></td> <td></td> <td></td> <td></td> <td>Upload</td> <td>New - 2</td> <td>;</td>	Select iter	ns to perform actions on th	nem.					Upload	New - 2	;
C capstone-particle-physics-domain       6 months ago         D phys141       3 months ago	c       capatone-particle-physics-domain       6 months ago         c       c       phys141       3 months ago	0	✓ ■ /					Name 🕹	Last Modified	File size	
D         phys141         3 months ago	phys141         3 months ago		capstone-particle-physic	s-domain					6 months ago	)	
			) phys141						3 months ago	)	

### Start coding!

Coding in Jupyter notebooks

Home Page - Select or create a X phys24	I_test - Jupyter Notebo 🗙 🖉 1.6 Appendix A. Complex numb 🗙 🏼 🄌 python - Does Numpy automati 🗙	G cupy version - Google Search × + ∽
$\leftarrow$ $\rightarrow$ C $\triangle$ $\triangleq$ datahub.ucsd.edu/user/jduarte/note	books/phys241_test.ipynb	☆ 🗖 👼 Incognito (3) 🚦
Cjupyterhub phys241_te	est Last Checkpoint: a few seconds ago (unsaved changes)	Logout Control Panel
File Edit View Insert Cel	Kernel Widgets Help	Trusted Python 3 (clean) O
	un 📕 C 🕨 Code 🗸 🖾 Show Usage Validate 🛄	
In [1]: print("hello world hello world	d")	
In [2]: import numpy as r	p	
<pre>In [3]: # matrix multipl: a = np.array([[1,</pre>	<pre>cation 2, 3], 3, 2], 1, 2], 2, 2]])</pre>	
Out[3]: (4, 3)		
<pre>In [4]: b = np.array([[1]</pre>	, , ])	
Out[4]: (3, 1)		
In [5]: a@b		
Out[5]: array([[ 8], [10], [ 5], [ 8]])		
In [6]: from scipy.specia from math import	l <b>import</b> hermite factorial	

## Plotting

• Plotting can be done easily with Matplotlib

🛛 🔍 💭 Но	ome Page - Select or create a 🗙 😑 phys241_test - Jupyter Notebo 🗙 🥑
$\leftarrow \rightarrow$ C $\triangle$	datahub.ucsd.edu/user/jduarte/notebooks/phys241_test.ipynb
	<b>Jupyterhub</b> phys241_test Last Checkpoint: 2 minutes a
	File Edit View Insert Cell Kernel Widgets He
	E + ∞ 2 E + ↓ Frances C → Code
	<pre># Range of x determine by classical tun xmin, xmax = -np.sgrt(2 * E(VMAX)), np.: x = np.linspace(xmin, xmax, 10000) fig, ax = plt.subplots(figsize=(5,5)) for v in range(VMAX): # plot potential V(x) ax.plot(x, V(x), color='black') # plot psi squared which we shift up ax.plot(x, psi(v,x) ** 2 + E(v), lw: # add lines and labels ax.axhline(E(v), color='gray', lines ax.set_xlabel('x') ax.set_ylabel('\$\psi^2_n(x)\$') plt.show()</pre>

1.6 Appendix A. Complex numb 🗙 🛛 🏄 python - Does Numpy automati 🗙	G cupy ve	rsion -	- Google Search	×	+	~
			*		lncognito (3)	):
go (autosaved)	<b>?</b>	Logo	ut Control Pane	əl		
2	Trusted	<b>A</b>	Python 3 (clean)	0		
Validate Show Usage Validate						
<i>ing points:</i> qrt(2 * E(VMAX))						
by values of energy 2)						
tyle='')						

# Installing a missing library

- Not all libraries are preloaded, but it's easy to install a new one
- Note: restart your "kernel" after doing this

In [1]:	<pre>!pip install cupy-cuda112user</pre>
	Collecting cupy-cudal12 Using cached cupy_cudal12-10.3.1-cp39-cp39-manylinux1_x86_6 Collecting fastrlock>=0.5 Using cached fastrlock-0.8-cp39-cp39-manylinux_2_5_x86_64.m _64.whl (49 kB) Requirement already satisfied: numpy<1.25,>=1.18 in /opt/cond 5) Installing collected packages: fastrlock, cupy-cudal12 Successfully installed cupy-cudal12-10.3.1 fastrlock-0.8

# t's easy to install a new one ng this

64.whl (78.9 MB)

manylinux1\_x86\_64.manylinux\_2\_12\_x86\_64.manylinux2010\_x86

da/lib/python3.9/site-packages (from cupy-cuda112) (1.19.

### Speed up numerical calculations with the GPU

Many libraries to do this: TensorFlow, CuPy, PyTorch, …

C ☆ 🌢 datahub.uc	d.edu/user/jduarte/notebooks/phys241_test.ipynb	☆ 🗖 👼 Incogi						
💭 jupyte	hub phys241_test Last Checkpoint: 16 minutes ago (autosaved)	Logout Control Panel						
File Edit	View Insert Cell Kernel Widgets Help	Not Trusted Python 3 (clean) O						
In [2]	import cupy as cp							
In [18]	a = cp.ones((10000, 10000)) b = cp.ones((10000, 10000))							
In [19]	%%time a@b							
	CPU times: user 2.34 ms, sys: 1.21 ms, total: 3.55 ms Wall time: 2.1 ms							
Out[19]	<pre>array([[10000., 10000., 10000.,, 10000., 10000., 10000.],       [10000., 10000., 10000.,, 10000., 10000., 10000.],       [10000., 10000., 10000.,, 10000., 10000., 10000.],       [10000., 10000., 10000.,, 10000., 10000., 10000.],       [10000., 10000., 10000.,, 10000., 10000., 10000.]])</pre>							
In [16]	<pre>import numpy as np a = np.ones((10000, 10000)) b = np.ones((10000, 10000))</pre>							
In [17]	In [17]: %%time a@b							
	CPU times: user 2min 21s, sys: 1min 21s, total: 3min 43s Wall time: 33.1 s							
<b>Out[17</b> ]	array([[10000., 10000., 10000.,, 10000., 10000., 10000.], [10000., 10000., 10000.,, 10000., 10000., 10000.], [10000., 10000., 10000.,, 10000., 10000., 10000.],							
	••••							



# **Exiting / shutting down server**

• When you're done; hit control panel and "stop your server" When you start it again, all your data will still be there



xa X C JupyterHub-	× C Home Pa	age ×   🗐 phy	vs241_tes × │ >	_ https://datah ×	G cupy	/s num X	+	~ :
					A			•
PLATFORM				UC Sa	n Di	ego		
es					Help 🗕	FAQ		
~							C Logout	
Stop My Server	My Server							

# **Data Science & Machine Learning Platform (DSMLP)**

- UCSD Data Science & Machine Learning Platform (DSMLP)
  - Based on docker and Kubernetes (K8s): open-source system for automating deployment, scaling, and management of containerized applications



- **Overview:** 
  - You log in to a remote "login" node
  - From that login node, you can launch a "pod" running a container
    - The pod automatically starts a "JupyterHub" web server (only accessible from campus so you must VPN if off campus)
    - You are also automatically logged into that pod so you can run interactive terminal commands, etc.



### **More Resources**

- UCSD DSMLP Cluster
  - Using the DSMLP server
    - <u>Customizing the DSMLP Containers</u>
- Terminal and Command-Line Interface
  - Bash Scripting Reference
  - <u>Git(Hub) Resources</u>