[2] The Vector

The Vector: William Rowan Hamilton

By age 5, Latin, Greek, and Hebrew
By age 10, twelve languages including Persian, Arabic, Hindustani and Sanskrit.

William Rowan Hamilton, the inventor of the theory of quaternions... and the plaque on Brougham Bridge, Dublin, commemorating Hamilton's act of vandalism.

$$i^2 = j^2 = k^2 = ijk = -1$$

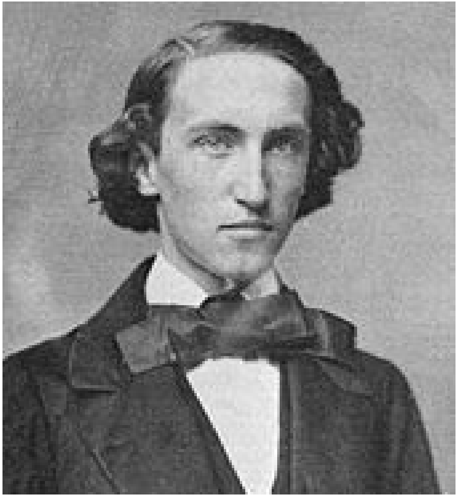
And here there dawned on me the notion that we must admit, in some sense, a fourth dimension of space for the purpose of calculating with triples ... An electric circuit seemed to close, and a spark flashed forth.

The Vector: Josiah Willard Gibbs

Started at Yale at 15
Got Ph.D. at Yale at 24
(1st engineering doctorate in US)
Tutored at Yale
Spent three years in Europe
Returned to be professor at Yale

Developed *vector analysis* as an alternative to quaternions.

His unpublished notes were passed around for twenty years.



“Professor Willard Gibbs must be ranked as one of the retarders of ... progress in virtue of his pamphlet on *Vector Analysis*; a sort of hermaphrodite monster.” (Peter Guthrie Tait, a partisan of quaternions)

What is a vector?

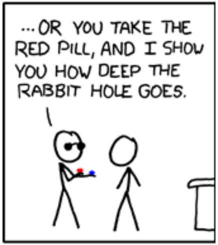
- This is a 4-vector over $\mathbb{R}$:

$$[3.14159, 2.718281828, -1.0, 2.0]$$

- We will often use Python's lists to represent vectors.

- Set of all 4-vectors over $\mathbb{R}$ is written $\mathbb{R}^4$.

- This notation might remind you of the notation $\mathbb{R}^D$: the set of functions from D to $\mathbb{R}$.



Vectors are functions

- Think of our 4-vector $[3.14159, 2.718281828, -1.0, 2.0]$ as the function

$$\begin{array}{rcl} 0 & \mapsto & 3.14159 \\ 1 & \mapsto & 2.718281828 \\ 2 & \mapsto & -1.0 \\ 3 & \mapsto & 2.0 \end{array}$$

- $\mathbb{F}^d$ is notation for set of functions from $\{0, 1, 2, \ldots, d-1\}$ to $\mathbb{F}$.
- **Another example:** $GF(2)^5$ is set of 5-element bit sequences, e.g. [0,0,0,0,0], [0,0,0,0,1], ...
- Let WORDS $=$ set of all English words
- In information retrieval, a document is represented ("bag of words" model) by a function

$$f : WORDS \longrightarrow \mathbb{R}$$

specifying, for each word, how many times it appears in the document.

Representation of vectors using Python dictionaries

- We often use Python's dictionaries to represent such functions, e.g. `{0:3.14159, 1:2.718281828, 2:-1.0, 3:2.0}`.
- What about representing a WORDS-vector over $\mathbb{R}$?
- For any single document, most words are *not* represented. They should be mapped to zero.
- Our convenient convention for representing vectors by dictionaries: allowed to omit key-value pairs when value is zero.
- **Example:** "The rain in Spain falls mainly on the plain" would be represented by the dictionary

  ```
  {'on': 1, 'Spain': 1, 'in': 1, 'plain': 1, 'the': 2,
   'mainly': 1, 'rain': 1, 'falls': 1}
  ```

Sparsity

- A vector most of whose values are zero is called a *sparse* vector.
- If no more than k of the entries are nonzero, we say the vector is k-*sparse*.
- A k-sparse vector can be represented using space proportional to k.
- **Example:** when we represent a corpus of documents by WORD-vectors, the storage required is proportional to the total number of words in all documents.
- Most signals acquired via physical sensors (images, sound, ...) are not exactly sparse.
- Later we study *lossy compression*: making them sparse while preserving perceptual similarity.

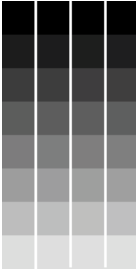
What can we represent with a vector?

- Document (for information retrieval)
- Binary string (for cryptography/information theory)
- Collection of attributes
  - Senate voting record
  - demographic record of a consumer
  - characteristics of cancer cells
- State of a system
  - Population distribution in the world
  - number of copies of a virus in a computer network
  - state of a pseudorandom generator
  - state of *Lights Out*
- Probability distribution, e.g. {`1:1/6, 2:1/6, 3:1/6, 4:1/6, 5:1/6, 6:1/6`}

What can we represent with a vector?

- Image

```
{(0,0): 0,      (0,1): 0,      (0,2): 0,      (0,3): 0,
 (1,0): 32,     (1,1): 32,     (1,2): 32,     (1,3): 32,
 (2,0): 64,     (2,1): 64,     (2,2): 64,     (2,3): 64,
 (3,0): 96,     (3,1): 96,     (3,2): 96,     (3,3): 96,
 (4,0): 128,    (4,1): 128,    (4,2): 128,    (4,3): 128,
 (5,0): 160,    (5,1): 160,    (5,2): 160,    (5,3): 160,
 (6,0): 192,    (6,1): 192,    (6,2): 192,    (6,3): 192,
 (7,0): 224,    (7,1): 224,    (7,2): 224,    (7,3): 224 }
```



What can we represent with a vector?

- Points
  - Can interpret the 2-vector $[x, y]$ as a point in the plane.
  - Can interpret 3-vectors as points in space, and so on.