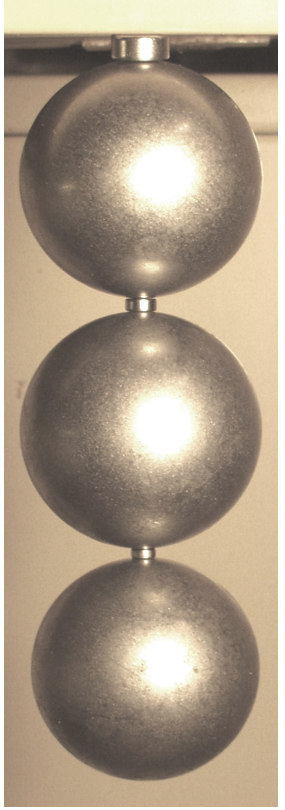
Mathematical Models of Imaging

Magnetic Resonance Imaging (MRI)

Suyash P. Awate

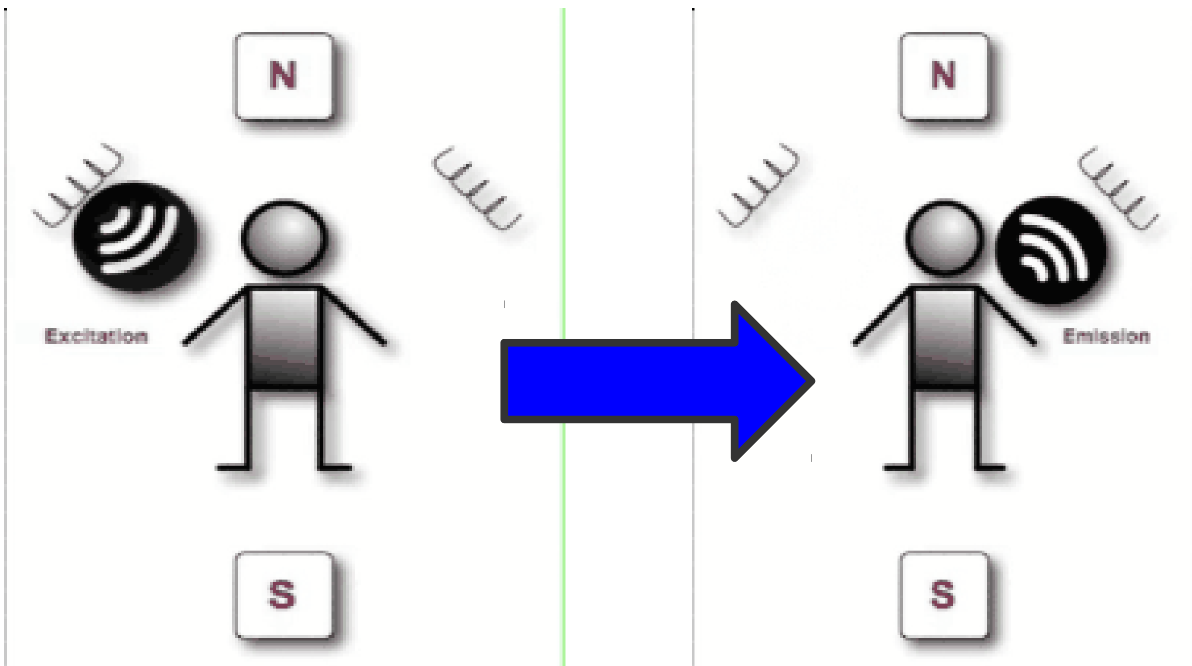
MRT / MRI

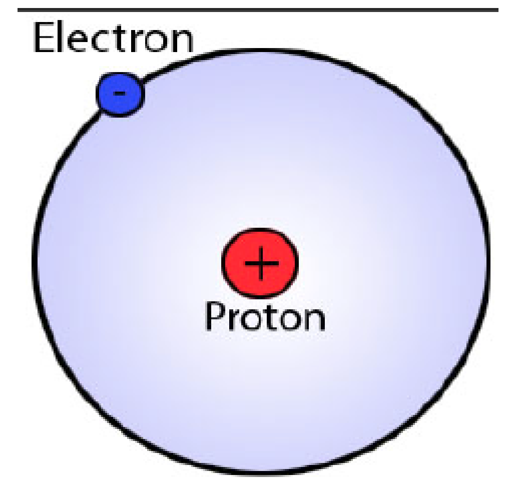
- Acquisition process
  - Place patient inside a large strong magnetic field
    - How strong ? (unit for magnetic-flux density: tesla)
      - Earth's magnetic field on surface is 25-65 micro-tesla
      - Refrigerator magnet (iron alloy) → 5 milli-tesla
      - Neodymium (alloy; in MR machine) magnet → 1.25 tesla
        - Coin sized magnet can lift 9 kg
    - How large ?
      - Bore size holds a human inside
  - Use additional time-varying and space-varying electromagnetic fields
    - Need precise control over fields and timings
  - Acquire data slice by slice (tomography)



MRI

- Acquisition process
  - Use radio-frequency (RF) electromagnetic waves to excite Hydrogen nuclei
  - Nuclei release energy, which is measured





  - Model effects of electromagnetic field on atomic nuclei using differential equations

MRI

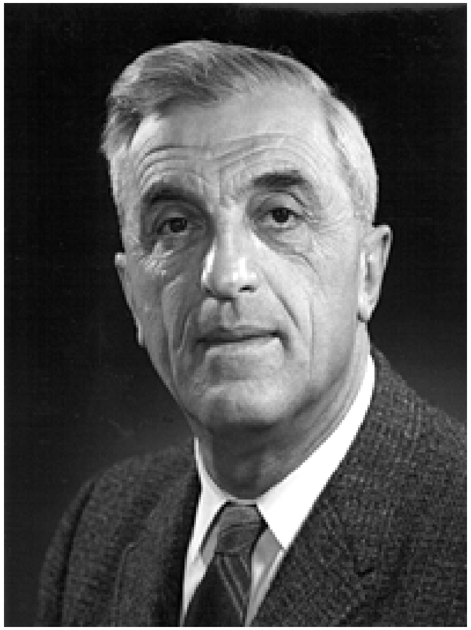
- Human body
  - Lots of Hydrogen nuclei distributed everywhere
  - We are ~53% water by mass
  - 11% Hydrogen by mass
    - 65% Oxygen, 18% Carbon
  - 67% Hydrogen by atomic percent
    - More Hydrogen atoms than any other element

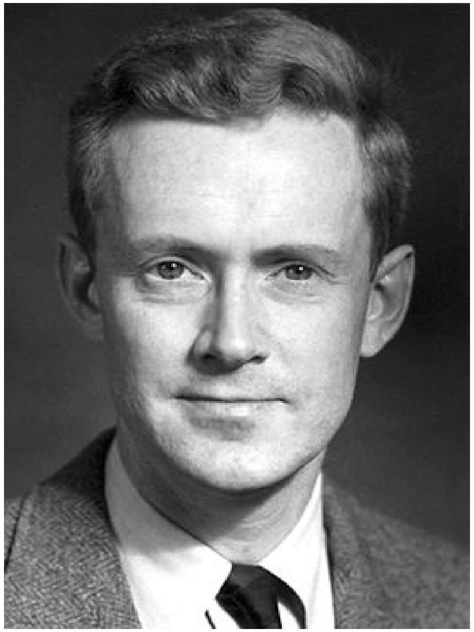
MRI

- Human body
  - Water content in:
    - Blood : 93%
    - Muscle : 70%
    - Brain
      - Gray matter : 71%
      - White matter : 84%
    - Heart : 80%
    - Liver : 70%
    - Kidneys : 80%
    - Bone : 12%

MRI

- History
  - Theory of nuclear magnetic resonance (1946)
    - Felix Bloch: classical physics
    - Edward Purcell: quantum physics
    - Interestingly
      - Bloch was trained as a quantum physicist (advisor Heisenberg)
      - Purcell was trained as a classical physicist
  - Bloch and Purcell got 1952 Nobel prize in physics





MRI

- Werner Heisenberg
  - Nobel Prize in Physics for 1932
    "for the creation of quantum mechanics"

CARNEGIE INSTITUTE OF TECHNOLOGY
SCHENLEY PARK
PITTSBURGH 13, PENNSYLVANIA

DEPARTMENT OF MATHEMATICS
COLLEGE OF ENGINEERING AND SCIENCE

February 11, 1948

Professor S. Lefschetz
Department of Mathematics
Princeton University
Princeton, N. J.

Dear Professor Lefschetz:

This is to recommend Mr. John F. Nash, Jr. who has applied for entrance to the graduate college at Princeton.

Mr. Nash is nineteen years old and is graduating from Carnegie Tech in June. He is a mathematical genius.

Yours sincerely,

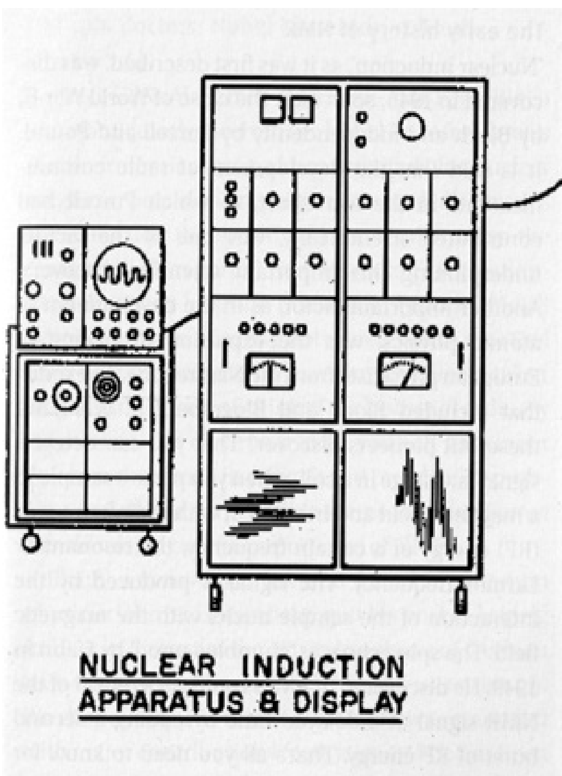
Richard J. Duffin

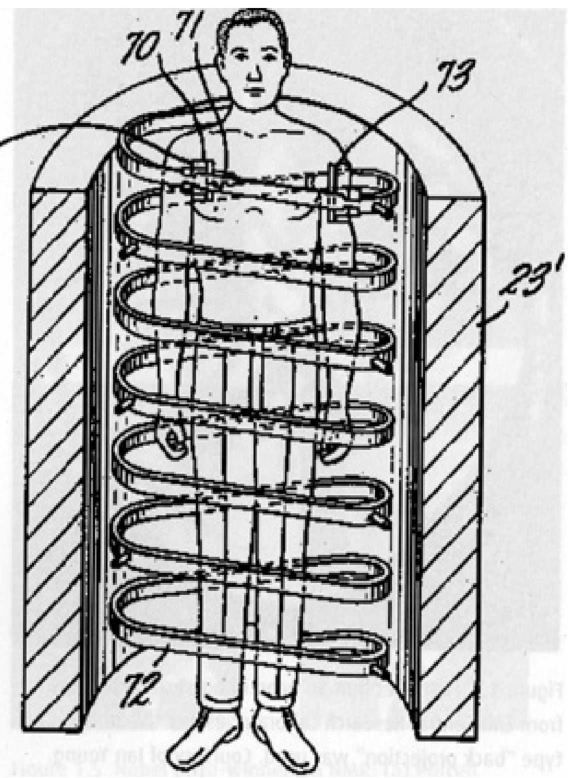
Richard J. Duffin

RJD:hl

MRI

- History
  - First 1D MR signal (1952)
    - Herman Carr (PhD Dissertation), Harvard



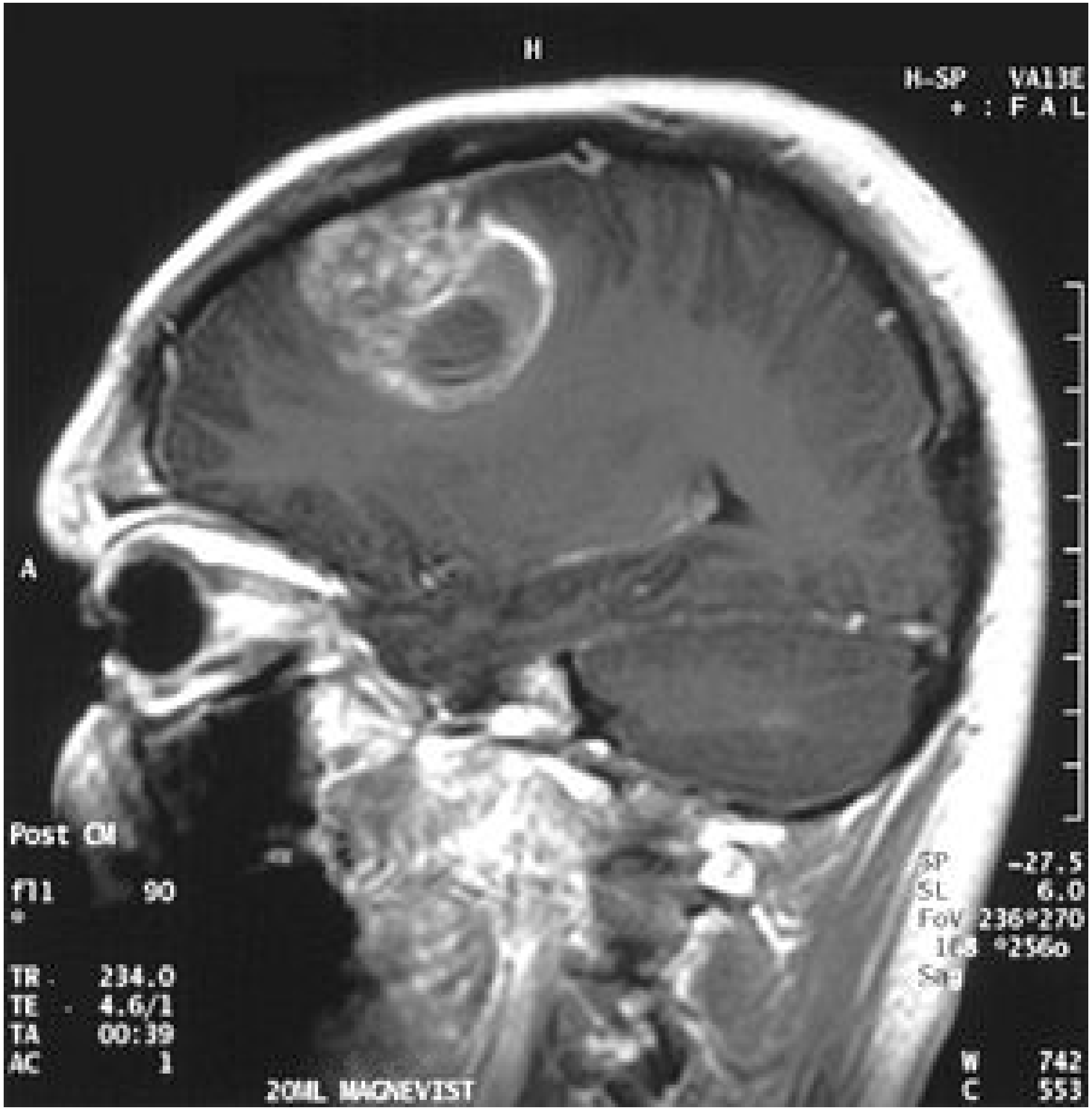


MRI

- History
  - Until ~1970, “images” weren't localized in space
  - First clinical application of MRI (1971)
    - Raymond Damadian (paper in *Science*)
  - First 2D-slice image (1970s)
    - Paul **Lauterbur** (Chemist)
    - Build on Herman Carr's work
    - Cross section of a living mouse (1974)
  - First mathematical analysis for faster imaging (1970s)
    - Peter **Mansfield** (Physicist)
    - Echo planar imaging (minutes instead of hours)
  - Lauterbur and Mansfield got 2003 Nobel prize in Medicine and Physiology

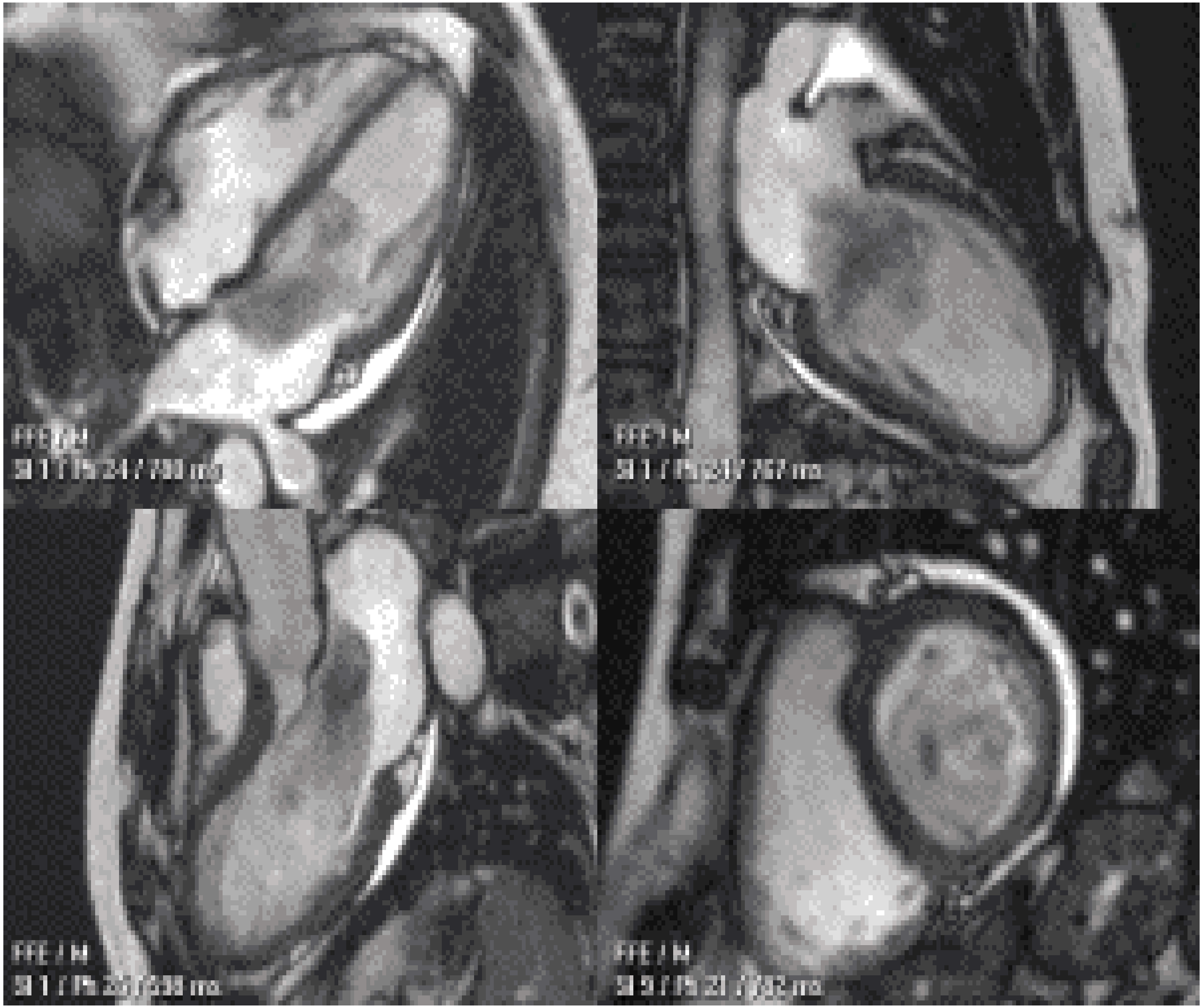
MRI

- Examples: Brain MRI
  - Investigate dementia, cerebrovascular disease, epilepsy, tumors



MRI

- Examples: Cardiac MRI
  - Cardiovascular disease



MRI

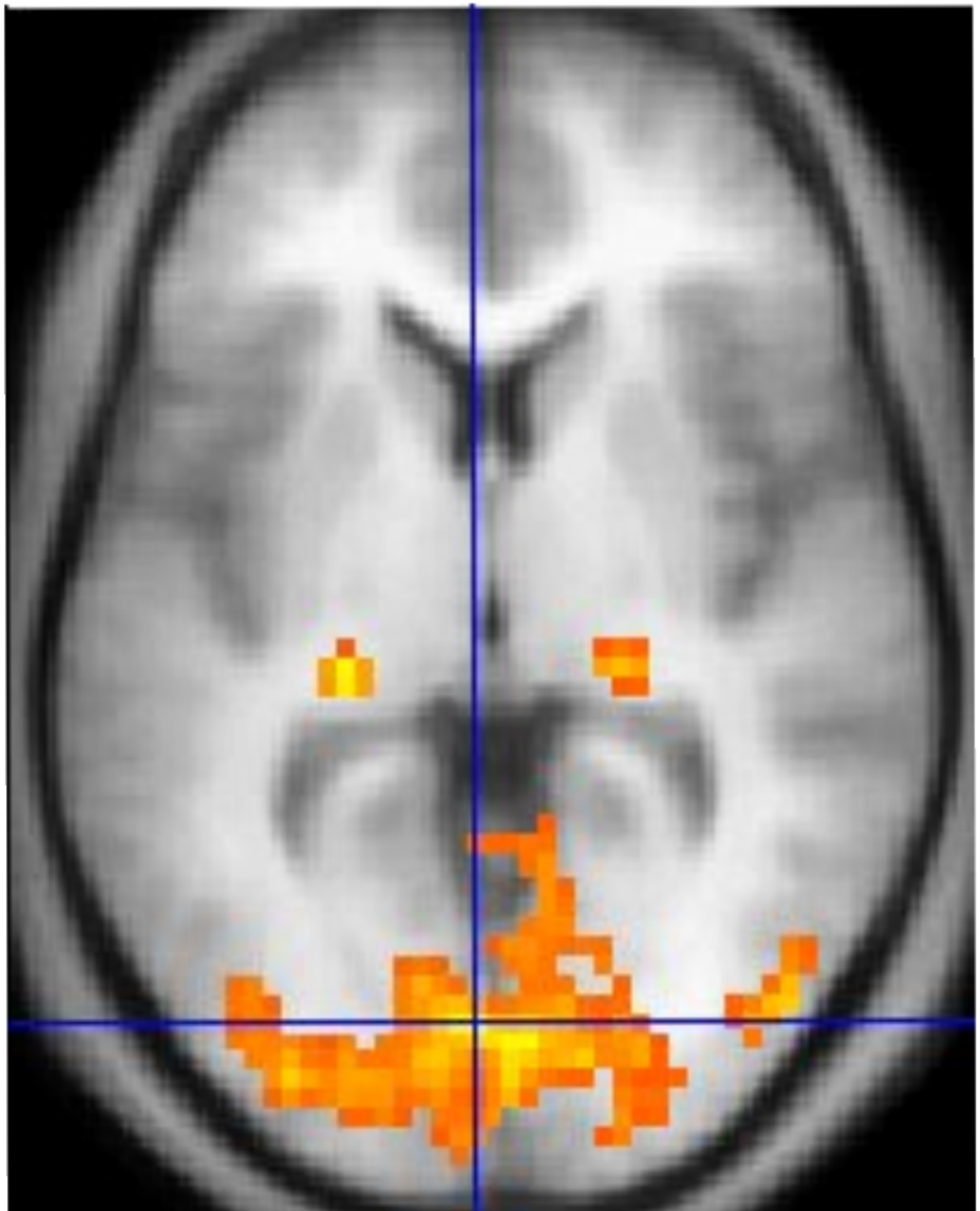
- Examples: Spinal MRI

MRI

- Examples: Angiography
  - Artery blockages

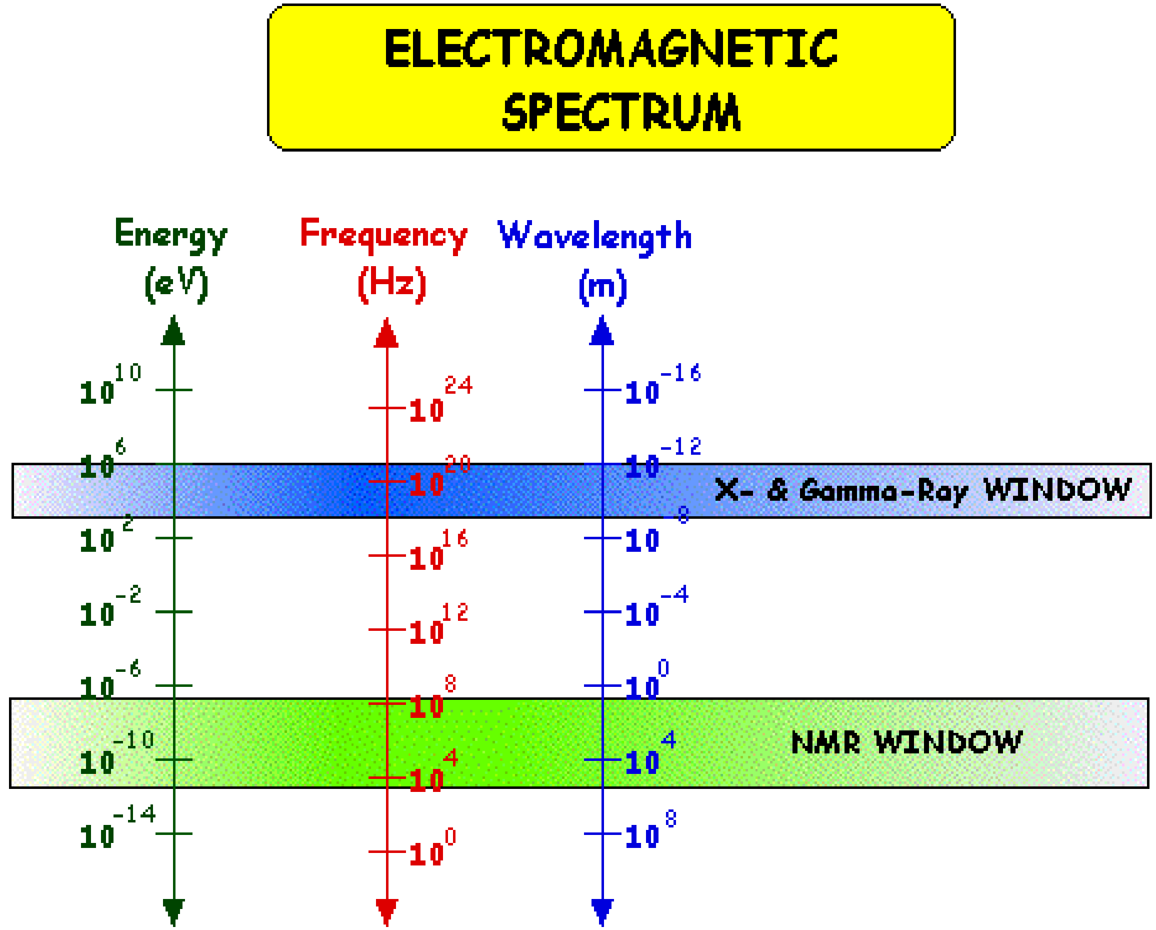
MRI

- Examples: Functional MRI
  - Activation in brain regions
  - Study diseased or injured brain



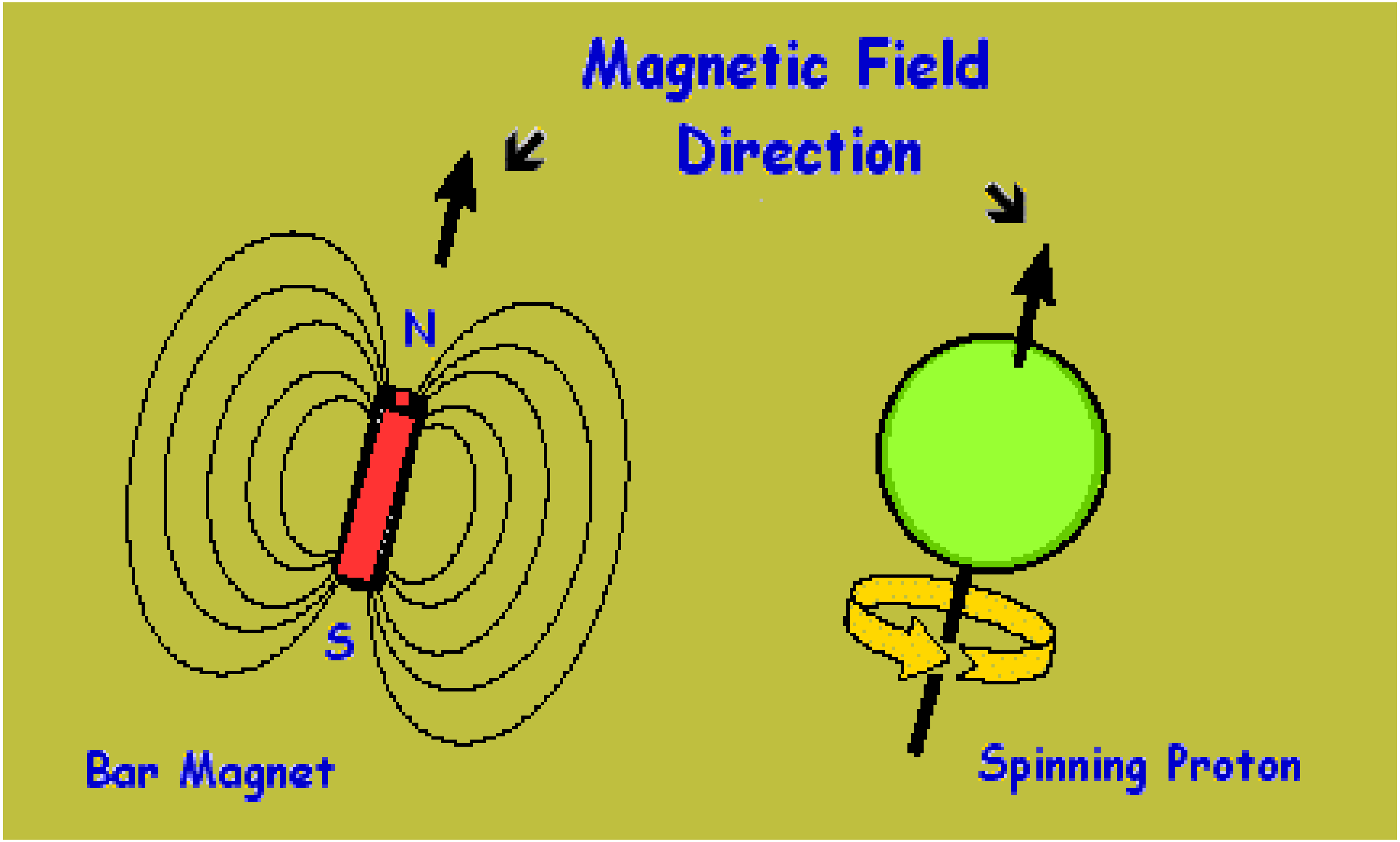
MRI

- MRI advantages
  - No ionizing radiation
  - Better details
    in soft tissues
    - Versus CT (bone)
- MRI disadvantages
  - Slower (20 minutes)
    - CT (5 minutes)
  - Machine expensive
  - Machine operation, maintenance needs more skill
  - What if patient has metallic implants ?
    - Cardiac pacemaker, ear implant



MRI

- Spinning proton
  - A moving electric charge acts like a magnet

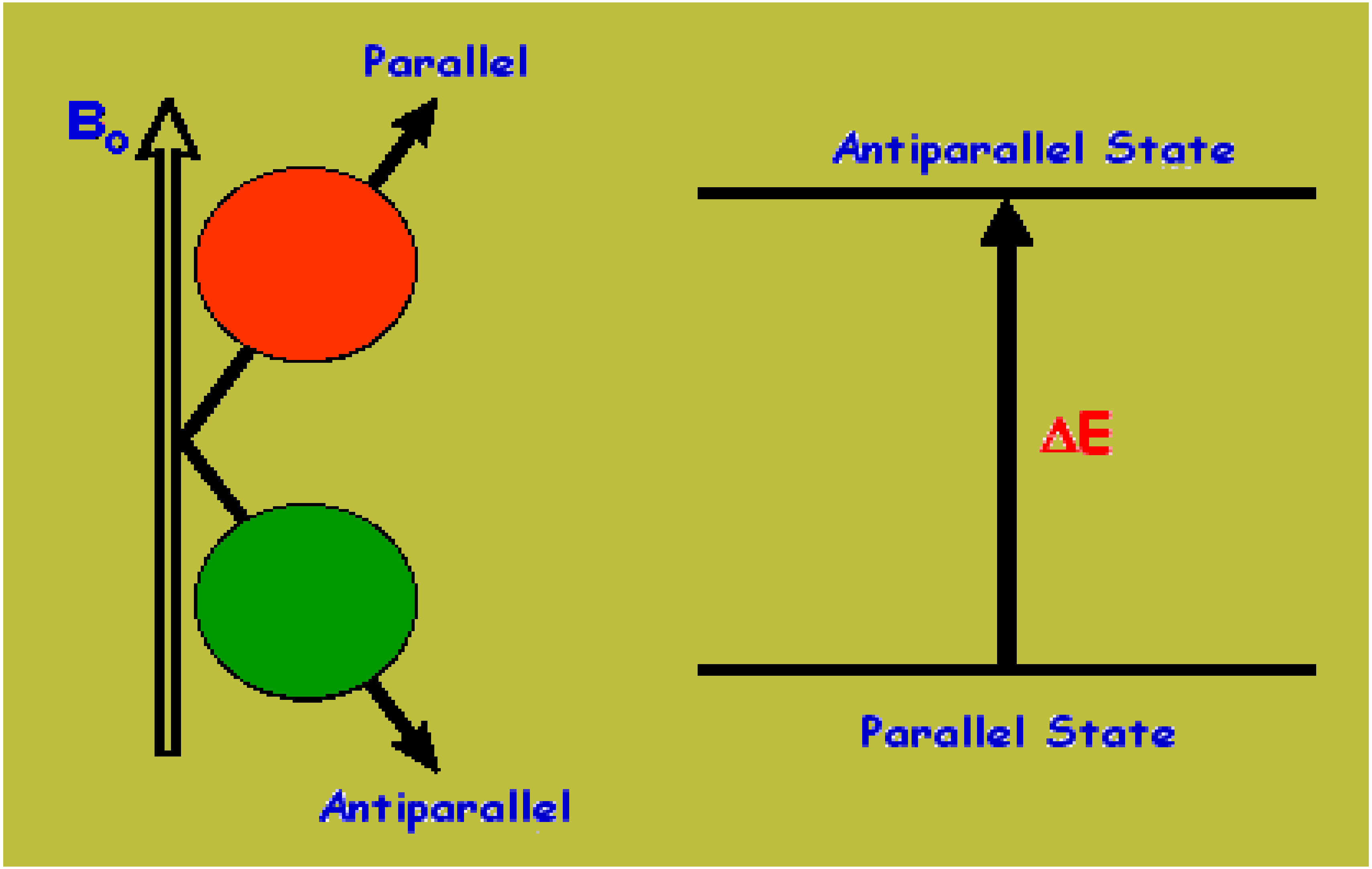


MRI

- Population of spinning protons
  - Assume non-magnetic material (body tissue)
  - Aggregate magnetization = 0
    - Magnets oriented randomly (uniformly) along directions

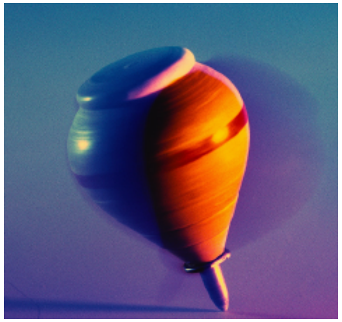
MRI

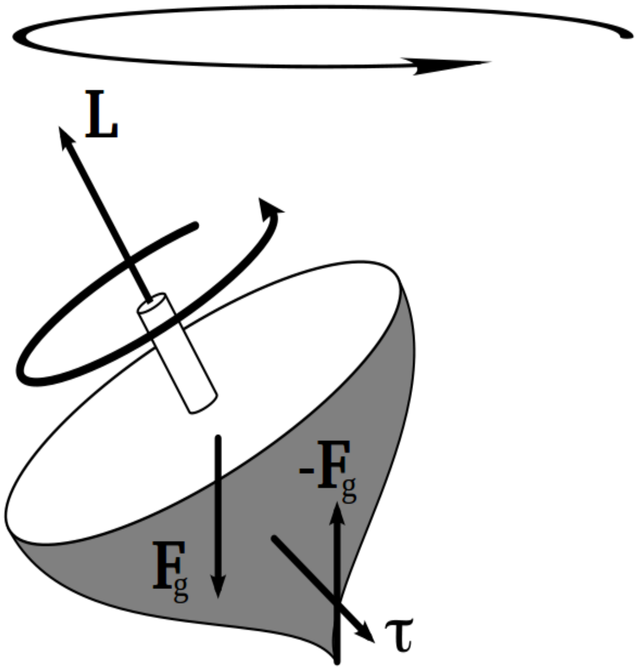
- Spinning proton in a static external magnetic field
  - (1) Starts aligning itself along or opposite the field
    - Alignment along the field is lowest energy state

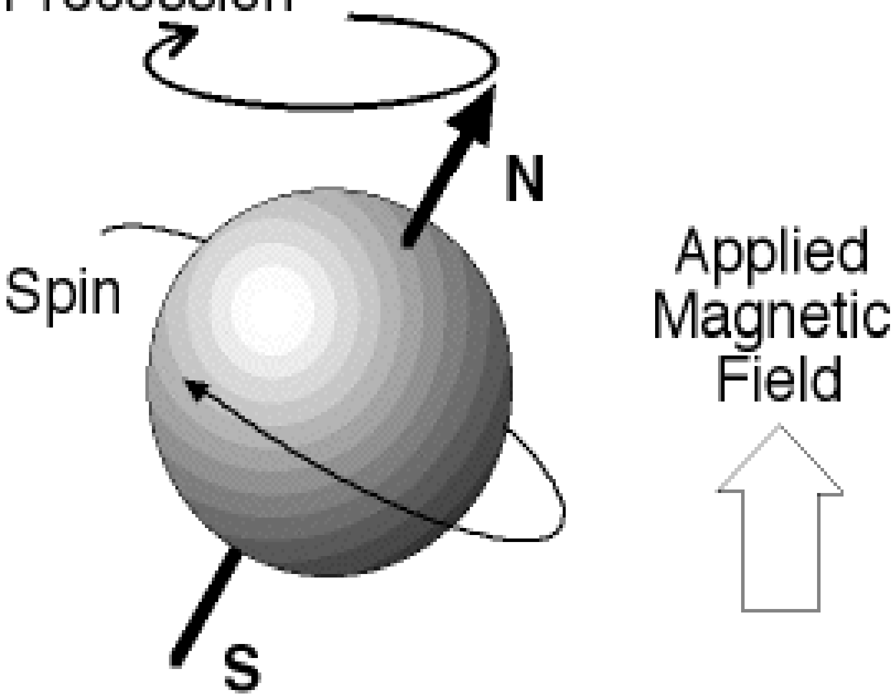


MRI

- Spinning proton in a static external magnetic field
  - (2) Precesses
    - Rotates around a central axis
    - Central axis itself rotates (= precesses) around axis of external field

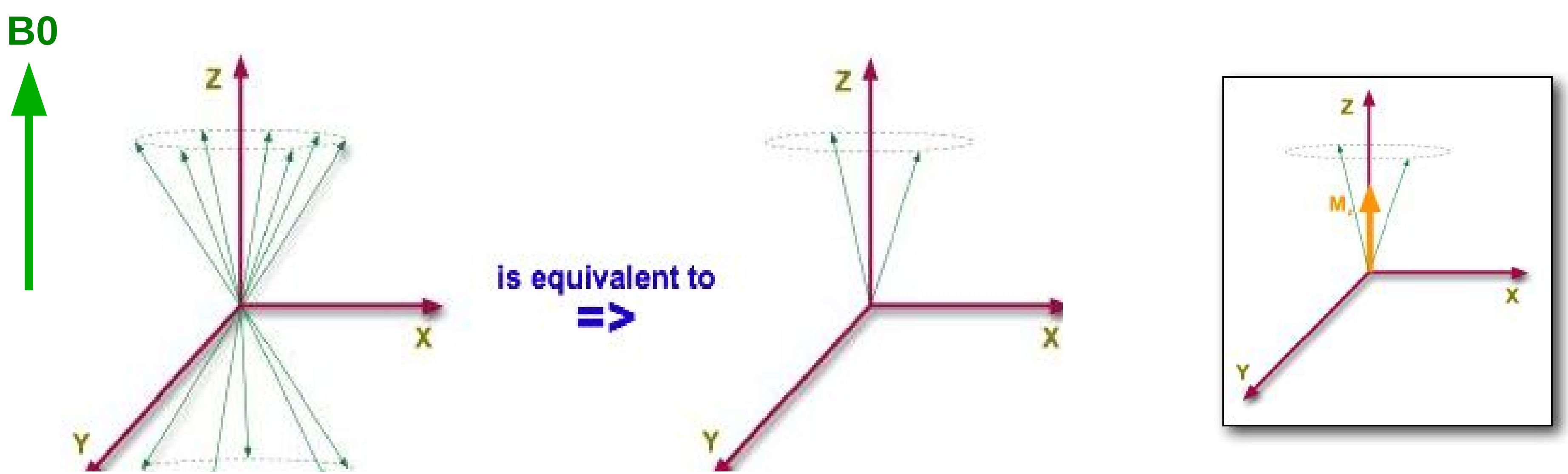






MRI

- A population of spinning protons in a static external magnetic field (equilibrium)
  - Spins become “polarized”
    - Precess about the +Z or -Z axes
    - More precess around the +Z axis
    - Each proton's precession has random phase
  - Aggregate **magnetization (M)** along static field (B0)



MRI

- **Bloch Equation**
  - What is the process for reaching equilibrium ?
  - What happens when the equilibrium is disturbed by introducing an additional time-varying magnetic field ?
    - How does the population react ?
    - What happens to the aggregate magnetization ?
    - Depends on the kind of tissue = chemical environment
  - Conventions and jargon
    - Axis along B0 = +Z axis = longitudinal axis
    - Plane perpendicular to B0 = XY plane = transverse plane

MRI

- Bloch Equation
  - Models rate of change of aggregate magnetization
  - Captures:
    - Precession of the aggregate magnetization
      - About the overall time-varying magnetic field
    - Transverse and longitudinal “relaxation”
    - Equilibrium state
  - Phenomenological equations
    - Consistent with fundamental theory
    - Not derived from fundamental theory
    - Modeling using empirical observations
    - Bloch 1946

MRI

- **Bloch Equation**
- **B0** := a strong uniform steady magnetic field; oriented along Z axis
- Now, consider additional magnetic fields are introduced that disturb the equilibrium
- **B(t,p)** := total **external** magnetic field experienced by sample at **time t** and **location p**
- **M(t,p)** := aggregate **magnetic moment** of nuclei in sample at time t and location p
- B(t,p), M(t,p) are vectors in 3D
  - B = [bx by bz]' ; M = [mx my mz]'

MRI

- Bloch Equation

$$\frac{dM}{dt} = \gamma M \times B - \frac{\langle m_x, m_y, 0\rangle}{T_2} - \frac{\langle 0, 0, m_z - m_{eq}\rangle}{T_1}$$

where γ, T_1, T_2, m_{eq} are constants.

- γ → biochemical nature of object
  - For H protons in $H_2 0$, γ ~ 43 MHz per Tesla
- T1, T2 → tissue properties
  - Tissue = magnetic nuclei contained within a specific chemical environment
- m_{eq} = constant
  = aggregate magnetic moment of nuclei in sample at equilibrium (i.e., $t \to \infty$) irrespective of tissue type

MRI

- Bloch Equation in **Static** Magnetic Field
- Let **B0 := <0, 0, b0>**
  - Independent of time t and location p
- B(t,p) = B0
- Then, Bloch equations reduce to:

$$\frac{dm_x}{dt} = \gamma b_0 m_y(t) - \frac{m_x(t)}{T_2}$$

$$\frac{dm_y}{dt} = -\gamma b_0 m_x(t) - \frac{m_y(t)}{T_2}$$

$$\frac{dm_z}{dt} = -\frac{m_z(t) - m_{eq}}{T_1}$$

$$\frac{dM}{dt} = \gamma M \times B - \frac{\langle m_x, m_y, 0 \rangle}{T_2} - \frac{\langle 0, 0, m_z - m_{eq} \rangle}{T_1}$$

where γ, T_1, T_2, m_{eq} are constants.

MRI

$$\frac{dm_x}{dt} = \gamma b_0 m_y(t) - \frac{m_x(t)}{T_2}$$

$$\frac{dm_y}{dt} = -\gamma b_0 m_x(t) - \frac{m_y(t)}{T_2}$$

$$\frac{dm_z}{dt} = -\frac{m_z(t) - m_{eq}}{T_1}$$

- Bloch Equation in **Static** Magnetic Field
- Solution of these equations is:

$$m_x(t) = e^{-t/T_2}\left(m_x(0)\cos(w_o t) - m_y(0)\sin(w_o t)\right)$$

$$m_y(t) = e^{-t/T_2}\left(m_x(0)\sin(w_o t) + m_y(0)\cos(w_o t)\right)$$

$$m_z(t) = m_z(0)e^{-t/T_1} + m_{eq}\left(1 - e^{-t/T_1}\right)$$

  where $w_0 = -\gamma b_0$ is called **Larmor frequency**
  - Depends on material (i.e., Hydrogen nuclei)
  - Angular frequency: +1 Hz = +2π radians / sec
- Initial magnetization: mx(0), my(0), mz(0)

MRI

- Bloch Equation in **Static** Magnetic Field

$$m_x(t) = e^{-t/T_2}\left(m_x(0)\cos(w_o t) - m_y(0)\sin(w_o t)\right)$$

$$m_y(t) = e^{-t/T_2}\left(m_x(0)\sin(w_o t) + m_y(0)\cos(w_o t)\right)$$

$$m_z(t) = m_z(0)e^{-t/T_1} + m_{eq}\left(1 - e^{-t/T_1}\right)$$

- Under **static magnetic field B0**, **during [t,t+Δt], aggregate magnetization M**:
  - Can have non-zero components along +Z axis, XY plane
    - Depends on initial conditions
  - Is **precessing** along +Z axis:
    - M_Z (t): magnitude stays same
    - M_{xy} (t): magnitude stays same, direction rotates around +Z
    - Note: individual nuclei precess around aggregate mag vector M

MRI

- Bloch Equation in **Static** Magnetic Field

$$m_x(t) = e^{-t/T_2}\left(m_x(0)\cos(w_o t) - m_y(0)\sin(w_o t)\right)$$

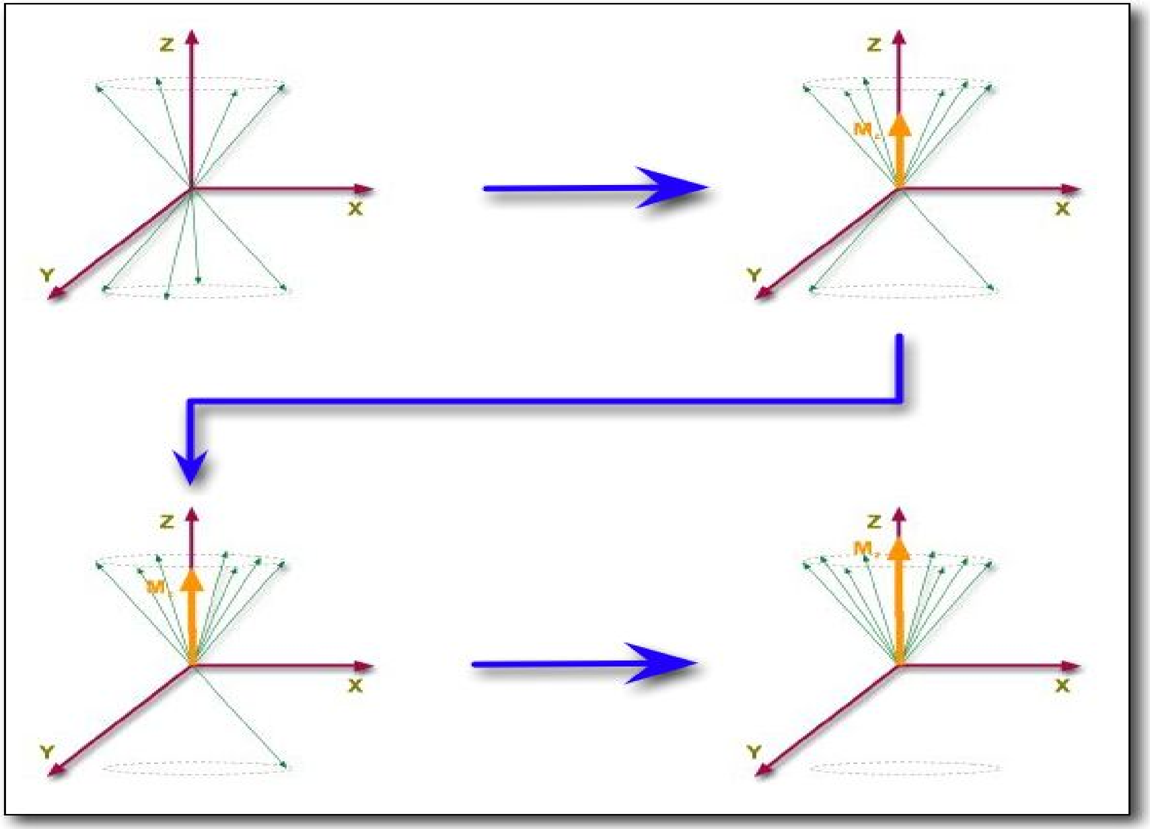
$$m_y(t) = e^{-t/T_2}\left(m_x(0)\sin(w_o t) + m_y(0)\cos(w_o t)\right)$$

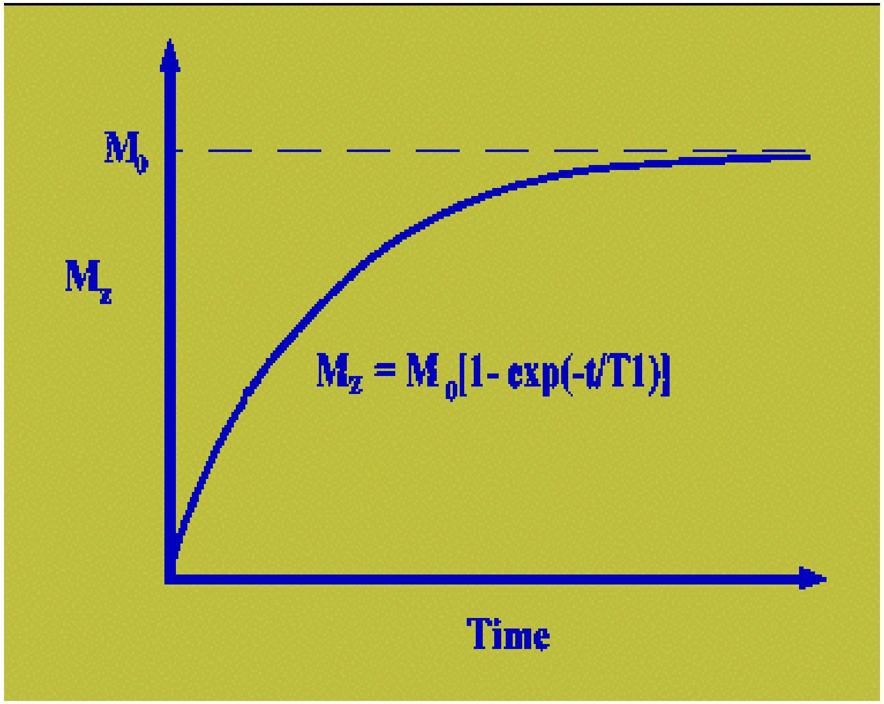
$$m_z(t) = m_z(0)e^{-t/T_1} + m_{eq}\left(1 - e^{-t/T_1}\right)$$

- Under **static magnetic field B0**, **during [t,t+Δt], aggregate magnetization M**:
  - Is precessing along +Z axis
- What happens to M(t) at equilibrium ($t \rightarrow \infty$) ?
  - Irrespective of tissue type (T1, T2) / initial conditions, aggregate magnetization aligned along B0

MRI

- Bloch Equation in a static magnetic field B0
  - Longitudinal (Z) relaxation (T1 dependent)
    - Mxy component not shown
    - Assuming mz(0) = 0





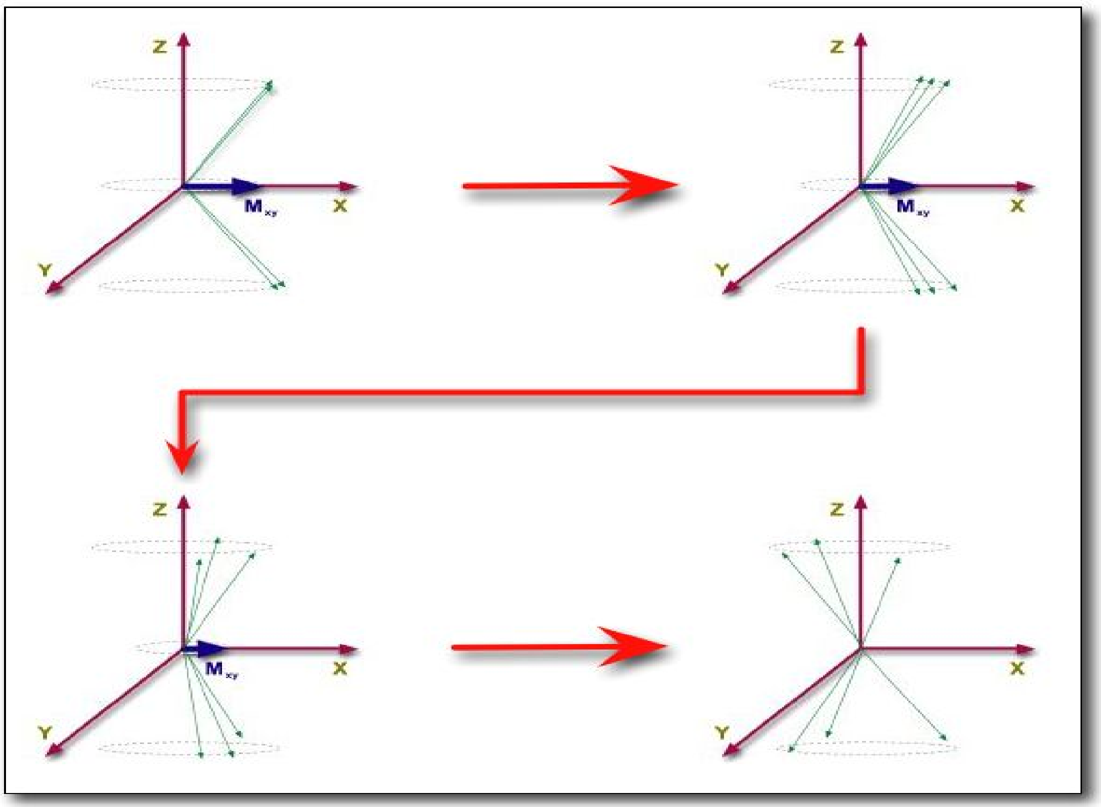
$$m_x(t) = e^{-t/T_2}\left(m_x(0)\cos(w_o t) - m_y(0)\sin(w_o t)\right)$$

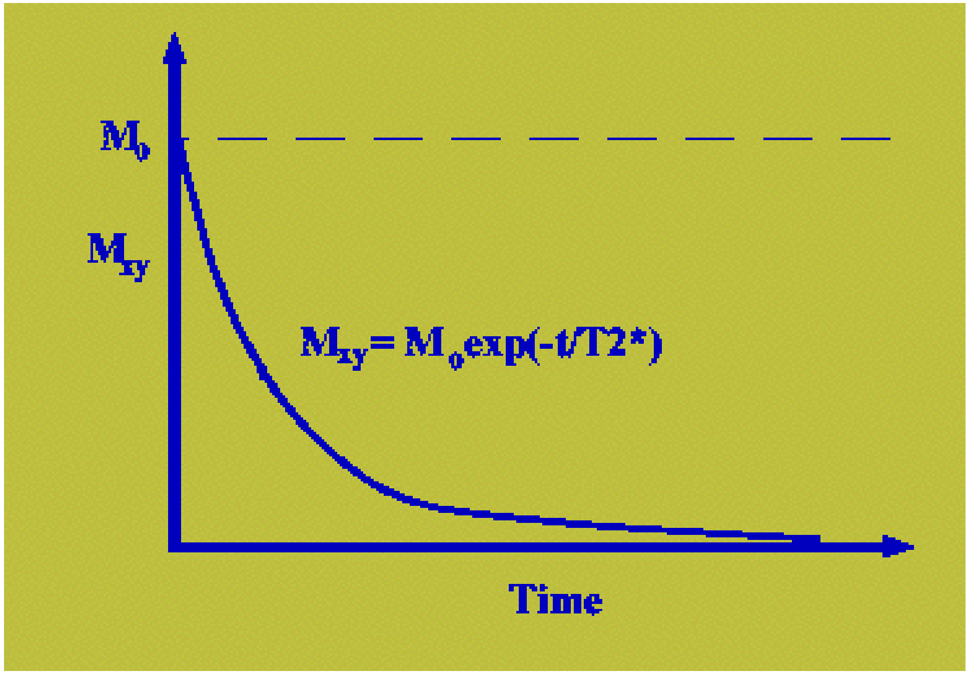
$$m_y(t) = e^{-t/T_2}\left(m_x(0)\sin(w_o t) + m_y(0)\cos(w_o t)\right)$$

$$m_z(t) = m_z(0)e^{-t/T_1} + m_{eq}\left(1 - e^{-t/T_1}\right)$$

MRI

- Bloch Equation in a static magnetic field B0
  - Transverse (XY) relaxation (T2 dependent)
    - Mz component not shown





$$m_x(t) = e^{-t/T_2}\left(m_x(0)\cos(w_o t) - m_y(0)\sin(w_o t)\right)$$

$$m_y(t) = e^{-t/T_2}\left(m_x(0)\sin(w_o t) + m_y(0)\cos(w_o t)\right)$$

$$m_z(t) = m_z(0)e^{-t/T_1} + m_{eq}\left(1 - e^{-t/T_1}\right)$$

MRI

- Bloch Equation
  - T1, T2 values at B0 = 1.5 Tesla

Tissue	T1 (ms)	T2 (ms)
Muscle	870	47
Liver	490	43
Kidney	650	58
Grey Matter	920	100
White Matter	790	92
Lung	830	80
CSF	2,400	160

MRI

- Contrast in T1-weighted MRI (Saturation Recovery)
- Consider two tissues
  - For tissue type A, let T1 value be T1A
  - For tissue type B, let T1 value be T1B
- Start from a state with NO magnetic field
  - Implying NO magnetization, mx(0) = my(0) = mz(0) = 0
  - State of NO magnetization = **Saturation**
- Place each tissue, one by one, in magnetic field B0
- What is M(t) ?

$$m_x(t) = e^{-t/T_2}\left(m_x(0)\cos(w_o t) - m_y(0)\sin(w_o t)\right)$$

$$m_y(t) = e^{-t/T_2}\left(m_x(0)\sin(w_o t) + m_y(0)\cos(w_o t)\right)$$

$$m_z(t) = m_z(0)e^{-t/T_1} + m_{eq}\left(1 - e^{-t/T_1}\right)$$

MRI

- Contrast in T1-weighted MRI (Saturation Recovery)
- Consider two tissues
  - For tissue type A, let T1 value be T1A
  - For tissue type B, let T1 value be T1B
- Start from a state with NO magnetic field
  - Implying NO magnetization, mx(0) = my(0) = mz(0) = 0
  - State of NO magnetization = **Saturation**
- Place each tissue, one by one, in magnetic field B0
- After time t, measure mzA (t) and mzB (t)
- Question: What value of t will give maximum contrast between two tissues ?

MRI

- Contrast in T1-weighted MRI (Saturation Recovery)
- What value of t will give maximum contrast between two tissues ?
  - Solve

$$\arg\max_t m_{zA}(t) - m_{zB}(t)$$

$$= \arg\max_t m_{eq}(1 - e^{-t/T_{1A}}) - m_{eq}(1 - e^{-t/T_{1B}})$$

  - Take derivative, equate it to zero → optimal time t

$$t = \frac{\log(T_{1B}/T_{1A})}{1/T_{1A} - 1/T_{1B}}$$

MRI

- Contrast in T1-weighted MRI (Inversion Recovery)
- Consider two tissues
  - For tissue type A, let T1 value be T1A
  - For tissue type B, let T1 value be T1B
- Place each tissue, one by one, in magnetic field B0
- Wait to reach equilibrium state: mx=my=0; mz=meq
- In some way, flip / **invert** magnetization by 180 deg.
  - mx = my = 0 and mz = **−meq**
  - Consider this time t=0
- After time t, measure mzA (t) and mzB (t)

MRI

- Contrast in T1-weighted MRI (Inversion Recovery)
- What value of t will give maximum contrast between two tissues ?
- Solve

$$\arg\max_t m_{zA}(t) - m_{zB}(t)$$

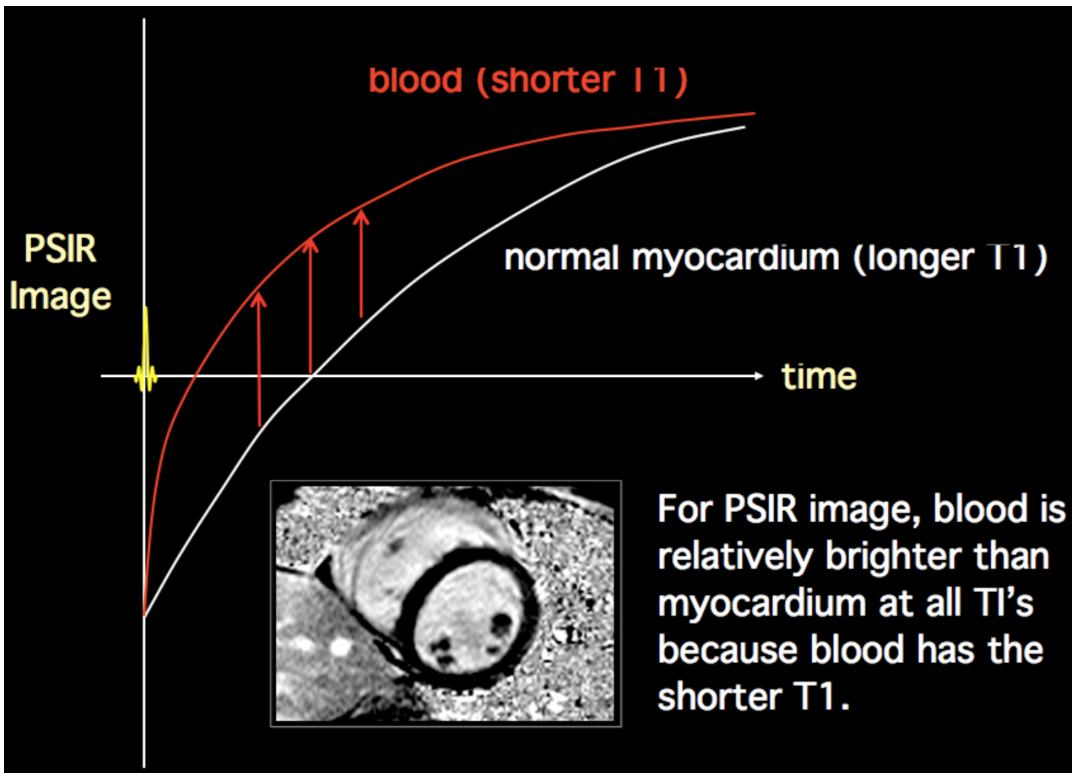
$$= \arg\max_t m_{eq}(1 - 2e^{-t/T_{1A}}) - m_{eq}(1 - 2e^{-t/T_{1B}})$$

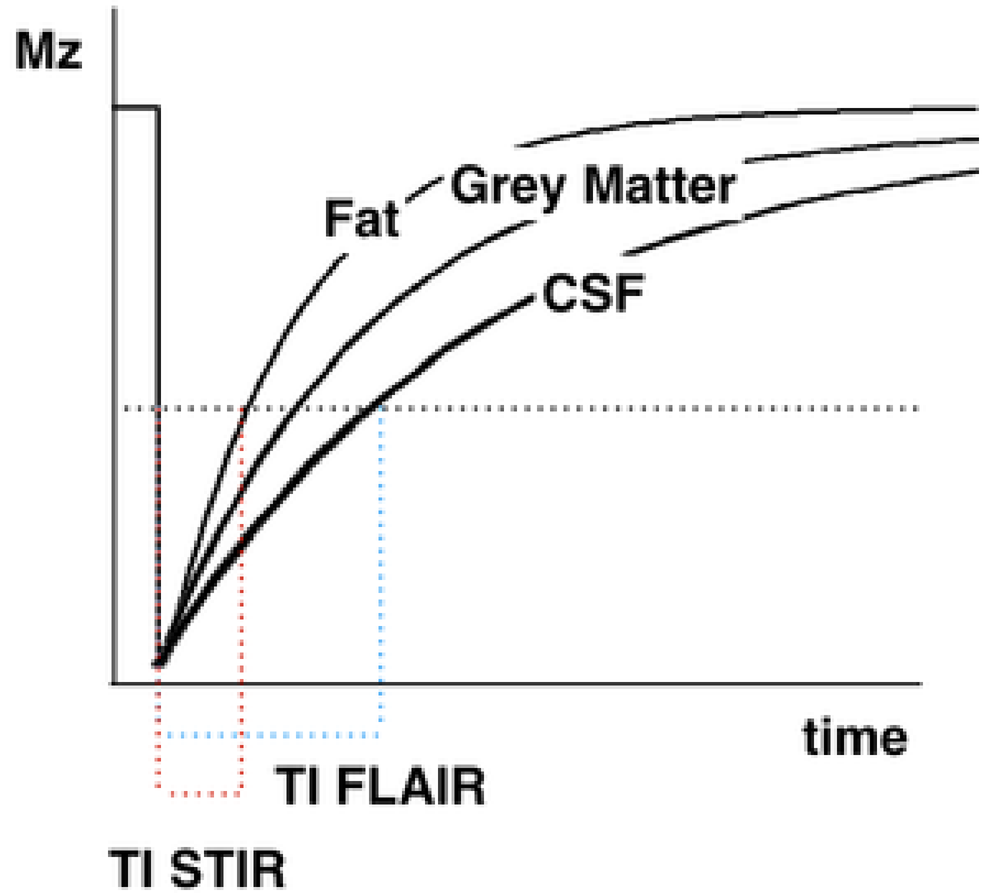
- Take derivative, equate it to zero → optimal time t

$$t = \frac{\log(T_{1B}/T_{1A})}{1/T_{1A} - 1/T_{1B}}$$

MRI

- Bloch Equation in a static magnetic field B0
  - Inversion Recovery
    - Optimal contrast



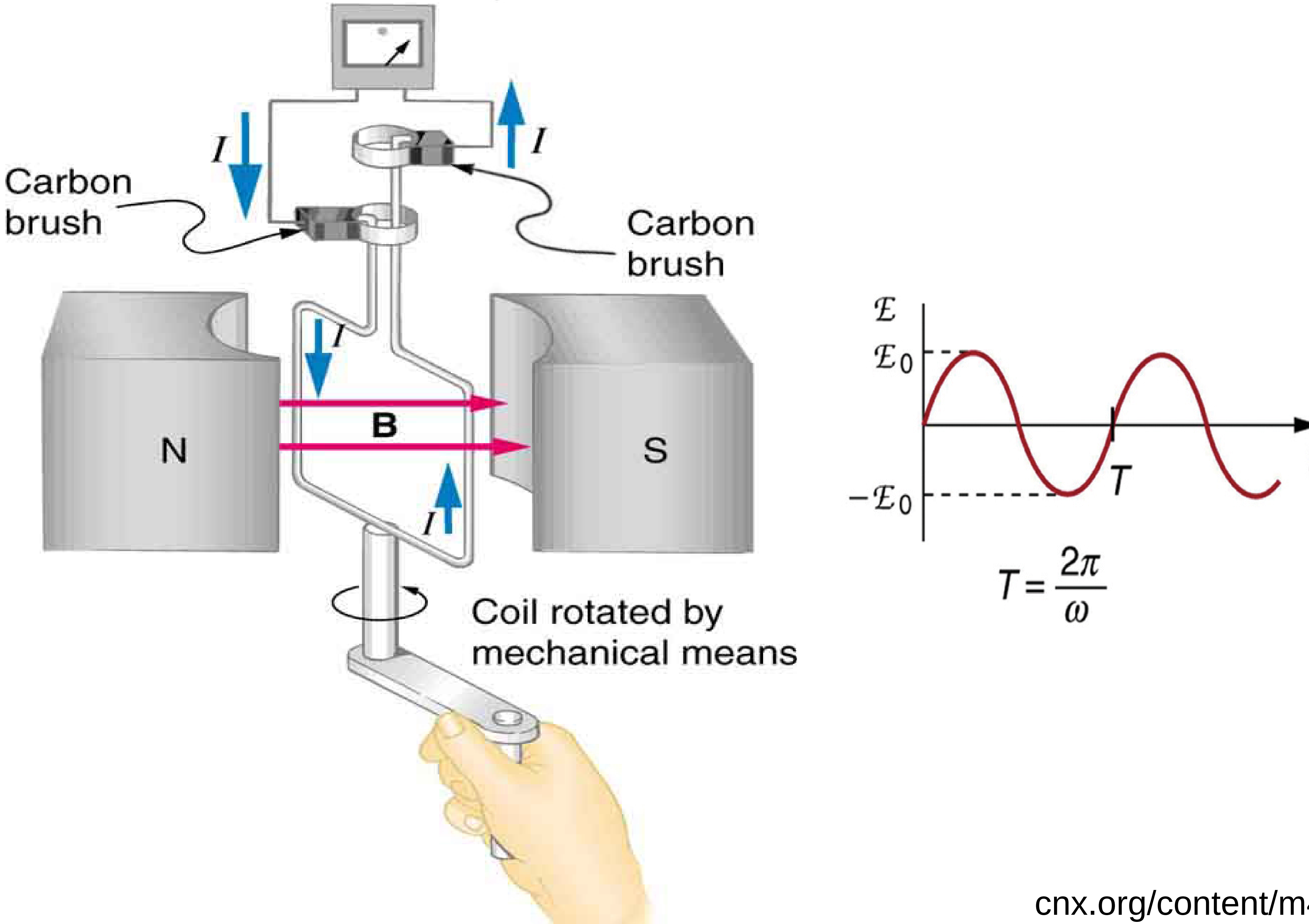


MRI

- Measuring signal in T1-weighted imaging
  - How to **measure** aggregate magnetization **mz(t)** that is a function of T1 ?
    - mx(t), my(t) aren’t functions of T1
  - Solution:
    - During [t,t+Δt], somehow **make vector mz(t) rotate in XY plane**
    - Faraday's Law:
      A changing magnetic field induces a current in a wire loop (orthogonal to XY plane)
    - Same principle as of electrical generator
    - Doesn't matter if magnet is moving or wire loop is moving

MRI

- Changing magnetic field induces current in wire loop



cnx.org/content/m42390

Michael Faraday

- Little formal education
- Electromagnetism, electrochemistry
  - Faraday's law of induction, Electrochemistry, Faraday effect, Faraday cage, Faraday constant, Faraday cup, Faraday's laws of electrolysis, Faraday paradox, Faraday rotator, Faraday-efficiency effect,
    Faraday wave, Faraday wheel, Lines of force
- "... there is no honour too great to pay to the memory of Faraday, one of the greatest scientific discoverers of all time" – Ernest Rutherford

MRI

- RF magnetic field B1(t) and resonance
- (1) Take a system precessing around B0 and
  (2) Apply an oscillating magnetic field in XY plane
- Static magnetic field: B0 := <0, 0, b0>
- Time-varying oscillating field in the XY plane:
  B1(t) := <b1 cos(wt), b1 sin(wt), 0>
  - Note: b1 = constant
- Net magnetic field:
  B(t) := B0 + B1(t) = <b1 cos(wt), b1 sin(wt), b0>

MRI

- RF magnetic field B1(t) and resonance
- **Assumption**: Time for which B1(t) is non-zero is very small compared to T1, T2
  - Ignore terms containing T1, T2 in Bloch equation

$$\frac{dM}{dt} = \gamma M \times B - \frac{\langle m_x, m_y, 0\rangle}{T_2} - \frac{\langle 0, 0, m_z - m_{eq}\rangle}{T_1}$$

where γ, T_1, T_2, m_{eq} are constants.

  - **RF "pulse"**
  - B(t) := B0 + B1(t) = <b1 cos(wt), b1 sin(wt), b0>
- Simplified Bloch equations

$$\frac{dm_x}{dt} = \gamma b_0 m_y - \gamma b_1 m_z \sin(wt)$$

$$\frac{dm_y}{dt} = -\gamma b_0 m_x + \gamma b_1 m_z \cos(wt)$$

$$\frac{dm_z}{dt} = \gamma b_1 m_x \sin(wt) - \gamma b_1 m_y \cos(wt)$$

MRI

- RF magnetic field B1(t) and resonance
- Introduce rotating coordinate frame <e1, e2> in XY

$$e_1(t) := \langle \cos(wt), \sin(wt) \rangle$$

$$e_2(t) := \langle -\sin(wt), \cos(wt) \rangle$$

  - Orthonormal basis
    - e1(t) orthogonal to e2(t)
    - || e1(t) || = 1
    - || e2(t) || = 1

MRI

- RF magnetic field B1(t) and resonance
- Introduce rotating coordinate frame <e1, e2> in XY

$$e_1(t) := \langle \cos(wt), \sin(wt) \rangle$$
$$e_2(t) := \langle -\sin(wt), \cos(wt) \rangle$$

- We want to write $\langle m_x(t), m_y(t) \rangle$ as $\langle u(t), v(t) \rangle_{e_1(t), e_2(t)}$
  - This implies: $\langle m_x(t), m_y(t) \rangle = u(t)e_1(t) + v(t)e_2(t)$
  - Equate X (and Y) components on LHS and RHS:

$$m_x(t) = u(t)\cos(wt) - v(t)\sin(wt)$$
$$m_y(t) = u(t)\sin(wt) + v(t)\cos(wt)$$

  - This gives: $u(t) := m_x(t)\cos(wt) + m_y(t)\sin(wt)$

$$v(t) := m_y(t)\cos(wt) - m_x(t)\sin(wt)$$

MRI

- RF magnetic field B1(t) and resonance
- Rewrite Bloch equations in the rotating frame
  - Start with
    $$u(t) := m_x(t)\cos(wt) + m_y(t)\sin(wt)$$
    $$v(t) := m_y(t)\cos(wt) - m_x(t)\sin(wt)$$
  - Differentiate LHS and RHS w.r.t. 't'
  - Substitute derivatives of mx(t), my(t) by their expressions in Bloch equations
  - Simplify
    $$\frac{dm_x}{dt} = \gamma b_0 m_y - \gamma b_1 m_z \sin(wt)$$
    $$\frac{dm_y}{dt} = -\gamma b_0 m_x + \gamma b_1 m_z \cos(wt)$$
    $$\frac{dm_z}{dt} = \gamma b_1 m_x \sin(wt) - \gamma b_1 m_y \cos(wt)$$
    to get →

MRI

- RF magnetic field B1(t) and resonance
- Rewrite Bloch equations in the rotating frame

$$\frac{du}{dt} = (\gamma b_0 + w)v$$
$$\frac{dv}{dt} = -(\gamma b_0 + w)u + \gamma b_1 m_z$$
$$\frac{dm_z}{dt} = -\gamma b_1 v$$

- Resonance
  - Choose frequency of RF pulse s.t. w := w0 = −γ b0 = Larmor frequency of H nuclei in B0

$$\frac{du}{dt} = 0$$
$$\frac{dv}{dt} = \gamma b_1 m_z$$
$$\frac{dm_z}{dt} = -\gamma b_1 v$$

MRI

- RF magnetic field B1(t) and resonance
- At resonance:

$$\frac{du}{dt} = 0$$
$$\frac{dv}{dt} = \gamma b_1 m_z$$
$$\frac{dm_z}{dt} = -\gamma b_1 v$$

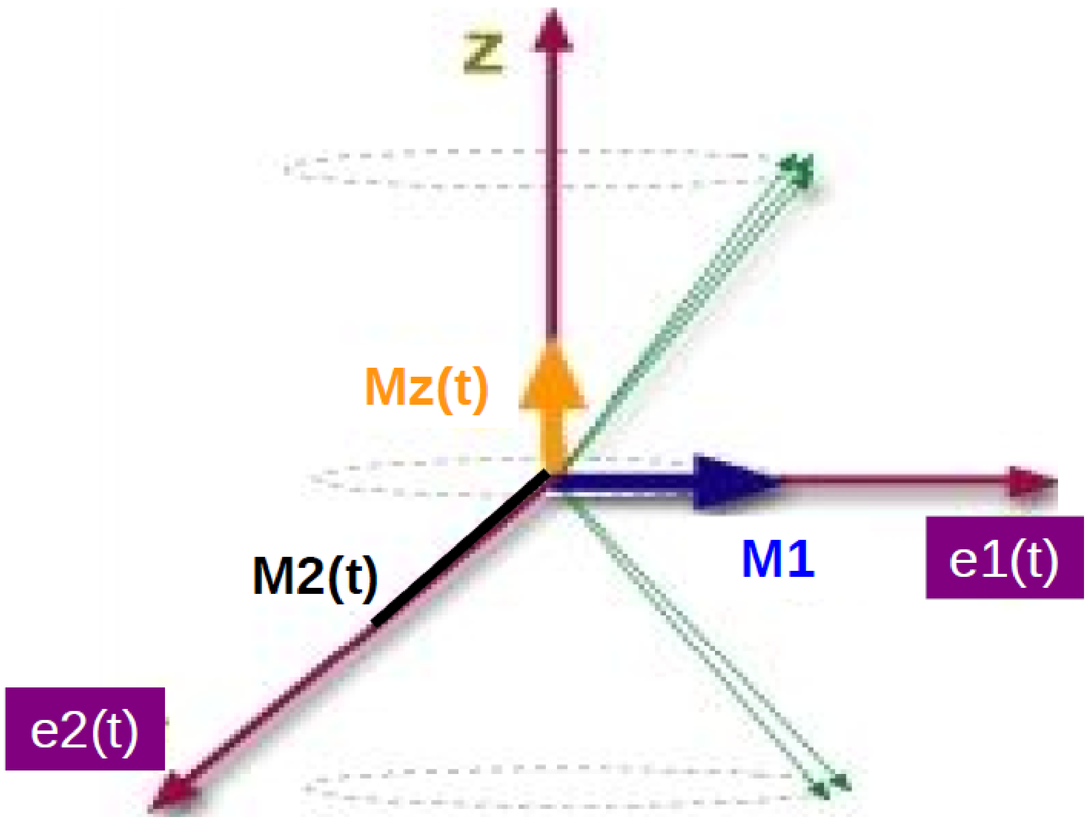
- Solution at resonance:

$$u(t) = u(0)$$
$$v(t) = v(0)\cos(-\gamma b_1 t) - m_z(0)\sin(-\gamma b_1 t)$$
$$m_z(t) = m_z(0)\cos(-\gamma b_1 t) + v(0)\sin(-\gamma b_1 t)$$

  - Aggregate magnetization precesses around e1(t) axis
  - e1(t) axis rotates at Larmor frequency w

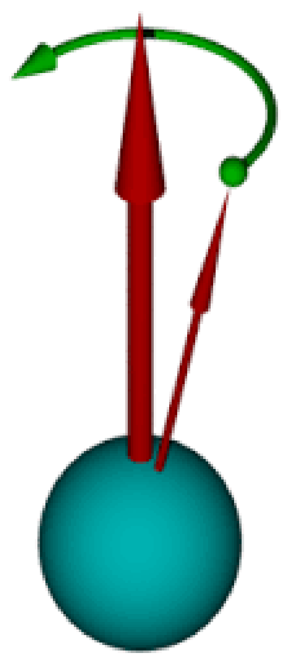
MRI

- RF magnetic field B1(t) …
  - Resonance: RF-field frequency = Larmor frequency
    - Rotating frame < e1(t), e2(t), Z > rotates at Larmor frequency
    - Aggregate magnetization precesses around e1(t)
    - Along e1(t)
      - Magnitude M1 = constant, time independent
    - Along e2(t), Z
      - Magnitudes M2(t), Mz(t) needn't be constant



MRI

- Joseph Larmor (Physicist, Mathematician)
  - Larmor precession
    - Magnetic field exerts torque on magnetic moment
    - $\vec{\Gamma} = \vec{\mu} \times \vec{B} = \gamma \vec{J} \times \vec{B}$
    - Angular momentum vector J precesses about external field B's axis with an angular frequency = Larmor frequency = $\omega = -\gamma B$
      - γ = gyromagnetic ratio



MRI

- B0 strength
  - https://www.youtube.com/watch?v=7g5UVrOt2CI
  - https://www.youtube.com/watch?v=byRIwDk21sw
  - https://www.youtube.com/watch?v=6BBx8BwLhqg

MRI

- **RF Pulse: 90-degree Pulse**
- (1) Take a system precessing around B0
  (2) Apply B1(t) in XY plane for a specific time τ1
- Assume initial state = equilibrium state
  - mx(0) = my(0) = 0; mz(0) = meq
- Apply B1(t) := <b1 cos(w0 t), b1 sin(w0 t), 0> for a carefully chosen time

  $\tau_1 := \frac{1}{\gamma b_1}\frac{\pi}{2}$ where $\tau_1 \ll T_1, T_2$
- Magnetization at time τ1 ?

$$u(t) = u(0)$$
$$v(t) = v(0)\cos(-\gamma b_1 t) - m_z(0)\sin(-\gamma b_1 t)$$
$$m_z(t) = m_z(0)\cos(-\gamma b_1 t) + v(0)\sin(-\gamma b_1 t)$$

$$u(\tau_1) = u(0) = 0$$
$$v(\tau_1) = m_z(0) = m_{eq}$$
$$m_z(\tau_1) = -v(0) = 0$$

MRI

- RF Pulse: 90-degree Pulse
- Initial state = equilibrium
- Apply B1(t) for τ1

$$u(\tau_1) = u(0) = 0$$
$$v(\tau_1) = m_z(0) = m_{eq}$$
$$m_z(\tau_1) = -v(0) = 0$$

- In the rotating coordinate frame (u(t), v(t), z), orientation of aggregate magnetization changes from <0, 0, meq> → <0, meq, 0>
  - Initial condition in rotated frame is: u(0) = v(0) = 0
    - Because mx(0) = my(0) = 0
  - Magnetization flipped by 90 degrees

MRI

- RF Pulse: 90-degree Pulse
- Revisit T1-w acquisition via Saturation Recovery

(1) Initial state : mx (0) = my (0) = mz (0) = 0

(2) Let magnetization recover towards equilibrium

 – At time t, aggregate magnetization is:
 mx (t) = my (t) = 0 and mz (t) = meq (1 − exp(−t/T1))

(3) Apply 90-deg pulse. At time t + τ1:

$$u(t+\tau_1) = u(t) = 0$$

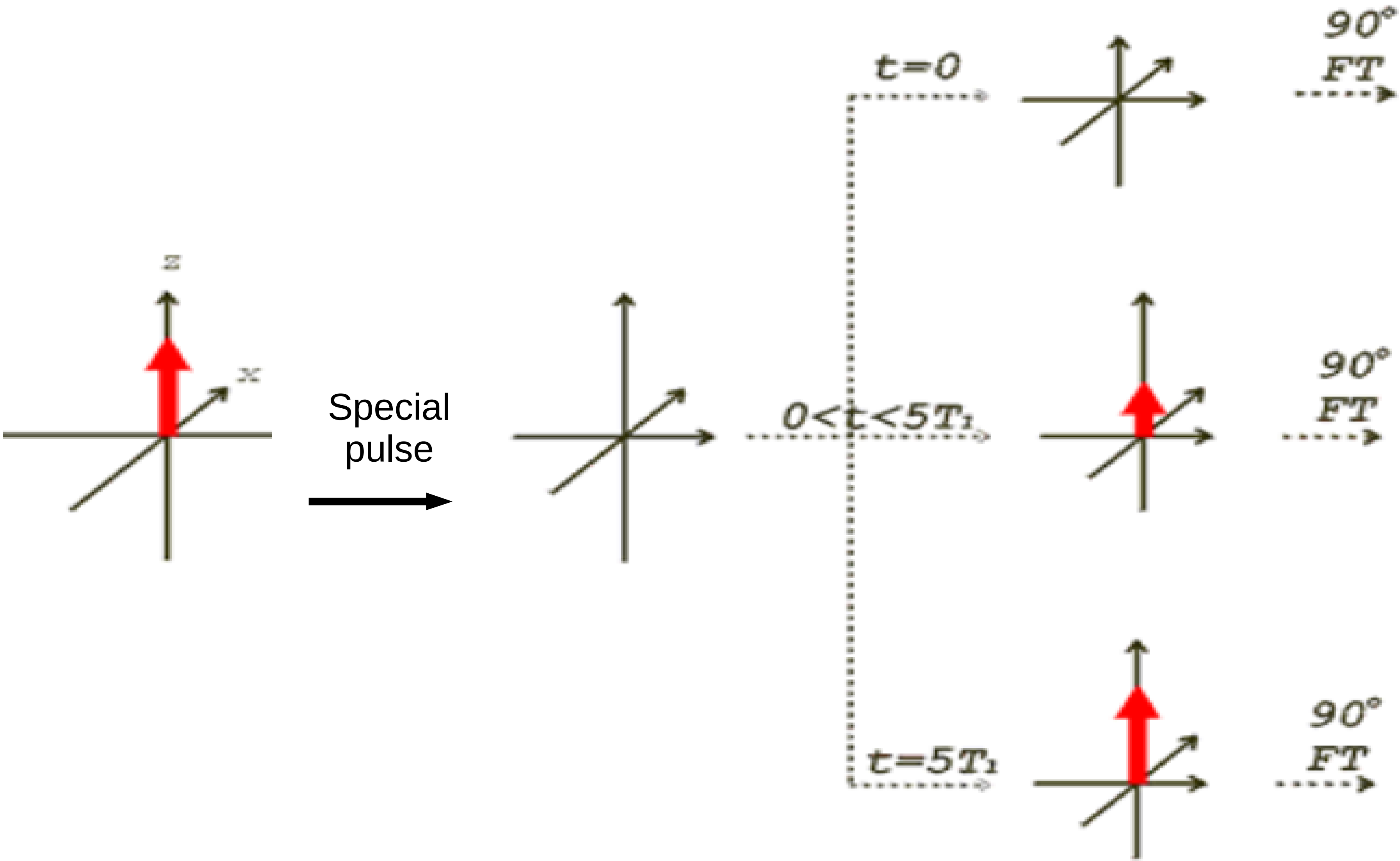
$$v(t+\tau_1) = m_z(t) = m_{eq}(1 - e^{-t/T_1})$$

$$m_z(t+\tau_1) = -v(t) = 0$$

(4) Measure current prop. to meq (1 − exp(−t/T1))

MRI

- Measuring the signal
  - Saturation Recovery



MRI

- **RF Pulse: 180-degree Pulse**
- (1) Take a system precessing around B0
  (2) Apply B1(t) in XY plane for a specific time τ2
- Assume initial state = equilibrium state
  - mx (0) = my (0) = 0 and mz (0) = meq
- Apply B1 (t) := <b1 cos(wt), b1 sin(wt), 0>
  for a carefully chosen time

$$\tau_2 := \frac{1}{\gamma b_1}\pi$$

- Magnetization at time τ2 ?

$$u(t) = u(0)$$
$$v(t) = v(0)\cos(-\gamma b_1 t) - m_z(0)\sin(-\gamma b_1 t)$$
$$m_z(t) = m_z(0)\cos(-\gamma b_1 t) + v(0)\sin(-\gamma b_1 t)$$

$$u(\tau_2) = u(0) = 0$$
$$v(\tau_2) = -v(0) = 0$$
$$m_z(\tau_2) = -m_z(0) = -m_{eq}$$

MRI

- RF Pulse: 180-degree Pulse
- Initial state = equilibrium
- Apply B1(t) for τ2

$$u(\tau_1) = u(0) = 0$$
$$v(\tau_1) = m_z(0) = m_{eq}$$
$$m_z(\tau_1) = -v(0) = 0$$

- In the rotating coordinate frame (u(t), v(t), z), orientation of aggregate magnetization changes from <0, 0, meq> → <0, 0, -meq>
  - <u,v> = 0 → <mx,my> = 0
  - Magnetization flipped by 180 degrees

MRI

- RF Pulse: 180-degree Pulse
- Revisit T1-w acquisition via Inversion Recovery

(1) Initial state : mx (0) = my (0) = 0 ; mz (0) = meq

(2) Apply 180-deg pulse

– At time τ2

$$u(\tau_2) = u(t) = 0$$
$$v(\tau_2) = -v(t) = 0$$
$$m_z(\tau_2) = -m_z(t) = -m_{eq}$$

– This implies mx (τ2) = my (τ2) = 0

(3) Let magnetization recover towards equilibrium

– At time t' = τ2 + t, aggregate magnetization is:
mx (t') = my (t') = 0 ;   mz (t') = meq (1 − 2exp(−t / T1))

MRI

- RF Pulse: 180-degree Pulse
- Revisit T1-w acquisition via Inversion Recovery

(1) Initial state : mx (0) = my (0) = 0 ; mz (0) = meq

(2) Apply 180-deg pulse

(3) Let magnetization recover towards equilibrium

(4) Apply 90-deg pulse

– At time τ2 + t + τ1

$$u(\tau_2 + t + \tau_1) = u(\tau_2 + t) = 0$$

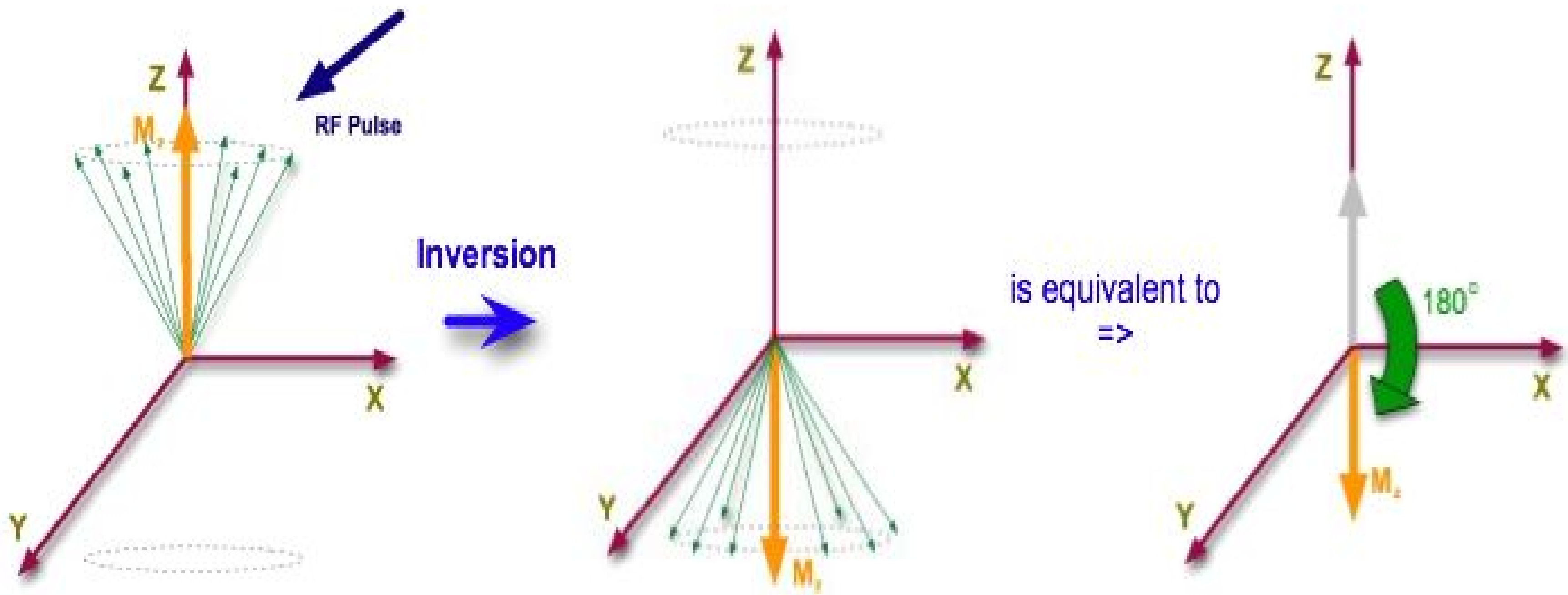
$$v(\tau_2 + t + \tau_1) = m_z(\tau_2 + t) = m_{eq}(1 - 2e^{-t/T_1})$$

$$m_z(\tau_2 + t + \tau_1) = -v(\tau_2 + t) = 0$$

(5) Measure current

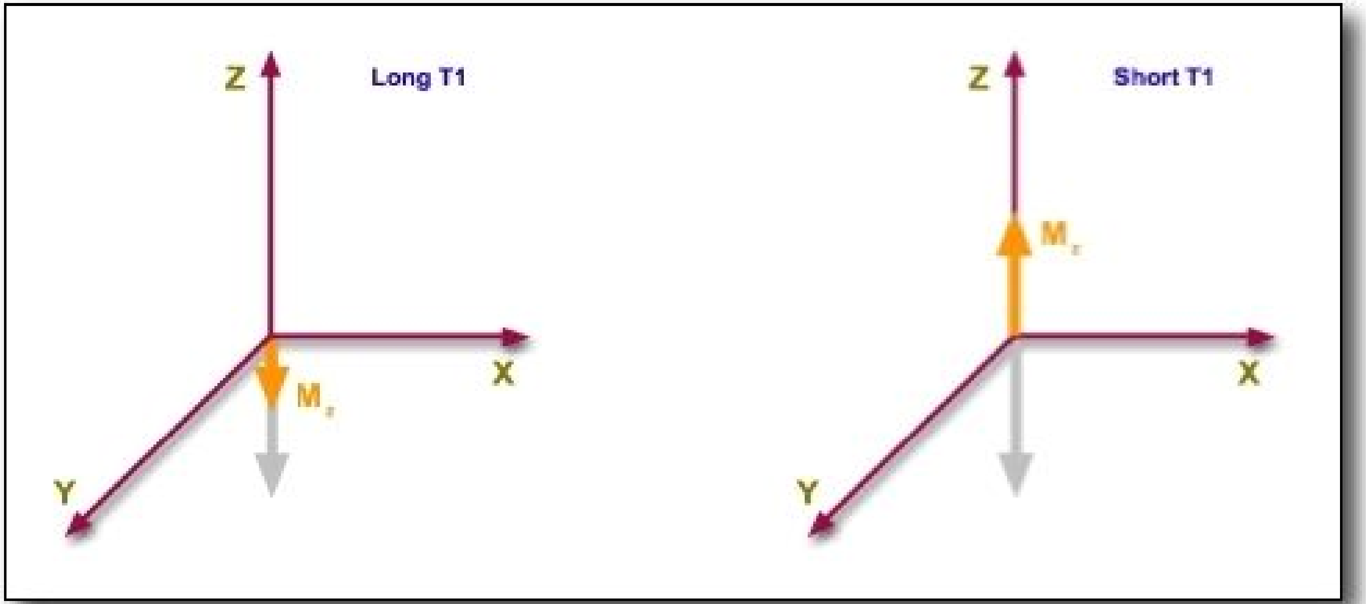
MRI

- Measuring the signal
  - Inversion Recovery
    - Initial condition: At equilibrium
    - Step 1 (of 3): Apply the 180 degree pulse



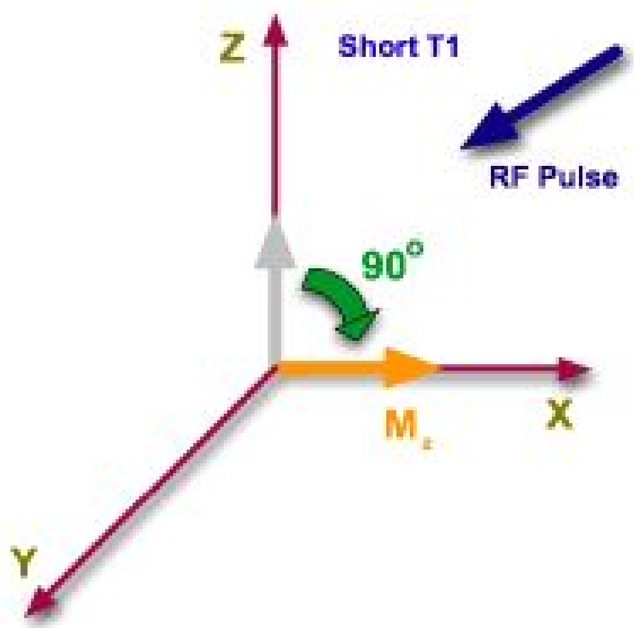
MRI

- Measuring the signal
  - Inversion Recovery
    - Initial condition: At equilibrium
    - Step 1 (of 3): Apply the 180-degree pulse
    - Step 2 (of 3): Wait for some time for T1 relaxation

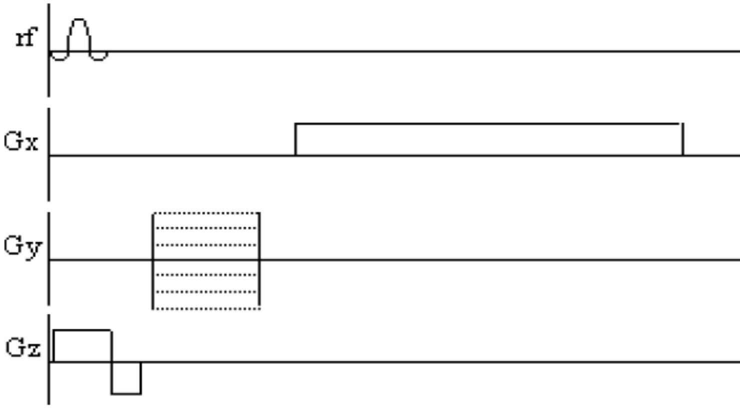


MRI

- Measuring the signal
  - Inversion Recovery
    - Initial condition: At equilibrium
    - Step 1 (of 3): Apply the 180-degree pulse
    - Step 2 (of 3): Wait for some time for T1 relaxation
    - Step 3 (of 3): Apply the 90-degree pulse and measure the signal



MRI



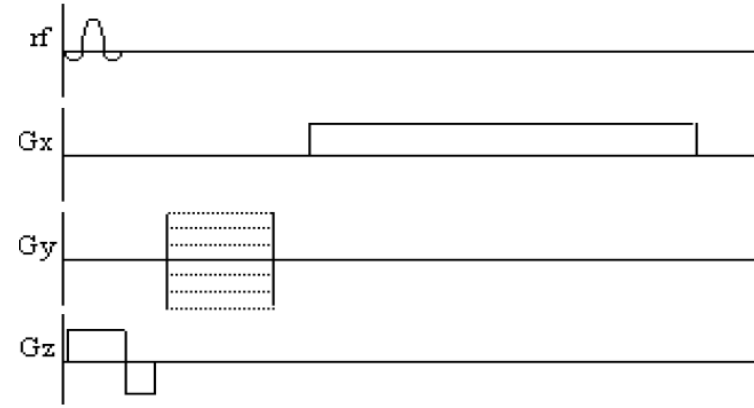
- **XY Imaging within Z slice**
  - Step 1: Static field B0 is on
    - Saturation / inversion recovery in progress
    - Mz (x,y,z, t) is a function of T1
  - Step 2: Slice Selection
    - Simultaneously, turn on RF field + Z-gradient field
    - Turn off both
    - Now, aggregate XY magnetization (rotating) non-zero in 1 slice
      - If slice around origin has thickness = t,

        Then

        | Mxy (x,y,z) | = 0 ; for |z| > t/2

        | Mxy (x,y,z) | ~ gamma Mz (x,y,z, t=0, T1) ; for |z| ≤ t/2

MRI



- XY Imaging within Z slice
  - Step 3 : Phase Encoding
    - Turn on Y-gradient field
      - Each Y column is subject to different **static** magnetic field **B0 + BG(y)**
      - Each Y-column's magnetization precesses at Y-proportional frequency
        - Recall: Bloch equation, Larmor frequency in **static** magnetic field

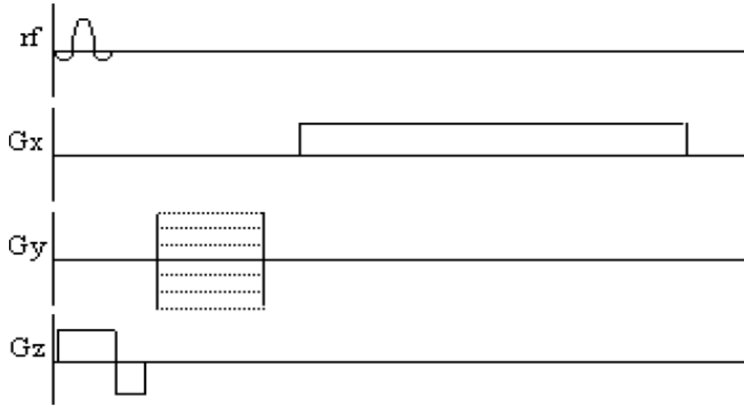
$$m_x(t) = e^{-t/T_2}\left(m_x(0)\cos(w_o t) - m_y(0)\sin(w_o t)\right)$$

$$m_y(t) = e^{-t/T_2}\left(m_x(0)\sin(w_o t) + m_y(0)\cos(w_o t)\right) \qquad w_0 = -\gamma b_0$$

$$m_z(t) = m_z(0)e^{-t/T_1} + m_{eq}\left(1 - e^{-t/T_1}\right)$$

    - Turn off Y-gradient field (after time t_2)
      - Each Y-column's magnetization has Y-proportional phase lag
    - Magnetization (rotating in XY plane)
      - Mxy (x,y) exp (i $G_Y y\ t_2$)

MRI



- XY Imaging within Z slice
  - Step 4 : Frequency Encoding
    - Turn on X-gradient field
      - Each X column is subject to different **static** magnetic field **B0 + BG(x)**
      - Each X-column's magnetization precesses at X-proportional frequency
        - Recall: Bloch equation, Larmor frequency in **static** magnetic field
  - At this moment :
    - Each XY element has an Y-proportional phase
    - Each XY element is precessing at a X-proportional frequency
    - Magnetization (rotating in XY plane)
      - Mxy (x,y) $\exp(i\, G_Y y\, t_2)$ $\exp(i\, G_X x\, t)$

MRI

- XY Imaging within Z slice
  - XY-magnetization from each slice element dXdY
    - Phase proportional to X
    - Frequency proportional to Y
    - Mxy (x,y) exp ($i\, G_Y t_2\, y + i\, G_X t\, x$)
  - Measure magnetization
    - Superposition of magnetization from each ΔXΔY element
      - Sum over XY plane **Mxy (x,y) exp ($i\, G_Y t_2\, y + i\, G_X t\, x$) $\Delta x\, \Delta y$** where

        Mxy (x,y) = gamma Mz (t=0,x,y,z,T1) → some function of T1
        - Via saturation recovery or inversion recovery

MRI

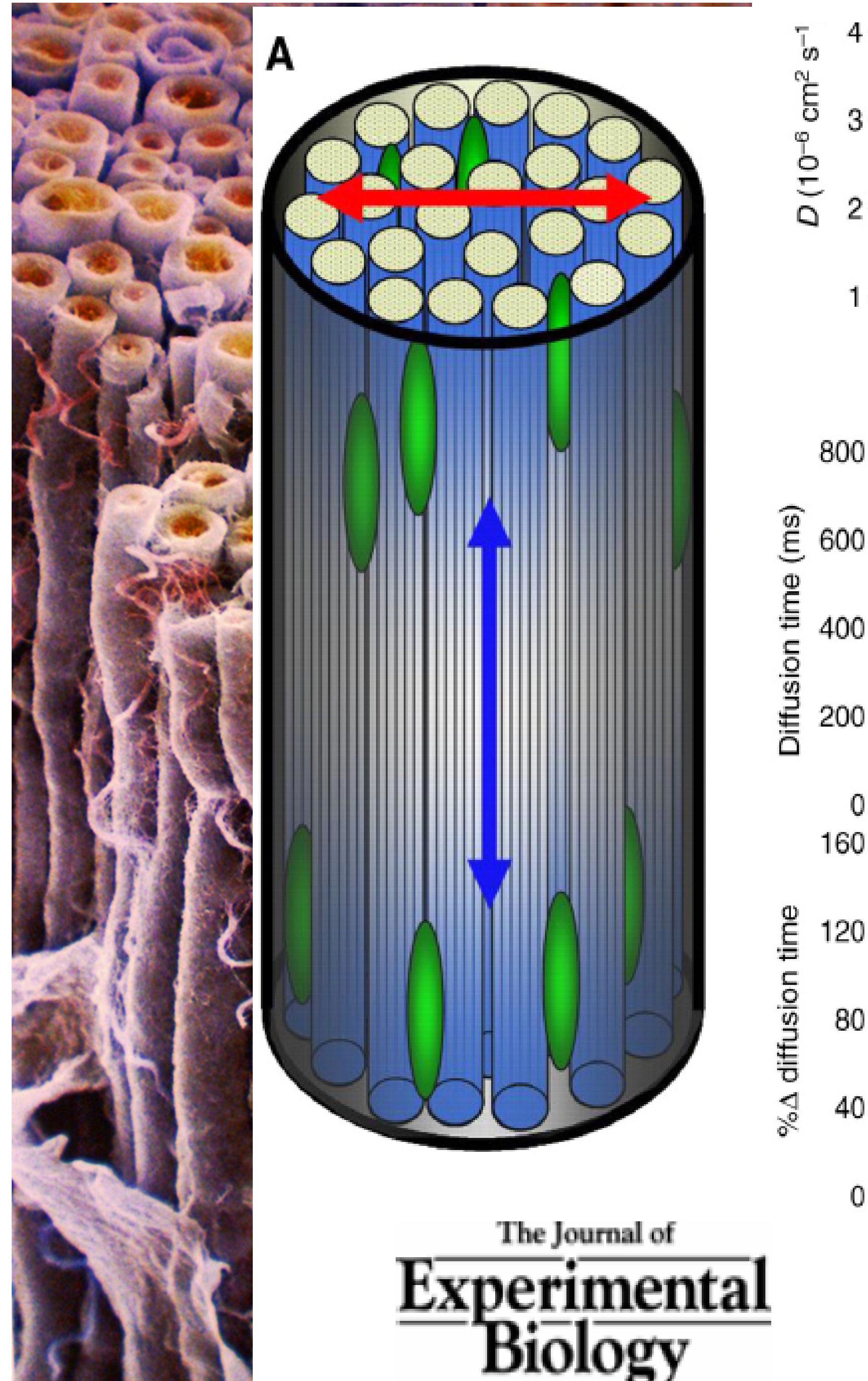
- XY Imaging within Z slice
  - Discrete Fourier transform of magnetization in XY plane
    - **FM (g_2t_2, g_1t)**
    - Note
      - In our definition of F(.), sign in exponential was negative
      - This is purely a matter of convention
    - What is more important is that this is an **integral over space**
  - For each value of $\mathbf{g_2t_2}$, measure **FM** for varying $\mathbf{t}$
  - Each value of $\mathbf{g_2t_2}$ gives 1 line of 2D DFT
  - Repeat sequence, with different $\mathbf{g_2}$ values,
    to fill Fourier coefficients in entire 2D plane

MRI

- XY Imaging within Z slice
  - For complex magnetization signal M(x,y)
    - Re-Im parts = X-Y components of (rotating) magnetization
    - Magnitude of M(x,y) is more important, for typical applications

Diffusion MRI

- Fibrous tissue
  - Muscle (heart), nerve (brain)
- Diffusion in fibrous tissue
  - Extracellular diffusion
  - Water (green)
  - More barriers to diffusion across axis of fibers (red →)
  - Less barriers to diffusion along axis of fibers (blue →)
  - More diffusion along axis of fibers (blue)



Diffusion MRI

- At each location, measures diffusion of water along chosen directions
  - Directions chosen in 3D
  - Number of directions between 10-100, typically

Diffusion MRI

- Early diffusion imaging was direction **independent**
- Mathematical model
  - Stejskal-Tanner equation (1965)

– Captures the diffusion-based signal at a particular location

$$S = S_0 \exp(-\gamma^2 G^2 \delta^2 (\Delta - \delta/3) D)$$

where

– $\gamma =$ gyromagnetic ratio

– $G =$ magnitude/strength of the diffusion gradient pulse

– $\delta =$ duration of the pulse

– $\Delta =$ time between two pulses

– $S_0 =$ signal with $G = 0$

Diffusion MRI

– Captures the diffusion-based signal at a particular location

$$S = S_0 \exp(-\gamma^2 G^2 \delta^2 (\Delta - \delta/3) D)$$

where

– $\gamma =$ gyromagnetic ratio

– $G =$ magnitude/strength of the diffusion gradient pulse

– $\delta =$ duration of the pulse

– $\Delta =$ time between two pulses

– $S_0 =$ signal with $G = 0$

Because $\gamma, G, \delta, \Delta$ are all scanner-related parameters, the equation can be simplified to

$$S = S_0 \exp(-b\text{ADC})$$

where

– $b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3)$

– ADC = apparent diffusion coefficient

– S_0 is the signal with $b = 0$

Diffusion MRI

- Mathematical model
  - Diffusion-tensor model (1994)
    - Tissue structure is fibrous
    - Diffusion is direction **dependent**
  - Key idea in diffusion-tensor imaging
    - Measure diffusion along multiple directions
    - Summarize direction-dependent diffusion using a tensor

$$S(g) = S_0 \exp(-bg^T D g)$$

where

– g is a direction vector, i.e., 3×1 unit vector.

– D is the diffusion tensor, i.e., 3×3 symmetric positive definite matrix

– b is a hardware-dependent constant

– S_0 is the signal with $b = 0$

Diffusion MRI

- Mathematical model
  - Consider a simpler (2D) case
    - D is 2x2
    - Given :
      Diffusion along X axis is twice as much as diffusion along Y axis
    - How can we represent this in matrix D ?
      - D (1,1) = 2
      - D (2,2) = 1
      - D (1,2) = D (2,1) = 0

Diffusion MRI

- Mathematical model
  - Does it work ?
    - If g = [1 0]' (X axis), then g' D g = 2
    - If g = [0 1]' (Y axis), then g' D g = 1
    - If g = [1/sqrt(2) 1/sqrt(2)]', then g' D g = 1.5
      - Direction g is between X and Y axes
      - Diffusion for g is between diffusion for X and diffusion for Y
    - If g = (- h), then g' D g = h' D h
      - Diffusion along positive axis is same as diffusion along negative axis

Diffusion MRI

- Mathematical model
  - Choice of XY axes is arbitrary
  - Consider rotated A_1A_2 axes
    - a1 := [cos(theta) sin(theta)]'
    - a2 := [-sin(theta) cos(theta)]'
  - Consider:
    Diffusion along a1 is twice as much as diffusion along a2
  - How can we represent this in a matrix E ?
    - E := [a1 , a2] D [a1 , a2]'
    - a1' g and a2' g = coordinates of g_{XY} in A_1A_2

Diffusion MRI

- Mathematical model
  - Does it work ?
    - If g = a1, then g' E g = 2
    - If g = a2, then g' E g = 1
    - If g = (a1 + a2) / sqrt(2), then g' E g = 1.5
    - If g = (- h), then g' E g = h' E h

Diffusion MRI

- Mathematical model
  - What is the structure of E = [a1 , a2] D [a1 , a2]' ?
    - D is a diagonal (implies symmetric)
      - Positive (> 0) diagonal elements
    - [a1 , a2] is an orthogonal matrix

  - What is an orthogonal matrix ?
    - Definition : A' A = A A' = I
    - Properties
      - Columns of A are unit vectors orthogonal to each other
        - Because A' A = I
      - Rows of A are unit vectors orthogonal to each other
        - Because A A' = I

Diffusion MRI

- Mathematical model
  - What are the properties of E ?
    - E is square (2x2)
    - E is symmetric
      - E' = ([a1 , a2] D [a1 , a2]')' = [a1 , a2] D [a1 , a2]' = E
    - E is positive definite
  - What is a positive definite matrix ?
    - A is square symmetric such that v' A v > 0, for any vector v
    - Show that v' E v > 0

Diffusion MRI

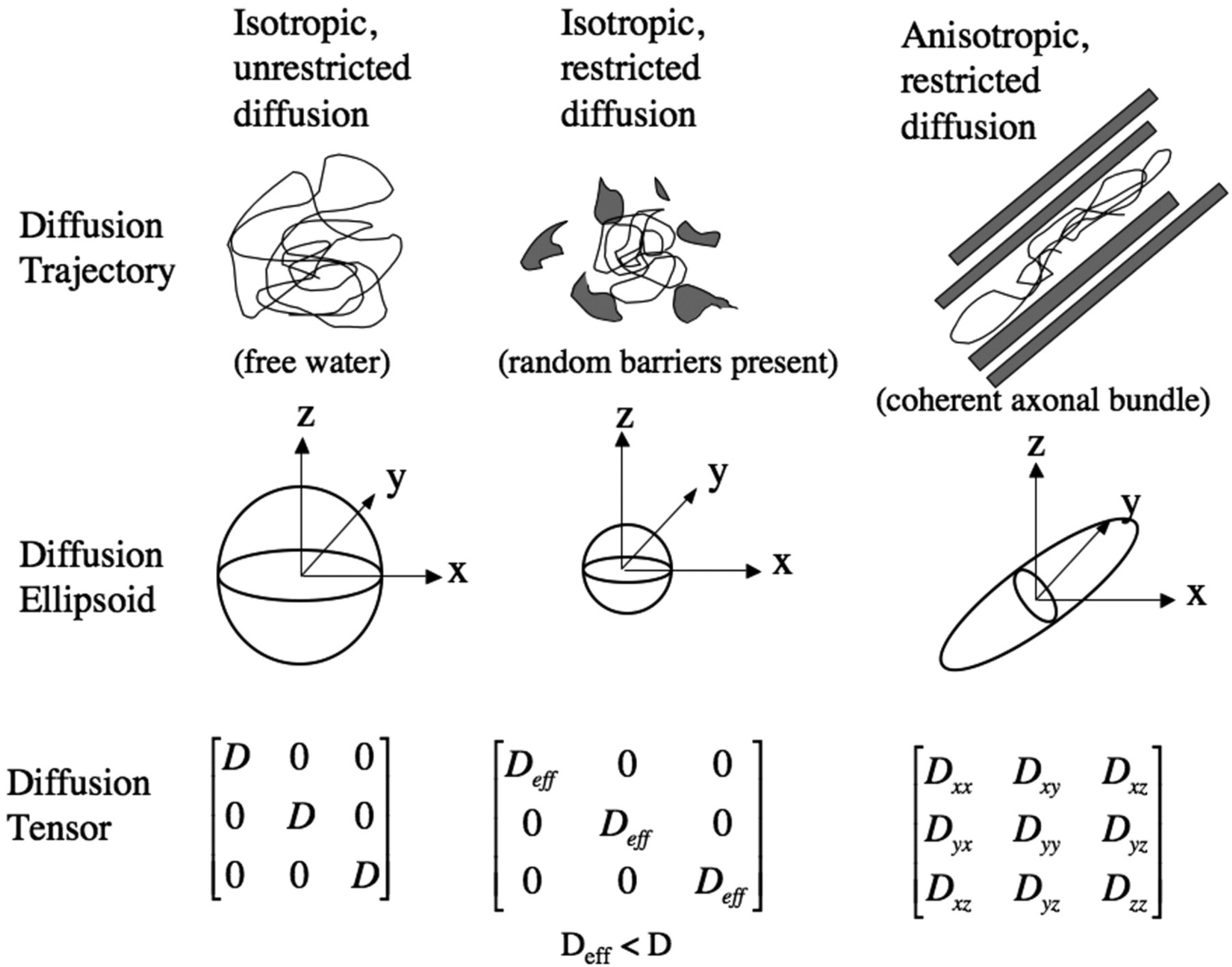
- Mathematical model
  - E is a symmetric positive-definite (SPD) matrix
  - Given E, how do we diagonalize it ?
    - Singular-value decomposition or eigen decomposition
      - E = V D V'
      - D contains the singular values or eigen values
        - For a SPD matrix, these values are all positive
      - V is an orthogonal matrix whose columns are eigenvectors
  - Eigenvalue of E gives amount of diffusion along (corresponding) eigenvector of E
    - Eigenvector with largest eigenvalue called principal eigenvector

Diffusion MRI

- Eigen Decomposition
  - Eigenvector definition : A x = lambda x
  - Assume A is square with linearly-independent eigenvectors x
  - Then A can be decomposed as $A = Q\ D\ Q^{-1}$
    - Q = Square NxN marix; each column = eigen-vector
    - D = Diagonal matrix; each element = eigen-value
  - If A is symmetric, then
    - Eigenvectors are orthogonal
    - Q is orthogonal
      - Q inverse = Q transpose

Diffusion MRI

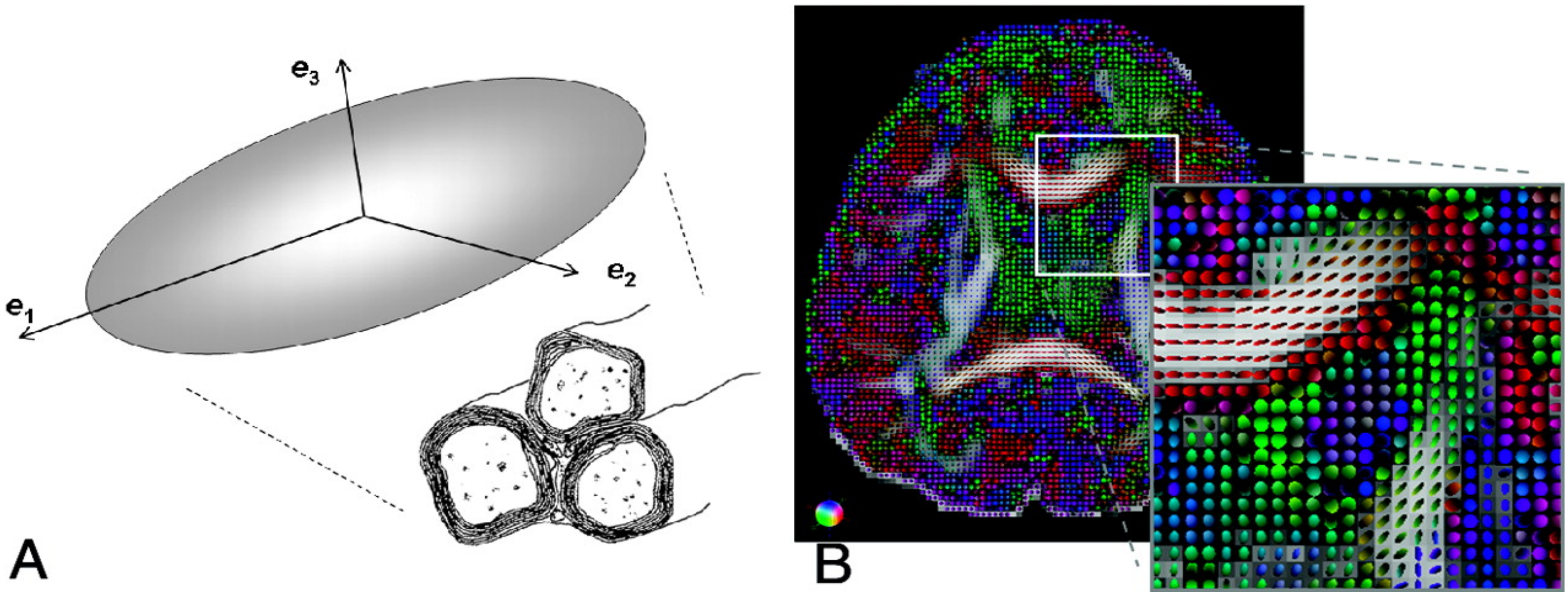
- Typical scenarios



Mukherjee P et al. Am. J. Neuroradiol. 2008;29:632-641

Diffusion MRI

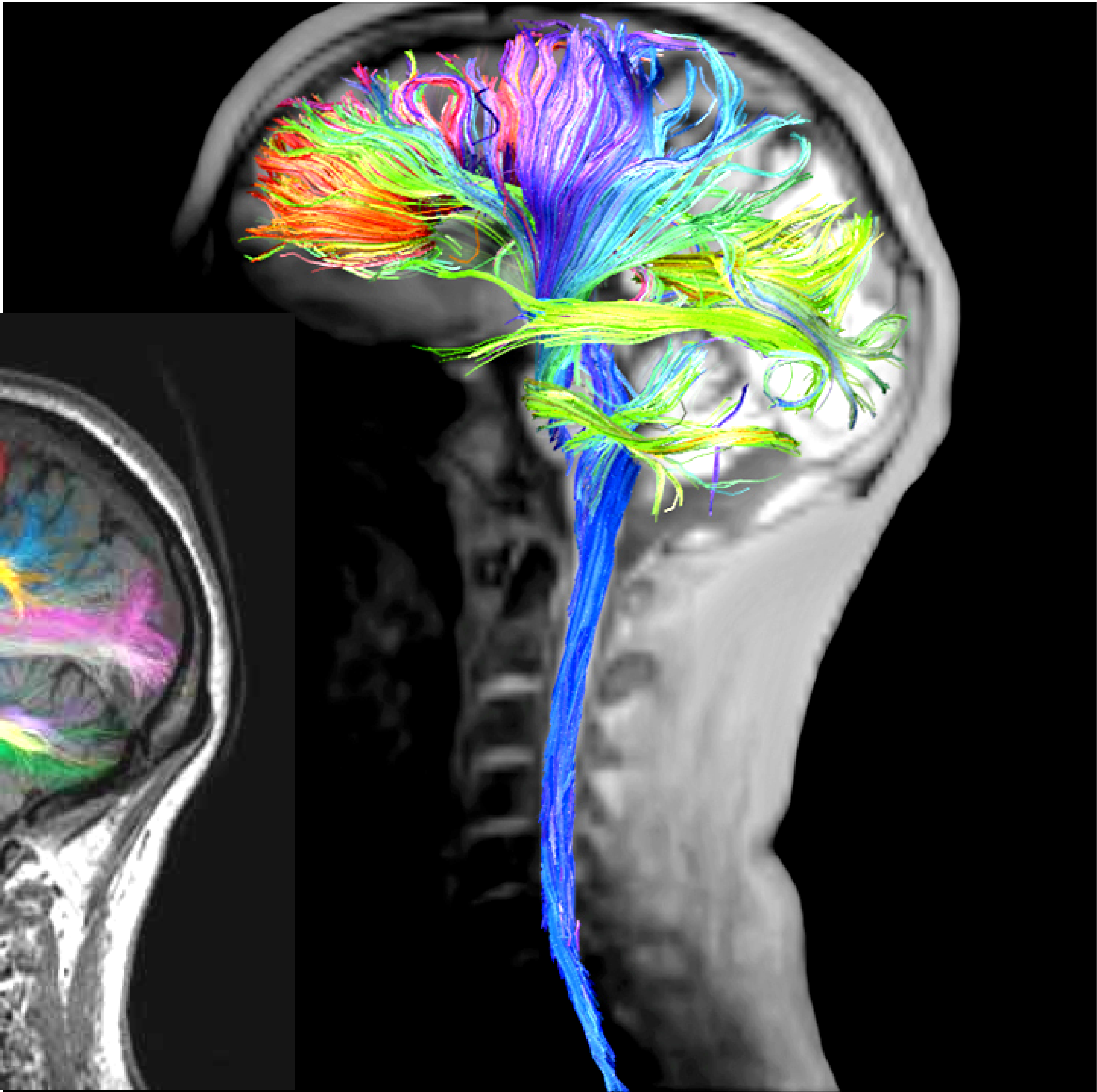
- Mathematical model
  - Visualize diffusion tensor as an ellipsoid
    - Ellipsoid's axes directions are vectors e1, e2, e3
    - Ellipsoid's axes lengths are diagonal elements of D
    - Ellipsoid's color based on direction of principal eigenvector

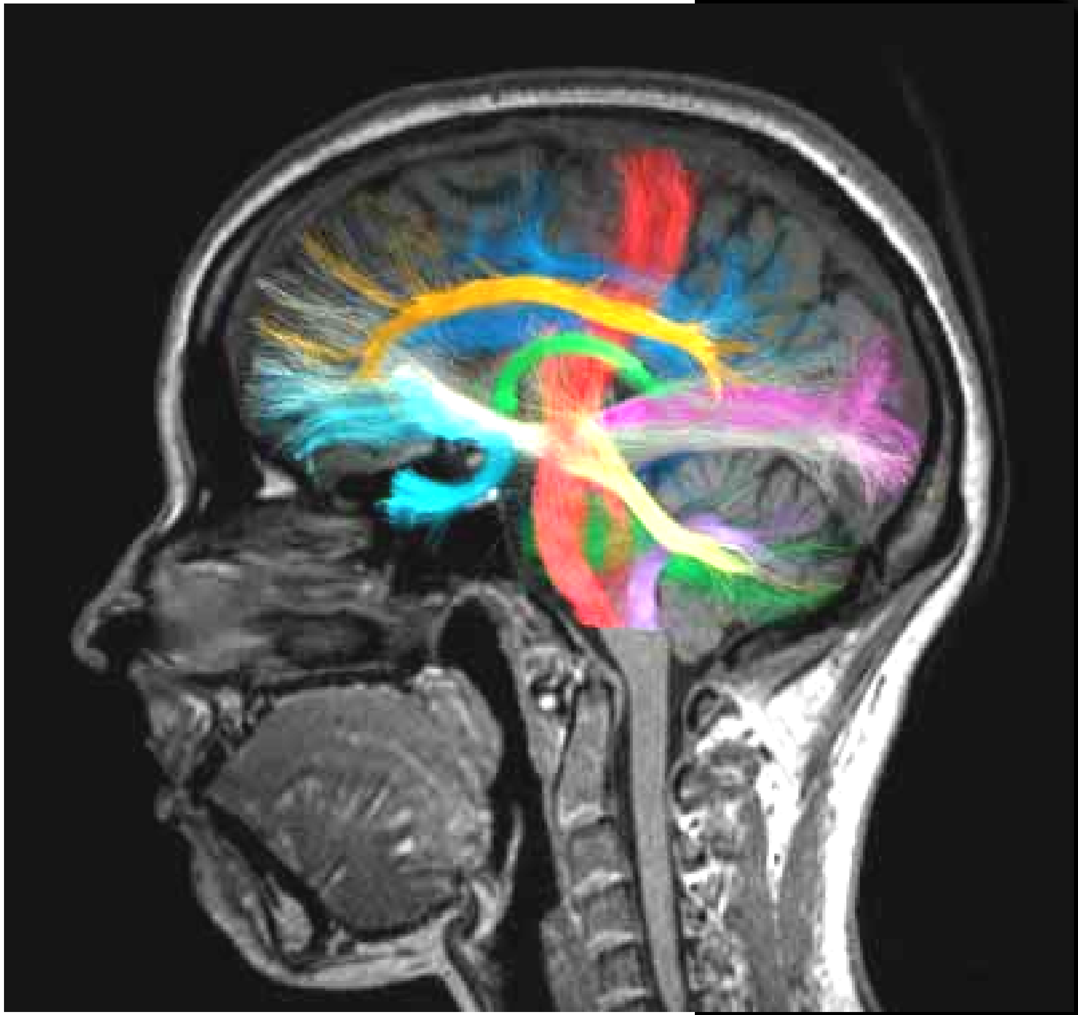


Mukherjee P et al. Am. J. Neuroradiol. 2008;29:632-641

Diffusion MRI

- Tractography
  - Streamlines in a vector field





Diffusion MRI

- Applications
  - Cerebral stroke
    - More sensitive than traditional MRI
  - Traumatic brain injury
  - Neurological disorders

Diffusion MRI

- Acquisition model
  - **S (g,p) = S0 exp (- b g' D(p) g)**
- How many measurements S (g,p) do we need to estimate E, for a fixed p ?
  - b is known
  - Unknowns
    - S0
    - 6 unknowns in D (3 x 3)
  - Need minimum 7 measurements
    - Typically, more than that are taken
      (due to noise / imperfections in measurements)

Diffusion MRI

- Estimating diffusion tensor from measurements
  - Assume S0 is known
    - Measure with b = 0
  - Given : S(gn) for n=1,2,...,N
  - Goal : Find D
  - Solve an optimization problem
    - Minimize over D
      Sum_n | S(gn,p) - S0 exp (- b gn' D gn) |^2
    - Constraint : D is SPD
  - How to ensure that D is SPD ?

Diffusion MRI

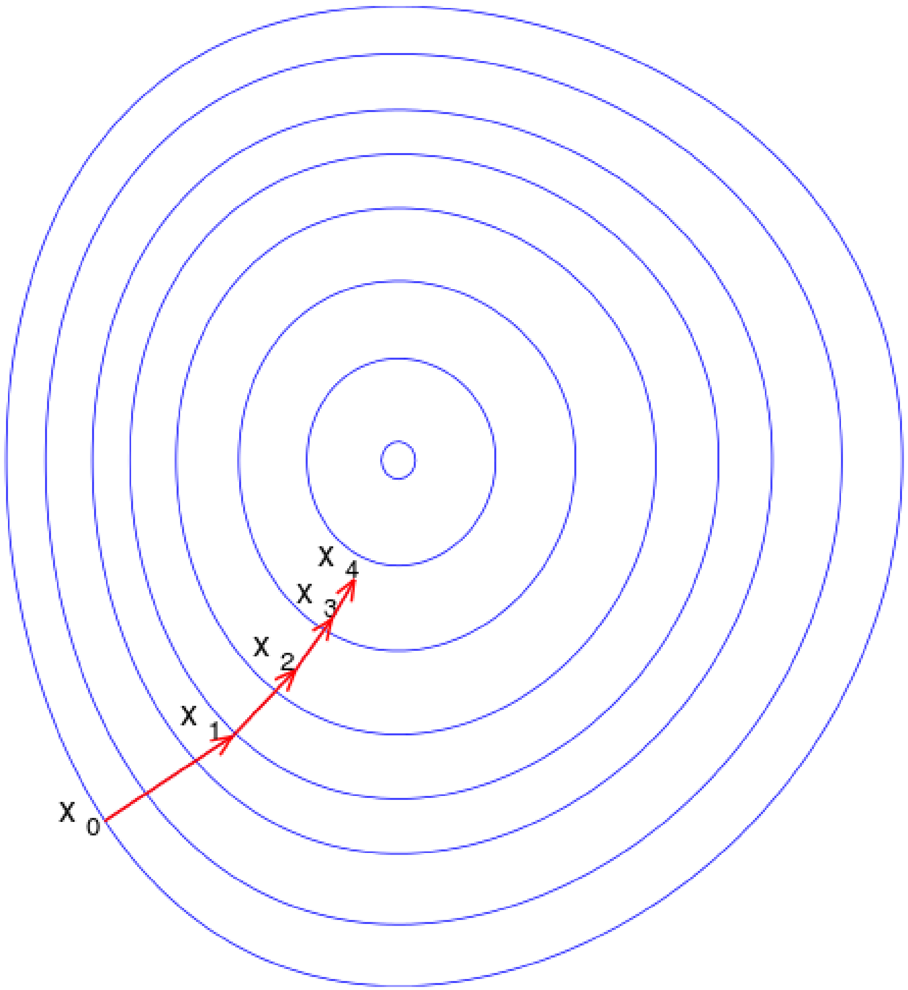
- Estimating diffusion tensor from measurements
  - How to ensure that D is SPD ?
  - Cholesky Decomposition
    - Every SPD matrix D can be written as
      - D = L L'
      - L = real lower-triangular matrix with positive diagonal entries
      - L has 6 unknowns (like D)
  - Optimization problem
    - Minimize over L
      Sum_n | S(gn) - S0 exp (- b gn' L L' gn) |^2
    - Constraint: L lower-triangular with diagonal elements > 0

Diffusion MRI

- Estimating diffusion tensor from measurements
  - Optimization using gradient-descent algorithm
  - Gradient-descent algorithm
    - Iterative optimization
    - Function to be optimized = “Objective Function”
    - Step 0: Start with an estimate of the solution
    - Step 1: Improve estimate a little bit
      - Update based on gradient of objective function
      - Handle constraint : don't let diagonal elements go < 0
    - Step 2: Repeat previous step until no improvement

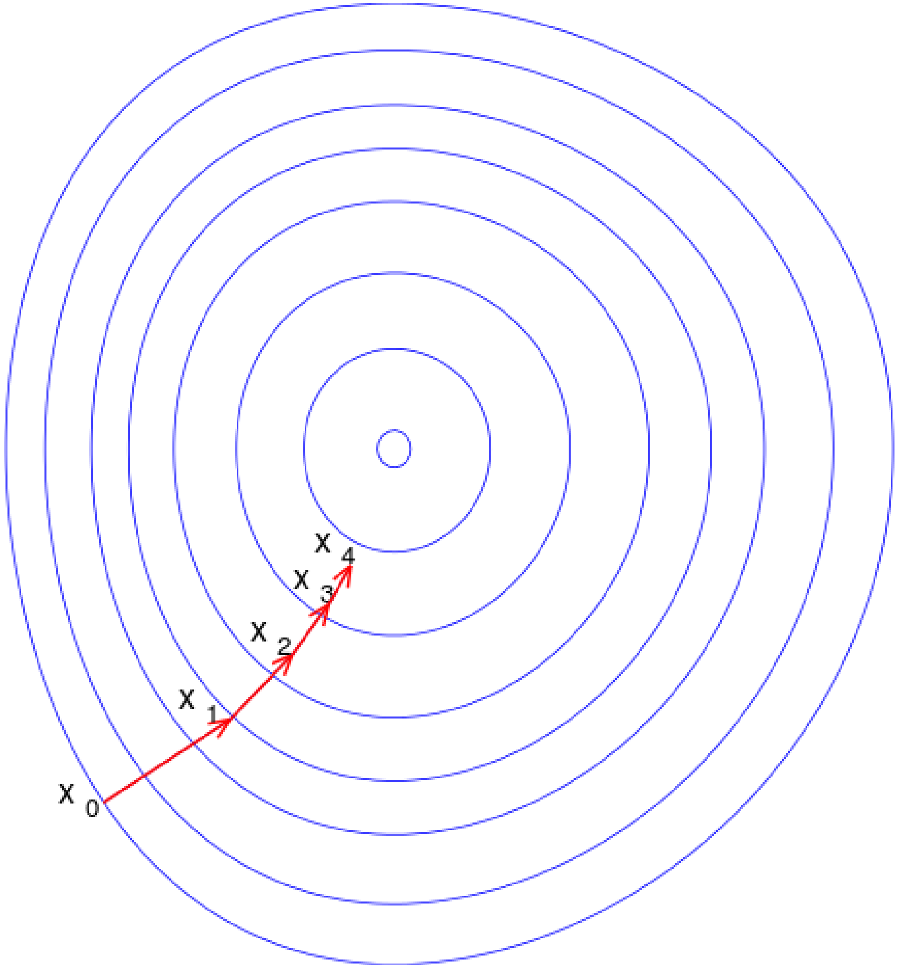
Diffusion MRI

- Estimating diffusion tensor from measurements
  - Gradient-descent algorithm
    - Blue Curves = Contours of objective function f(A,B)
    - Let current solution = (a,b)
    - Improved solution
      - Compute gradient at (a,b)
        - g(a,b) = (df/dA, df/dB) evaluated at (A=a,B=b)
      - Improved solution (a',b') = (a,b) – s g(a,b)
      - s > 0 is a small step size
      - How to choose s ?



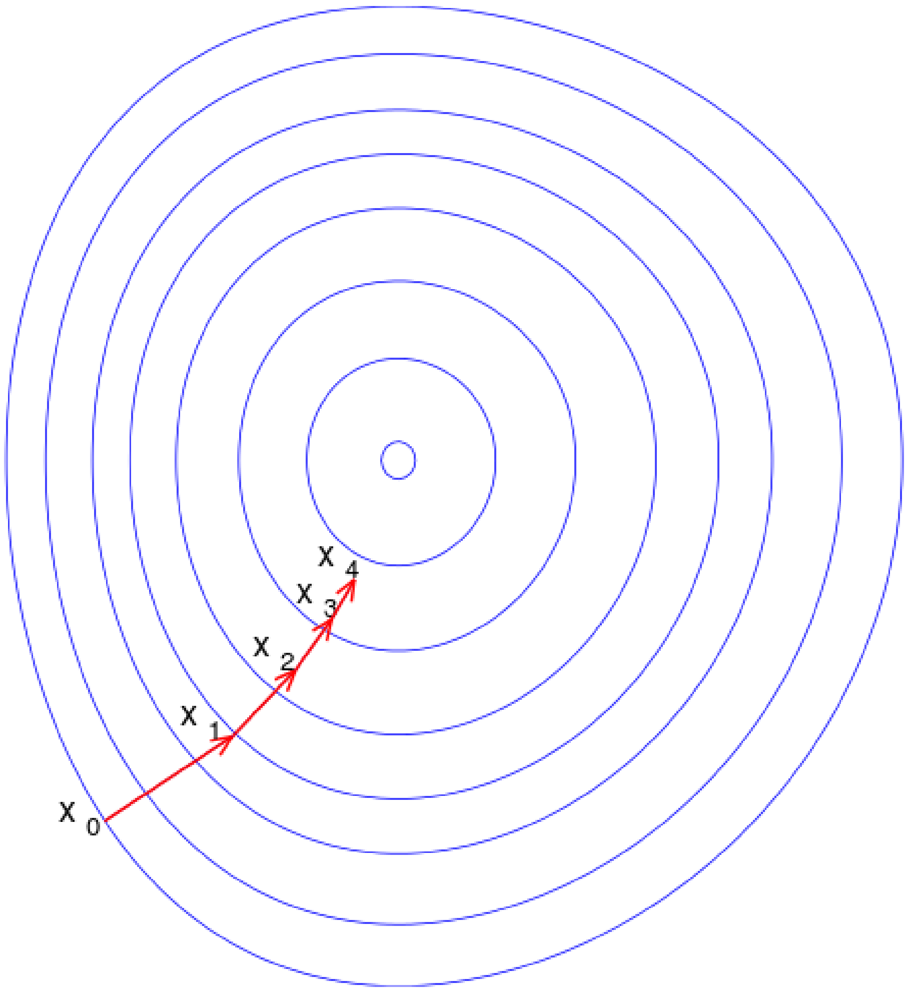
Diffusion MRI

- Estimating diffusion tensor from measurements
  - Gradient-descent algorithm (how to choose step size)
    - Start with some 's'
    - Guess updated solution (a',b') = (a,b) – s g(a,b)
    - Check if f(.) value reduces
      - If f(a',b') < f(a,b), improved solution found increase s by 10%
      - If f(a',b') >= f(a,b), reduce s by 50%
    - Repeat Guess+Check until improved solution found OR s becomes zero



Diffusion MRI

- Estimating diffusion tensor from measurements
  - Gradient-descent algorithm
    - Let current solution = (a,b)
    - Improved solution
      - Compute gradient at (a,b)
        - g(a,b) = (df/dA, df/dB) evaluated at (A=a,B=b)
      - Improved solution (a',b') = (a,b) – s g(a,b)
      - s = chosen step size
    - Repeat last step until improvement is negligible



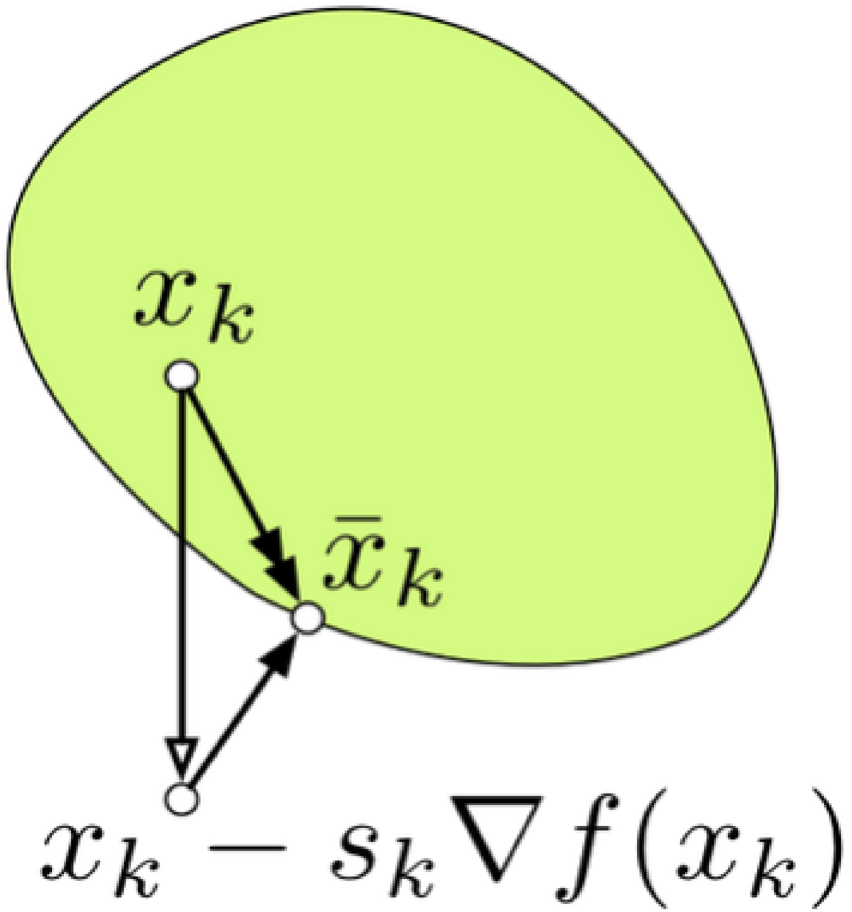
Diffusion MRI

- Estimating diffusion tensor from measurements
  - Gradient descent guaranteed to convergence
    - Given sufficiently small step size
    - If maximum curvature of objective function is known, then step size can be fixed appropriately

Diffusion MRI

- **Projected gradient descent**
  - Algorithm:
    (1) Take a step along negative gradient
    (2) Project back onto constraint set
  - When does this converge ?

    (1) When gradient is zero
    OR
    (2) When ...



Diffusion MRI

- **Newton's method**
  - Scalar-valued function f(.) of single-variable x
  - Taylor series expansion (upto 2nd-order terms)

$$f_T(x_n + \Delta x) = f(x_n) + f'(x_n)\Delta x + \frac{1}{2} f''(x_n) \Delta x^2$$

  - Minimize $f(x_n + \Delta x)$ w.r.t. Δx
    - Take derivative w.r.t. Δx
    - Equate it to 0

$$f'(x_n) + f''(x_n)\Delta x = 0$$

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}$$

Diffusion MRI

- Newton's method
  - Scalar-valued function f(.) of multi-variable x
  - Taylor series expansion (upto 2nd-order terms)

$$f(\mathbf{x}+\Delta\mathbf{x}) \approx f(\mathbf{x}) + J(\mathbf{x})\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^{\mathrm{T}} H(\mathbf{x})\Delta\mathbf{x}$$

    - J = Jacobian of f(.)
      - Transpose of gradient (column) vector G(.)
    - H = Hessian matrix of f(.)

Diffusion MRI

- **Jacobian**
  - $F : R^n \rightarrow R^m$
    - m-valued function of n variables

$$\mathbf{J} = \frac{d\mathbf{F}}{d\mathbf{x}} = \left[\frac{\partial \mathbf{F}}{\partial x_1} \quad \cdots \quad \frac{\partial \mathbf{F}}{\partial x_n} \right] = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_n} \end{bmatrix}$$

Diffusion MRI

- **Hessian**
  - For real-valued n-variable functions f(.) only
  - H f(.) = J (G f(.))
    - Gf (.) is a n-valued function of n variables

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\,\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\,\partial x_n} \\ \frac{\partial^2 f}{\partial x_2\,\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\,\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n\,\partial x_1} & \frac{\partial^2 f}{\partial x_n\,\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Diffusion MRI

- Newton's method
  - Scalar-valued function f(.) of multi-variable x
  - Taylor series expansion (upto 2nd-order terms)

$$f(\mathbf{x}+\Delta\mathbf{x}) \approx f(\mathbf{x}) + J(\mathbf{x})\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^{\mathrm{T}} H(\mathbf{x})\Delta\mathbf{x}$$

    - J = Jacobian of f(.)
      - Transpose of gradient (column) vector G(.)
    - H = Hessian matrix of f(.)
  - Minimize f(x+Δx) w.r.t. Δx: $\mathbf{x}_{n+1} = \mathbf{x}_n - [Hf(\mathbf{x}_n)]^{-1}\nabla f(\mathbf{x}_n)$
    - Take derivative w.r.t. Δx
    - Equate it to 0
    - $0 = G(x)^T + 0.5\, \Delta x^T (H(x) + H(x)^T)$
    - $0 = G(x) + H(x)\, \Delta x$

Carl Gustav Jacob Jacobi

- Mathematician
- Prodigy
  - Home schooled by his uncle
  - Age 12, joins advanced secondary school (gymnasium)
  - In 6 months, moved to senior (4^{th}) year
  - Remained in same class until 16 (when Univ. accepted)
  - Age 21, got PhD (Univ. Berlin)
  - Age 25, full prof at Konisberg Univ.
  - Age 38, breakdown from overwork
  - Visits Italy to relax, back to Berlin, retires, pension
  - Age 47, smallpox, death

Otto Hesse

- Mathematician
- Advisor: Jacobi
- Students
  - Gustav Kirchhoff
  - Kirchhoff students
    - Max Planck → Nobel prize in Physics, 1918
    - Dmitri Mendeleev → Periodic table

Diffusion MRI

- Newton's method
  - If objective function is quadratic:
    - Converges in 1 iteration
  - If starting point sufficiently near a minimum:
    - Can converge faster then gradient descent
  - In general, convergence NOT guaranteed
    - Can diverge
    - Introduce a step size < 1 (can tune adaptively)
  - Requires more work:
    - 2^{nd} derivatives matrix, $O(n^2)$

Diffusion MRI

- **Gauss-Newton algorithm**
  - For non-linear least-squares optimization
    - n variables: *β = (β1, …, βn)^T*
    - m-valued function (e.g., m measurements): *r = (r1, …, rm)^T*
    - **Residual** i: $$r_i(\boldsymbol{\beta}) = y_i - f(x_i, \boldsymbol{\beta})$$
      - For fitting tensor D, n-th residual = S(gn,p) - S0 exp (- b gn' D gn)
    - Objective function: $$S(\boldsymbol{\beta}) = \sum_{i=1}^{m} r_i(\boldsymbol{\beta})^2$$
      - Minimize || r^T r ||
      - Real-valued
    - Gradient of real-valued obj. function
      - J-th component = $$g_j = 2\sum_{i=1}^{m} r_i \frac{\partial r_i}{\partial \beta_j}$$

Diffusion MRI

$$g_j = 2\sum_{i=1}^{m} r_i \frac{\partial r_i}{\partial \beta_j}$$

- Gauss-Newton algorithm
  - For non-linear least-squares optimization
    - Hessian of real-valued obj. function:

$$H_{jk} = 2\sum_{i=1}^{m}\left(\frac{\partial r_i}{\partial \beta_j}\frac{\partial r_i}{\partial \beta_k} + r_i \frac{\partial^2 r_i}{\partial \beta_j \partial \beta_k}\right)$$

    - Approximate Hessian of real-valued obj. function:
      - Ignore 2nd term: $\left| r_i \frac{\partial^2 r_i}{\partial \beta_j \partial \beta_k} \right| \ll \left| \frac{\partial r_i}{\partial \beta_j}\frac{\partial r_i}{\partial \beta_k} \right|$

$$H_{jk} \approx 2\sum_{i=1}^{m} J_{ij} J_{ik} \qquad J_{ij} = \frac{\partial r_i}{\partial \beta_j}$$

      - ***J*** = “Jacobian” of m-valued function *r(.)*
        - NOT objective function

Diffusion MRI

$$g_j = 2\sum_{i=1}^{m} r_i \frac{\partial r_i}{\partial \beta_j} \qquad H_{jk} = 2\sum_{i=1}^{m} \left(\frac{\partial r_i}{\partial \beta_j} \frac{\partial r_i}{\partial \beta_k}\right.$$

- Gauss-Newton algori
  - For non-linear least-squares optimization

$$J_{ij} = \frac{\partial r_i}{\partial \beta_j}$$

    - Apply Newton's method using approximate Hessian
    - Gradient and approx. Hessian

$$\mathbf{g} = 2\mathbf{J_r}^{\mathsf{T}}\mathbf{r}, \quad \mathbf{H} \approx 2\mathbf{J_r}^{\mathsf{T}}\mathbf{J_r}$$

    - Update: $\boldsymbol{\beta}^{(s+1)} = \boldsymbol{\beta}^{(s)} + \Delta; \quad \Delta = -\left(\mathbf{J_r}^{\mathsf{T}}\mathbf{J_r}\right)^{-1}\mathbf{J_r}^{\mathsf{T}}\mathbf{r}$
    - Works like Newton's method when:
      - Residuals have small magnitude in region near minimum, starting point
      - Functions aren't very nonlinear, i.e., 2^{nd} derivatives small in magnitude

$$\left| r_i \frac{\partial^2 r_i}{\partial \beta_j \partial \beta_k} \right| \ll \left| \frac{\partial r_i}{\partial \beta_j} \frac{\partial r_i}{\partial \beta_k} \right|$$

Diffusion MRI

- Gauss-Newton algorithm
  - For non-linear least-squares optimization
  - Advantages over Newton's method:
    - 2nd derivatives of residuals aren't required
  - In general, convergence NOT guaranteed
    - Can diverge
    - Introduce a step size < 1 (can tune adaptively)
    - A better way is the Levenberg-Marquardt algorithm

Sir Isaac Newton

- Mathematician, physicist, astronomer, 1642 - 1726
- His epitaph, suggested by Alexander Pope,
  - Nature and nature's laws lay hid in night;
    God said "Let Newton be" and all was light.
- Newton wrote in his memoir
  - I do not know what I may appear to the world, but to myself,
    I seem to have been only like a boy playing on the sea-shore,
    and diverting myself, in now and then,
    finding a smoother pebble or a prettier shell than ordinary,
    whilst the great ocean of truth lay all undiscovered before me.
- In a letter to rival Robert Hooke
  - If I have seen further, it is by standing on the shoulders of giants.

Sir Isaac Newton

- Born premature, father died, wasn't expected to live
- In school, loved chemistry.
  Mother pulled him out of school. Planned to make him a farmer. Newton found farming monotonous
- Went to Cambridge Univ., to study, later professor
- *Philosophiae Naturalis Principia Mathematica*
  - Single most influential book on physics & perhaps science
- As a person
  - Never married or had (m)any friends. In later years, mentally unstable due to pride, insecurity, alchemy

Diffusion MRI

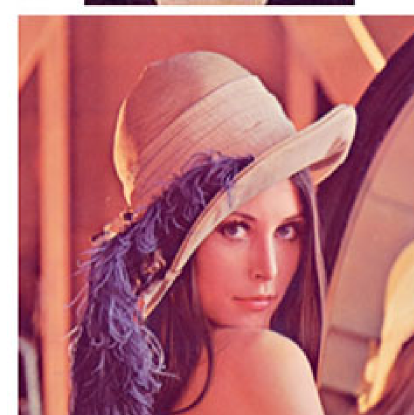
- **Levenberg-Marquardt algorithm**
  - For non-linear least-squares optimization
  - Gauss-Newton update (step size 's'):
    - $\Delta = - s\ (J^T J)^{-1}\ J^T r$
  - **Modification 1**
    - $\Delta = - s\ (J^T J + \boldsymbol{\lambda I})^{-1}\ J^T r$
    - Damping factor $\lambda > 0$
    - Large $\lambda \rightarrow$ move in direction of gradient
    - Small $\lambda \rightarrow$ Gauss-Newton update

Diffusion MRI

- Levenberg-Marquardt algorithm
  - For non-linear least-squares optimization
  - Gauss-Newton update (step size 's'):
    - $\Delta = - s (J^T J)^{-1} J^T r$
  - **Modification 2**
    - $\Delta = - s (J^T J + \boldsymbol{\lambda}\ \mathbf{diag(J^T J)})^{-1} J^T r$
    - Diagonal of $J^T J$
      ~ diagonal of Hessian of real-valued obj. function
      = 2nd derivatives of real-valued obj. function
      = curvature of obj. function in each variable's dimension
    - Large $\lambda \rightarrow$ gradient scaled differently for each variable s.t. scaling factors for gradient are small when curvature is large & scaling factors for gradient are large when curvature is small

Sonnet for Lena – Thomas Colthurst

- O dear Lena, your beauty is so vast
- It is hard sometimes to describe it fast.
- I thought the entire world I would impress
- If only your portrait I could compress.
- Alas! First when I tried to use VQ
- I found that your cheeks belong to only you.
- Your silky hair contains a thousand lines
- Hard to match with sums of discrete cosines.
- And for your lips, sensual and tactual
- Thirteen Crays found not the proper fractal.
- And while these setbacks are all quite severe
- I might have fixed them with hacks here or there
- But when wavelets took sparkle from your eyes
- I said, "Heck with it. I'll just digitize.”



The “Lenna” Image

Summer is Here

Winter is Here Too !

Summer is Here !