Analyzing Neural Time Series Data

1.1 What is Cognitive Electrophysiology?

1.2 What is the purpose of this book?

- Gain a deeper understanding of data analysis, without requiring formal training in math and computer science.

1.3 Why You Shouldn't Use EEG Analysis Packages

- Lack of flexibility

1.4 Why Program Analysis, and Why in Matlab?

- Easy to use
- Matlab has many EEG analysis toolboxes
- Easily Sharable with others
- **Octave –** Free Matlab alternative
- Matlab Toolboxes for EEG:
  - o Signal-processing Toolbox
  - o Statistics Toolbox
  - o Image-processing Toolbox

1.5 How to Best Learn from and Use This Book

- In front of a computer, running matlab
- In order from simple to more complicated analysis

1.6 Sample Data and Online Code

- www.mikexcohen.com/book

1.7 Terminology in this Book

- Always refer to as EEG, but also applies to other data collection techniques

1.8 Exercises

- Do them

1.9   Is Everything There is to Know about EEG Analysis in this Book?

- No, but it has the most useful, promising, and accepted approaches for linking EEG dynamics to cognitive processes.

2. Advantages and Limitations of Time- and Time-Frequency-Domain Analysis

2.1 Why EEG?

- Captures cognitive dynamics in the time frame the cognition occurs
  - Theta-band (4-8Hz)
    - Memory / Cognitive Control
    - Slower frequency
  - Gamma-band (30-80Hz)
    - Faster frequency
- Measures neural activity
  - Voltage fluctuations
- Multi-Dimentional
  - Voltage changes over time and space
    - Time
    - Space
    - Frequency
    - Power (strength of frequency-band-specific-activity)
    - Phase (timing of activity)

2.2 Why Not EEG?

- Not suited for studies where precise functional localization is important
- Not suited for testing hypotheses of deep brain structures
- Not suited for questions concerning slow processes with uncertain and variable time courses.

2.3 Interpreting Voltage Values from the EEG Signal

- Recorded in microvolts
  - Change in measured electrical potential between the electrode and a reference electrode placed elsewhere on the head
    - Readings change based on choice of reference and time period for baseline subtraction
    - Values differ across subjects because of

- Skull shape / thickness
- Scalp preparation
- Orientation of dipole in brain
- Cortical folding
- If subject washed their had the morning of
  - Raw Values are difficult to compare and should not be overinterpreted
  - Interpret general pattern of effects and time-frequency-electrode characteristics of effects, rather than difference in microvolt values
- Analysis using scale transformations are advantageous
  - Individual differences in raw voltage values are eliminated

2.4 Advantages of Event-Related Potentials

1. Simple, fast to compute, require few analysis assumptions
2. High temporal precision and accuracy
   a. Applying low / high pass filters decreases precision
3. Extensive literature of ERP findings
4. Quick and useful data quality check of single subject data

Should be inspected for each subject to make sure data was properly collected.

2.5 Limitations of ERPs

1. Null results – many EEG data dynamics that aren't represented in ERP
2. Provide limited opportunities for linking results to physiological mechanisms
   a. Mechanisms that produce ERPs are less well understood than those producing oscillations
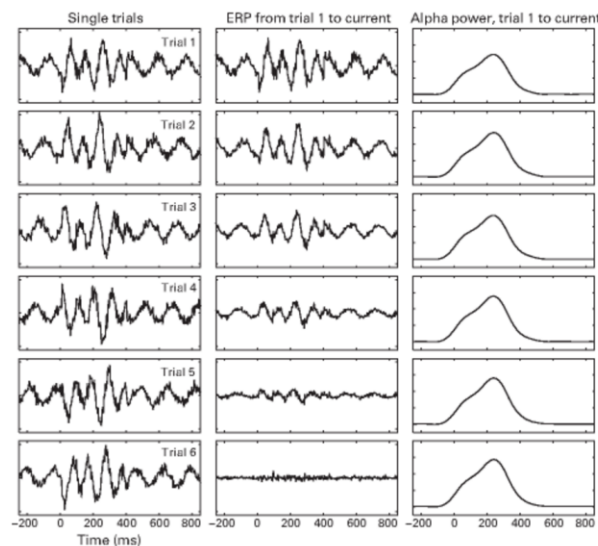

Figure 2.1

Simulated data showing how time-locked but not phase-locked activity (left column) is lost in ERP averaging (middle column) but is visible in band-specific power (right column). Each row in the left column shows a different trial, and each row in the middle and right columns shows averages from the first until the current trial.

2.6 Advantages of Time-Frequency-Based Approaches

1. Results can be interpreted in terms of neurophysiological mechanisms of neural oscillations
2. Oscillations are the most promising bridge that links findings from multiple disciplines within neuroscience and across multiple species
3. Many task-relavent dynamics that are retrievable using only time-frequency-based approaches
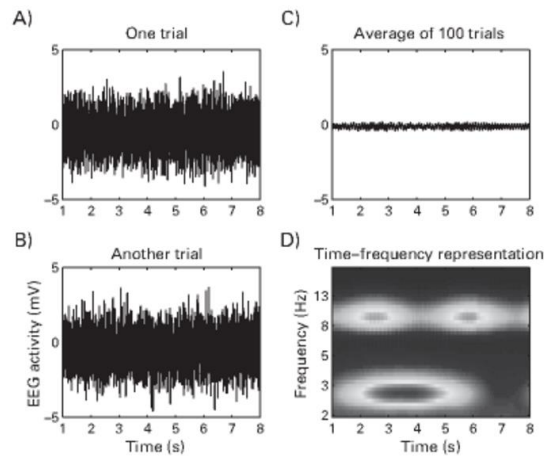

Figure 2.2
Simulated data showing that complex and multifrequency information contained in EEG data may have no representation in the ERP, if that information is non-phase-locked. One hundred trials were simulated; panels A and B show example trials. Panel C shows the ERP of those 100 trials, and panel D shows the time-frequency power. (This figure is adapted from Cohen 2011b).

2.7 Limitations of Time-Frequency-Based Approaches

1. Decrease of temporal precision → from time-frequency decomposition
   a. Lower frequency = more loss of temporal precision
   b. Low pass filtering of ERP data diminishes temporal precision
2. Large number of analyses that can be applied to EEG data
   a. Complexity of those analyses is intimidating

2.8 Temporal Resolution, Precision, and Accuracy of EEG

- Resolution → # of data samples per unit time
- Precision → certainty of measurement at each time point
- Accuracy → Relationship between timing of EEG signal and timing of biophysical events that lead to EEG signal
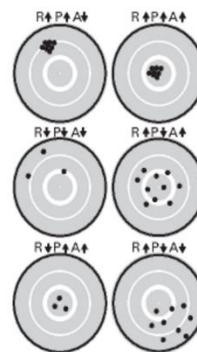 o Distance of the dots to the center of the bull's eyes

Figure 2.4
Bull's-eye illustration of the differences among resolution (R), precision (P), and accuracy (A). Up-and-down arrows indicate high and low levels. Resolution is illustrated by the number of dots, precision is illustrated by the spread of the dots, and accuracy is illustrated by the distance of the dots away from the center of the bull's-eye.

- Temporal resolution (Generally 250Hz – 1000Hz)
    - Determined by sampling rate of acquisition (100s – 1000s of samples per second)
    - Allows for extraction of frequency-band-specific information
    - Low temporal resolution
        - Extracting Delta-Band Power
    - High temporal resolution
        - Cross-frequency Coupling
- Temporal Precision
    - Depends on applied analysis / Selected Parameters / Frequency Band
        - Higher frequency (generally) = higher temporal precision
    - High Precision
        - Unfiltered ERPs
    - Low Precision
        - 1Hz bandpass-filtered activity
- When temporal precision is decreased by analysis, temporal resolution can decrease to match precision
- Temporal Accuracy – Is extremely high because
    - Brain electrical activity travels instantaneously from neurons to electrodes

2.9 Spatial Resolution, Precision, and Accuracy of EEG

- Spatial Resolution
    - Depends on # of electrodes
- Spatial Precision
    - Low, but can be improved by spatial filtering
        - Surface Laplacian / Adaptive source-space-imaging techniques
- Spatial Accuracy
    - Low
    - Electrodes collect data on cluster of neurons
- Organization of Brain networks
    - Microscoping Scale → less than a few cubic millimeters
    - Mesoscoping Scale → patches of cortex of several cubic (mm – cm)
    - Macroscopic Scale → large regions of cortex (many cubic cm)

2.10 Topographical Localization vs. Brain Localization

- Topographical → identify electrodes that show max effect
    - Description of observation
- Brain → identify regions in brain that generate activity (measured on scalp)
    - Interpretation of a result, supported by combination of theory, previous research, and data analysis in combination with spatial filtering

2.11 EEG or MEG?

- MEG → better at detecting high-frequency activity
    - Also better for source localization
- EEG → works better for radial sources

2.12 Costs of EEG Research

- Good EEG headsets are expensive
- If you have no equipment, the cost of EEG per subject gets close to, and sometimes supersedes the cost of MEG or MRI research

Interpreting and Asking Questions about Time-Frequency Results

3.1 EEG Time Frequency: Basics

- Rhythmic activity reflects neural oscillations
    - Fluctuations in excitability of populations of neurons
- Frequency
    - Speed of oscillation (Hz) – cycles per second
- Power
    - Energy in a frequency band that is squared amplitude of oscillations
- Phase
    - Position along sine wave at any given time point (Radians or Degrees)
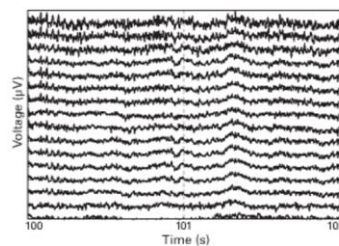- Power and phase are independent of eachother

Figure 3.1
Raw EEG data (after 0.1-Hz high-pass filtering) showing oscillations at different speeds and for different lengths of time. Each line corresponds to an electrode.
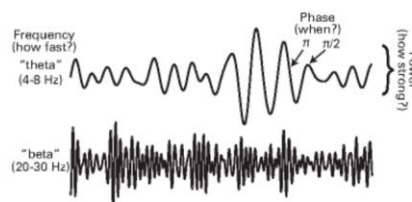

Figure 3.2
The three dimensions that define oscillations: frequency, power, and phase.

- EEG measures meso / macro – scopic cortical electrical activity
- Brain rhythms
    - Delta (2-4Hz)
    - Theta (4-8Hz)
    - Alpha (8-12Hz)
    - Beta (15-30Hz)
    - Lower Gamma (30-80Hz)
    - Upper Gamma (80-150Hz)
    - Subdelta (<600Hz)
- Generally: Better precision in time or frequency → Poorer precision on other domain

- Background activity exists
  - So we require baseline normalization
  - To remove artifacts that are consistently present, but not relevant to area of focus
  - Phase-locked
    - Phase is same or similar on each trial
  - Non-phase locked
    - Phase is different on each trial
- Spatial Autocorrection
  - Sometimes electrodes record activity from the same brain sources

3.2 Ways to View Time-Frequency Results

- Frequency Slice
  - Power (energy at each frequency band) vs. Frequency
  - No use of time
  - Useful when little time varying changes in frequency characteristics are expected
    - Ex. Resting State or Sleep Stage
- Time Slice
  - Select one frequency band
  - Plot its activity over time
  - Useful when comparing activity across multiple conditions or electrodes
    - And a prior reason to focus on a specific frequency
- Space Slice
  - One time-frequency point (or average over mult. Adjacent time freq. points)
  - Over electrodes on a topographical plot
  - Useful in visualizing topographical distribution of effect and facilities of topographical localization
- Time-Frequency Slice
  - Frequency vs. Time
  - Typically higher frequencies are plotted at the top
  - Color can be used to reflect:
    - Power
    - Phase clustering
    - Connectivity
    - Correlation Coefficient
- Other Slices
  - Activity of each electrode over time
  - Activity over different frequencies as a function of physical distance of subject to a target
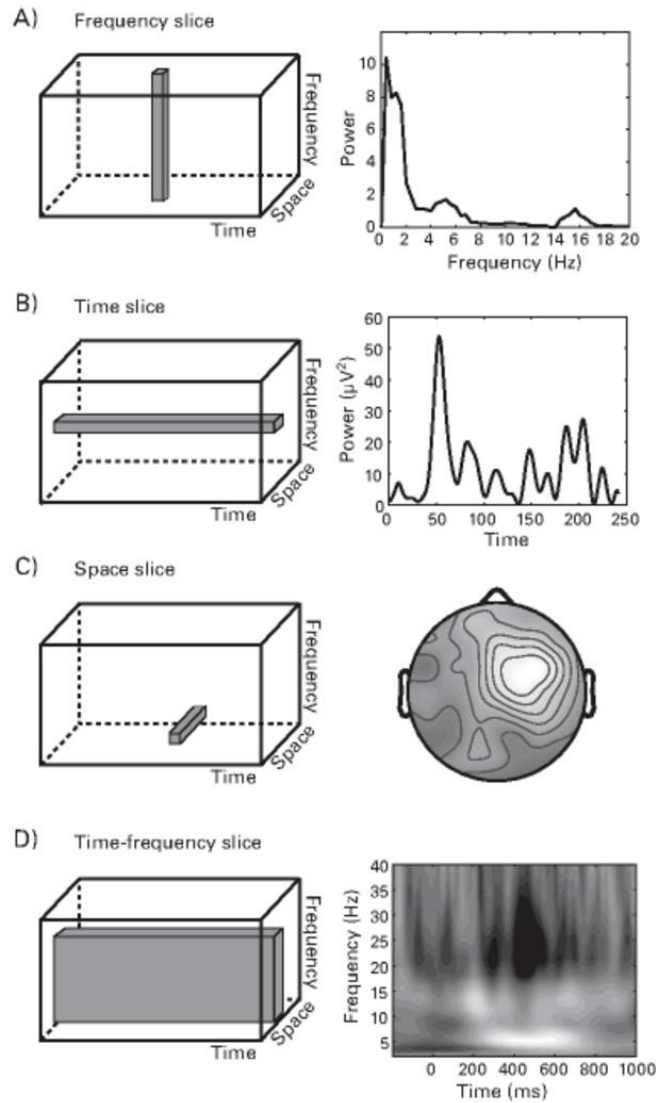
Figure 3.3
The data cube, containing information over time, frequency, and space, is difficult to view or conceptualize and therefore is sliced in different ways to illustrate 1-D or 2-D snapshots of the results.

3.3 Tfviewerx and erpviewerx

- Simultaneously shows time-frequency plot from one electrode and the topographical map at a selected time-frequency point

3.4   How to View and Interpret Time-Frequency Results

1. Determine what is shown in the plot
   - Power
   - Phase Clustering
   - Connectivity
   - Correlation with behavior
   - Understand conceptually what is being plotted
2. Inspect the ranges and limits of the plot
   - What are the time and frequency ranges?
   - Is there activity that is cutoff by boundaries?
   - Color limits? Symmetric or asymmetric or bounded by 0?
3. Inspect the Results
   - Activity at multiple freq. and time windows? Or all centered at one time-frequency?
   - Activity duration short or long? Freq. band-limited or spans multiple frequency bands?
   - Activity during prestimulus period?
   - Topographical specificity (are effects present selectively at some parts of scalp)?
   - Which electrode(s) are shown and why?
4. Link Results to Experiment
   - What does time = 0 refer to?
   - Mult. Events in experiment? How are they represented in time-freq. Results?
   - Results make sense? Are they consistent?
   - What do results suggest about cog. Process under investigation?
   - Results prove any new information about brain function?
5. Understand Statistical Procedures Used to Support the Interpretations
   - Statistical Threshold?
   - Hypothesis-driven or exploratory and data-driven?
   - Exploratory approaches generally require conservative statistical thresholds and corrections for multiple comparisons over time, frequency, and electrodes.
   - Hypothesis driven → increase sensitivity and theoretical relevance. Less stringent thresholds ($p < 0.05$ is acceptable)
   -If hypothesis driven, how were time-frequency-space windows selected for statistical analysis?

3.5 Things to Be Suspicious of When Viewing Time-Frequency Results

- Horizontal / Vertical stripes in Time-Frequency Plot
  - Ripple artifacts from poor filter construction
    - Filter widths are:
      - Too narrow
      - Applied to too little data
- Brief and large-power effects at high frequencies
  - EEG artifacts such as:
    - Amplifier saturation
    - Noise spike from bad electrode
- Broadband effects
  - Mechanical noise or excessive muscle activity from jaw or neck
- Fast color changes over time or frequency
  - Mistake in analysis – real part of analytical signal plotted instead of power
  - Fast change in lower frequencies are more suspect than in high frequencies
    - Increased temporal smoothing at lower frequencies
- Strange topographical distributions
  - Noisy or bad electrodes
  - Incorrect mapping between electrode label and physical location
  - High-pass spatial filters (Laplacian) increase topographical localization and highlight local spatial features
- High-frequency activity (over 100Hz)
  - Has low signal-to-noise ratio and may require many trials and special analysis techniques to enhance signal noise
- Low-Frequency Activity (<1z)
  - High pass filters that attenuate activity in lower frequencies
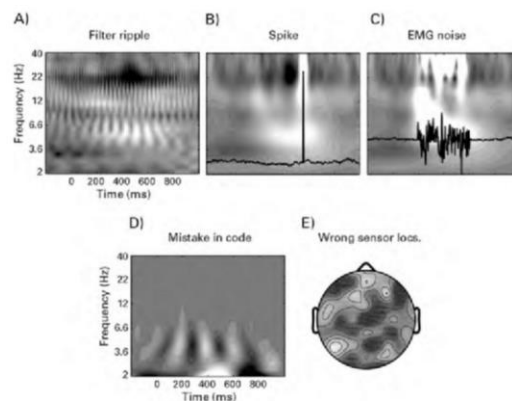 - Apply high pass filter of 0.1 or 0.5 Hz to eliminate slow fluctuations

Figure 3.5
Some features of time-frequency results that should arouse suspicion, although they are not necessarily artifacts. In panels B and C, the offending single trial (out of 99 otherwise good-data trials) is superimposed on the time-frequency plot (EEG trace amplitude is arbitrarily scaled). The topographical map in panel E was produced by randomly swapping electrode label-location mappings.

3.6 Do Results in Time-Frequency Plots Men That There Were Neural Oscillations

- On one hand, EEG measures summed field potentials of populations of neurons
    - Strongly oscillatory
- On the other, Fourier's theorem specifies that any signal can be represented using sine waves, and thus, even nonoscillatory signals have a representation in a time-frequency plot

4 Introduction to Matlab Programming

4.1 Write Clean and Efficient Code

- Clean code is easy to read and understand
- 1. Write Brief Comments before the code
- 2. Group lines of code by their common purpose
- 3. Use sensible, interpretable variable names
- Perform matrix manipulations instead of loops when possible
- Verbally plan out your code on paper before coding it

4.2 Using Meaningful File and Variable Names

- Put "I" at end of counting variables in loops

4.3 Make Regular Backups of Your Code and Keep Original Copies of Modified Code

4.4 Initialize Variables

- Reserve space in Matlab buffer by creating variable before populating it with data
  - o Helps avoid memory crashes
  - o Helps prefent data from previous iterations of loop contaminating current iterations
  - o Helps you think about size, dimensions, and contents of large and important variables in advance

5 Introduction to the Physiological Basis of EEG

5.1 Biophysical Events That Are Measurable with EEG

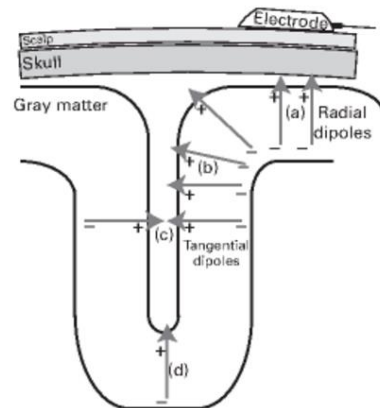- Magnetic fields are perpendicular to electric fields and pass through skull/scalp unimpeded

Figure 5.1

Illustration of dipoles in different orientations with respect to the skull. The dipoles illustrated in (a) will contribute the strongest signal to EEG, whereas the dipoles illustrated in (b) will contribute the strongest signal to MEG. The dipoles illustrated in (c) are unlikely to be measured because the dipoles on opposing sides of the sulcus produce electrical fields that are likely to cancel each other. The dipole illustrated in (d) will make a smaller contribution to EEG than dipole (a) because it is further away from the electrode. (This figure is inspired by figure 1 of Scherg 1990.)

- Populations of neurons in subcortical structures are not arranged in geometrically parallel orientation
  - In synchronous population activity, electrical fields generated by individual neurons are likely to cancel eachother out at the macroscopic scale
- Slow fluctuations (<1Hz) are difficult to measure with EEG
  - Most amplifiers have built-in high-pass filters that attenuate very slow fluctuations because they may cause amplifier saturations
    - There are DC-coupled amplifiers for fluctuations below 1Hz
- Fast fluctuations (>100Hz)
  - High frequency activity generally has low power → difficult to distinguish from noise

5.2 Neurobiological Mechanisms of Oscillations

- Oscillation – Rhythmic alteration of states
  - Rhythmic fluctuations in excitability of neuron populations
- Interaction between inhibitory interneurons and excitatory pyramidal cells
  - Oscillations can be produced by excitatory or inhibitory neurons

5.3 Phase-Locked, Time-Locked, Task-Related

- Phase-locked
  - Aligned with time = 0
  - Observed in time-domain averaging and in time-frequency-domain averaging
- Non-Phase-Locked
  - Time locked, but not phase locked with time = 0
  - Time-frequency-domain averaging
- Time and/or frequency characteristics change as a function of engagement in task events
  - Background activity does not
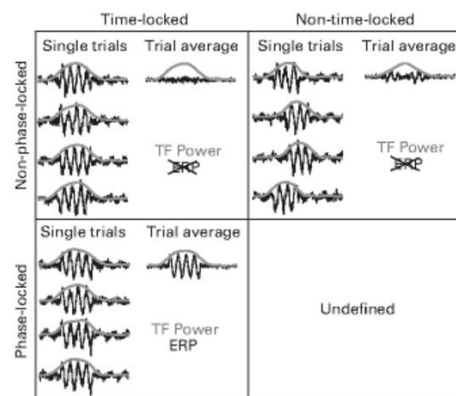 - Apply baseline normalization to remove background activity

Figure 5.2

Illustration of whether time-frequency (TF) power and the ERP can measure phase-locked, non-phase-locked, time-locked, and non-time-locked activity. The left column of each cell shows four trials of simulated data, and the right column of each box shows the average of those four trials. Black lines show the raw time series, and gray lines show the time course of 10-Hz power. The ERP captures only phase-locked and time-locked activity. Time-frequency power can measure time-locked activity regardless of whether it is phase-locked or non-phase-locked. Activity that is not time-locked can be measured with time-frequency power, although the results will be smoothed and thus less temporally precise.

5.4 Neurophysiological Mechanisms of ERPs

- Additive
  - ERP reflects a signal elicited by an external stimulus (picture / sound)
  - Or internal event (manual response)
- Phase reset
  - ERP results from sudden alignment of phases of ongoing oscillations
  - Stimulus appears → ongoing oscillation at a particular frequency band is reset to a specific phase value
- Amplitude Asymmetry or Baseline Shift
  - Outward-going currents are less detectable from scalp
    - Producing asymmetry in oscillations measured by EEG
      - Unequally distributed peaks / troughs
  - Changes in overall power could produce asymmetries in ongoing oscillations

5.5 Are Electric Fields Causally Involved in Cognition?

- Theory 1: Long-term potentiation occurs at theta-band oscillations
- Theory 2: Timing of many neurons is constrained by local field potential
- Theory 3: Interregional oscillatory synchronization is a mechanism underlying the transmission of information across neural networks, and the synchronization-mediated connectivity is crucial for perceptual and cognitive processes

5.6 What if Electric Fields Are Not Causally Involved in Cognition?

- Doesn't stop progress

6 Practicalities of EEG Measurement and Experiential Design

6.1 Designing Experiments: Discuss, Pilot, Discuss, Pilot

- Discuss with colleagues before data collection
- Test run without EEG headset
- Perform on 2 people and fully analyze those datasets

6.2 Event Markers

- Square-wave pulses sent from stimulus-delivering computer to EEG amplifier
- Recorded as separate channel in raw data file
    - Amplitude used to encode specific events (stimulus onset or response)
    - During data importing, markers are converted to labeled time stamps
        - Indicate when different events occurred
- Used to time lock EEG data
- Used to reconstruct different conditions and responsebetters
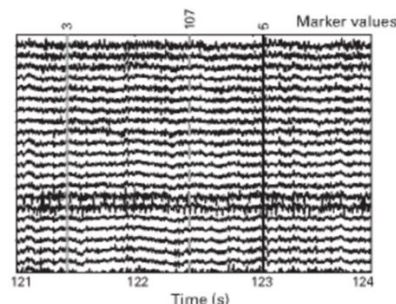   - **Better more detail in these markers, More =**

Figure 6.1

Example EEG data showing 3 s of data and three experiment markers. The experiment markers are represented as vertical lines, and the numbers on top of the vertical lines correspond to particular events. In this case the numbers 3 and 5 refer to two response buttons being pressed by the subject, and the number 107 corresponds to a particular stimulus. This picture was made using the eeglab lab function eegplot.

- Check for overlapping and dropped markers
- Temporal duration
    - Time when marker has non-zero value
    - (Should be at least a few samples ~5 ms)
- Data is useless without markers
    - Test by:
    - Sending codes 1-256 with 10ms spaces between markers

6.3   Intera- and Intertrial Timing

- Space out time between tasks by a few seconds (Intertrial Interval ~1000ms)
  - Allow brain response to subside after a task
  - Different tasks will take different times to subside
    - Ex. Pictures that evoke emotion
- Period of time for baseline normalization of task related data?
- What frequencies to analyze.
- Time Freq. Decomposition
  - End before trial onset (-500 to -200ms)
    - Bc. Temporal filtering may cause early poststimulus activity to leak into the prestimulus baseline period
- ERPs
  - End at time = 0
- Subjects nearly always generate temporal expectations about when the next trial will occur
- Constant Time Intervals
  - People can mentally prepare for an upcoming event
- Random Time Intervals
  - People can try to guess when the next event is upcoming

6.4   How Many Trials You Will Need

- Depends on:
  - Signal-to-noise ratio (how clean vs. how noisy) of the data
  - Size of the effect
  - Type of analysis to be performed
- Usually minimum of:
  - 50 trials per condition per subject
  - But there are unique cases where less is needed

6.5   How Many Electrodes You Will Need

- Depends on Experiment and what you're looking for:
  - Brain Source Reconstrucion Analysis
    - 100+ electrodes
  - Measure P3 Amplitude
    - 3 electrodes
      - Central parietal cortex for P3
      - Reference
      - Ground
  - At least 64 usually
    - Consider time to prepare the subject with an EEG cap
    - Storage and processing capacity required for more data

6.6 Which Sampling Rate to Use When Recording Data

- Times per second the data are acquired
- Defines data's temporal resolution
- Depends on:
  - Type of analysis
  - Frequencies to analyze
  - Available desk space
  - Processor Speed
- Nyquist Theorem
  - Only frequencies below half the sampling rate can be recovered
    - Looking at 50Hz, need to sample 100Hz
- Generally use (500Hz to 2000Hz) sampling rate
  - Better higher sampling rate and then downsizing
- Make it easy to convert between time in samples and time in ms
  - 1000Hz is optimal
    - 14ms is 14 samples
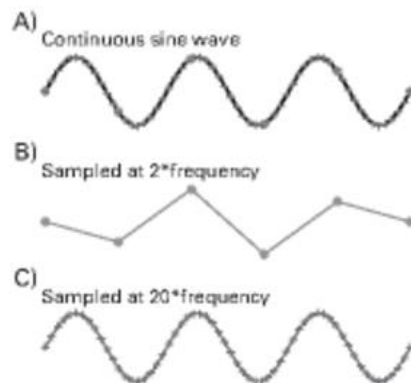
Figure 6.2

A continuous sine wave (panel A) and an illustration of the effect of subsampling that sine wave. Panel B shows that sampling the sine wave at twice its frequency (see gray dots along the sine wave in panel A) can reconstruct some features of the sine wave but fails to reconstruct the finer features, in particular the precise peak and trough times and the ongoing phases. Panel C shows that sampling at 20 times the frequency (see gray plus signs in panel A) can reconstruct the time-varying features of the sine wave with much higher accuracy.

6.7 Other Optional Equipment to Consider

- Response EMG or Force Grips
  - Provide data on muscular movement
- Eyetracker
  - Lets you remove trials where subjects looked away from the fixation spot
  - Use saccades and looking times as dependent measures
  - Remove oculomotor artifacts from EEG data
  - Changes in pupil dilation

- Electrode Localization Equipment
  - Have the precise location of the electrodes on the dead
- Comfortable Chair for Subject to Sit In
  - More comfort = less movement artifacts
- Good Response Device
  - Response device with good timing, comfortable, and intuitive
  - Easy to hold, intuitive layout
  - Not too easy to press (subjects must know that response was registered)
  - Not too hard to press (subject gets tired)
  - Clicks to show that a response was registered
    - Might also create artifacts

7 Preprocessing Steps Necessary and Useful for Advanced Data Analysis

7.1   What is Preprocessing?

- Any processing between collecting and analyzing the data
  - o Organize data
    - ▪ Extract epochs from continuous data
  - o Removing bad or artifact-ridden data w/ out changing clean data
    - ▪ Remove bad electrodes
    - ▪ Reject epochs with artifacts
  - o Modifying Clean Data
    - ▪ Temporal filters
    - ▪ Spatial transformations
- Keep track of all details of preprocessing for each subject
  - o Trials rejected
  - o Electrodes interpolated
  - o Independent components removed from the data

7.2   Balance Between Signal and Noise

- Signal and noise overlap, so it is hard to remove noise without removing signal too
- Noise example:
  - o Amplifier Saturations (produce spikes)
- Threshold for noise depends on the experiment at hand

7.3   Creating Epochs

- Continuous data are cut into segments surrounding particular experimental events
- Must decide what to call "time = 0"
  - o Options:
    - ▪ Time-lock to earliest event in each trial
    - ▪ Time-lock to the data of focus
- Must decide how much time before and after the "time = 0" event
  - o Epoch must be at least as long as duration of trial
- Compute only ERPs
  - o Epoch as long as time period to analyze + a baseline period
- Time-Frequency-Based Analysis
  - o Longer epochs to avoid contaminating results with edge artifacts
    - ▪ Edge artifact:
      - • Apply temporal filters to sharp edges → producing a high-amplitude broadband power artifact lasting hundreds of ms
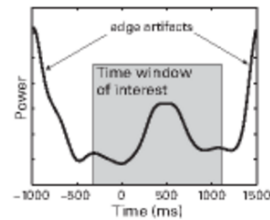
Figure 7.2
Edge artifacts resulting from discontinuous breaks in the time series between trials can contaminate the results if there are insufficient buffer zones to allow those edge artifacts to subside. In this case the edge artifacts are easily identifiable, and it is also clear that those artifacts subside before the time window of interest (gray area). In general, edge artifacts will contaminate up to three cycles of activity, but this could be less or more depending on the magnitude of the edges.

- The buffer zone you choose to include depends on frequencies you intend to extract
  - Longer epochs allow for edge artifacts to subside before and after experiment events
  - Lower frequency band to extract = more buffer zone to avoid edge artifacts
- Analyze one subject closely
  - Check if artifacts affect time period of interest
- General Rule:
  - 3 Cycles at the lowest frequency
  - Ex. 1500 ms for 2Hz activity
- Caveat: Large Epochs
  - Overlapping data in each epoch
  - A problem for Independent Component Analysis
    - Don't expose ICA to the same data more than once
- Sufficient Buffer Required for:
  - Time-Frequency-Decomposition via:
    - Complex Morlet Wavelet Convolution
    - Filter-Hilbert Method
- Sufficient Buffer Not-Required:
  - Time-Frequency Decomposition via:
    - Short-Time FFT
    - Multitaper
- If analyzing data that's already EEG epoched
  - Use Reflection – only when necessary
    - EEG data from each trial and electrode are reversed and put in beginning and end of trial
  - Makes Epoch 3 times longer
  - Discard reflected data after analysis
- NEVER – Taper the entire epoch time period

7.4 Matching Trial Count across Conditions

- Ideal for all conditions to have the same number of trials
- Analysis backed on phase:
  - More sensitive to trial count
  - Small # of trials introduces positive bias
- Analysis based on power or ERP
  - Less sensitive to trial count
  - Low trials ERP
    - Mean amplitude in time range is more robust to noise compared to peak times
- Large Differences in Trial Count Across Conditions
  - (less than 30 trials) – consider matching trial count across conditions
- Matching Trial Count
  - Identify condition with fewest trials
  - Selecting trials from other conditions
    - All conditions end up with equal # of trials
  - Method 1:
    - Select 1st N trials from each condition
    - N is # of trials in smallest condition
    - Biases conditions to have more trials earlier in experiment (when subjects are less tired, more patience/motivated)
  - Method 2:
    - Select trials at random
    - No bias in terms of when trial occurred in experiment
    - But reanalyzes same data multiple times (unless store which trials were used already)
  - Method 3:
    - Select trials based on relevant behavioral or experimental variable (ex. Reaction time)
    - Select subset of trials from all conditions
      - Distributions of reaction times from the retained trials are similar across conditions
    - Disadvantage:
      - If there are reaction time differences between conditions, matching reaction times across conditions may bias trial selection from different regions of the reaction time trial distribution
    - Selective sampling based on any relevant behavioral measure:
      - Saccade speed
      - Pupil response
      - Subject difficulty rating
      - Luminance
      - Location of stimulus
- Trial matching when comparing EEG results across subjects with a behavioral variable that might be related to trial count

7.5 Filtering

- Notch filter at 50Hz or 60Hz
    - Attenuate electric line noise
- Not always necessary
    - Ex.
        - Don't need to low-pass filter time-domain data at 40Hz if perform time-frequency analysis to extract power from 2 – 20 Hz
- High Pass Filters
    - Only applied to continuous data (NOT EPOCH DATA)
        - Since edge artifacts of 0.5Hz may last up to 6 seconds (longer than epochs)
    - High-pass filter at 0.1 or 0.5Hz will minimize slow drifts

7.6 Trial Rejection

- Automatic rejection procedures
    - Fast
    - No user bias
    - Same trials rejected every time
    - Cons:
        - May use criteria only appropriate for some subjects but not others
- Time-frequency Decomposition
    - Sharp edges are more detrimental in TFD than ERPs
    - Small edges may not be picked up by automatic rejection procedures
        - Will have huge effect on TFD results

7.7 Spatial Filtering

- 1. Help to localize results
    - Ex. Confirm activity peak corresponds to left motor cortex
        - Surface Laplacian or fit a single dipole
- 2. Isolate topographical feature of data filtering low-spatial-frequency features
    - Ex. Visual stimuli that requires spatial attention
        - Difficult to separate visual processing (occipital cortex) from
        - Attention-related processing (parietal cortex)
        - Surface Laplacian (or distributed source imaging)
            - Minimize spatial overlap between brain regions
                - Increasing confidence in functional / anatomical distinctions
- 3. Preprocessing for connectivity analyses
    - Surface Laplacin (or distributed source imaging)
        - Minimize volume conduction (artifact contaminating connectivity analyses)
- When to use spatial filtering?
    - ERP on response preparation

- Is ERP peak consistent with source in motor cortex?
  - Single dipole fit to grand-averaging ERP
- Surface Laplacian to minimize volume conduction for connectivity analysis?
  - Laplacian applied to single-trial time-domain data before time-freq. decomposition
- PCA on time-freq. power
  - PCA on single trials within subject
  - Or trial-averaged power across subjects

7.8 Referencing

- Only issue with EEG
  - Voltage values recorded from each electrode are relative to voltage values recorded elsewhere
  - Any activity in reference electrode is reflected as activity in all other electrodes
- Average mastoid (bone behind ear)
  - Reference electrode should not be close to brain region you expect to be activated
- Data can be re-referenced online because is a linear transformation
- Bipolar reference
  - One electrode is measured relative to another
  - Common measuring the eye
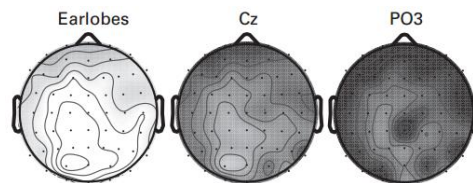
Figure 7.4
The effect of different reference electrodes on the same data. Earlobes refers to the average of electrodes placed on the two earlobes. In many situations, using one of the scalp electrodes as the reference is suboptimal.

7.9 Interpolating Bad Electrodes

- Data from missing electrodes are estimated based on the ativity and locations of other electrodes
  - Weighted Distance Metrics usually used include:
    - Nearest-neighbor
    - Linear
    - Spline
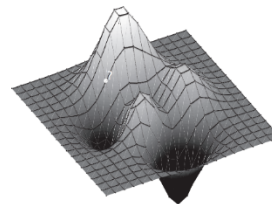
Figure 7.5
Topographical illustration of interpolation. Displayed is a smooth topographical landscape (analogous to a scalp-measured voltage) that is discretely sampled (black dots; analogous to electrodes). Interpolation involves estimating the activity at the white "electrode," given the activities and distances of all other electrodes. This topography was generated with the Matlab   function.

- Interpolation is perfect weighted sum of all other electrodes
  - May cause problems with analysis requiring matrix inverses
- Another Option: Remove them altogether
- Interpolation important for spatial filters:
  - Surface Laplacin
  - Source reconstruction
  - Spherical Spline
  - Re-reference to average of all electrodes
    - "Activity of one bad electrode may contaminate signal of other electrodes"
- Do I interpolate?
  - 1. Inspect data
    - If a lot of noise, try filtering noise without interpolating
  - Can this data be salvaged?
    - Low pass filter of 30 Hz
    - If activity looks similar to surrounding electrodes, it can be salvaged. No interpolation.
    - If activity looks **different** (much smaller / larger magnitude) → unlikely real signals were recorded. Should be interpolated.

7.10 Start with Clean Data

- Noisy data is crap and won't give you anything at all

8 EEG Artifacts: Their Detection, Influence, and Removal

- Main EEG Artifacts
  - Blinks
  - Muscle movements
  - Amplifier Saturations
  - Line noise
  - Cognitive Artifacts

8.1 Removing Data Based on ICA

- Decomposes EEG time series into components
  - Components try to identify independent sources of variance
  - Set of weights for all electrodes so each component is a weighted sum of activity at all electrodes
    - Weights designed to isolate sources of brain activity
  - Analogy:
    - Cocktail party where everyone has a microphone. Tuning your system to dull the input of some microphones more than others to hear whom you want
- ICA
  - Clean EEG Data
    - Identify components that isolate artifacts
      - Artifacts found by:
        - Topographies
        - Time courses
        - Frequency Spectra
      - Take conservative stance when removing artifacts of noise
        - Don't want to be removing part of single in the process
        - Best case scenario, Only remove component regarding blink artifacts
  - Data Reduction
    - Analyze component time series (instead of electrode TS)
      - Max number of components = # of electrodes used
    - If Have 100+ Electrodes
      - Extract fewer components than electrodes
        - Faster analysis
        - Unlikely to be over 100 independent sources in the brain that are active AND can be statistically isolated with EEG

A) Topographical map of component weights

B) One trial of component time course
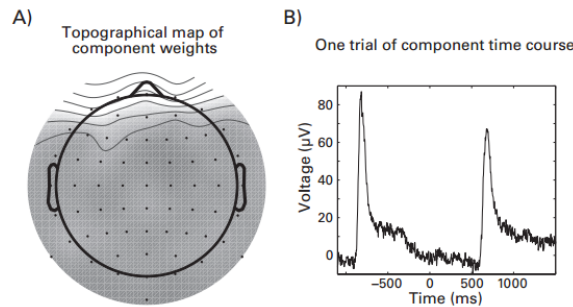
Figure 8.1

Example topographical map and example single-trial time course of an independent component that isolated a blink artifact. Panel A shows that the weights from this component are maximal at anterior electrodes, and panel B shows the time course of this component from one trial. You can see that on this trial, the subject blinked before and after the trial. The time course is a weighted sum of the activity of all electrodes, and the weights are defined by the results of the independent components analysis.

8.2 Removing Trials because of Blinks

- Blink artifacts linearly sum on top of the brain-generated EEG
- Removed with:
  - ICA (Jung et al. 2000)
    - Appears to work better
  - Regression-based Techniques (Graton, Coles, and DOnchin 1983)
- Other reasons to remove trials with blinks
  - Multi-millisecond blinks can indicate subject is tired and cannot focus on task
  - Poorly timed blink means subject didn't see visual stimulus
- Asking people to refrain from blinking
  - Suppression of blinking is a demanding task (creates other signals)
    - Task-unrelated but stimulus-locked activity in frontoparietal oculomotor circuits
    - Distracted by inhibiting blinks and unfocused on stimuli
    - If subjects can blink during intertrial interval, the time period may not be useful
      - For baseline for normalization of time-frequency
- Best bet
  - Ask subjects to avoid time-locking blinks with experimental events (pressing button)

8.3 Removing Trials because of Oculomotor Activity

- Saccades and microsaccades also contaminate data
  - Frontal and lateral frontal electrodes
  - Minimized by having centered fixation point on monitor
    - Tell subjects that eye movement will impact data
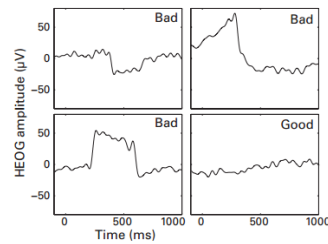- Eye trackers or vert / horiz. EOG electodes

Figure 8.2
Horizontal EOG activity indicates eye movements after stimulus onset (time = 0). The first three panels are taken from trials that were removed prior to analyses.

- Microsaccades minimized by:
  - Small stimuli so subject does not need to saccade to see entire stimulus
  - Display on screen for short time
- Reference impact on data
  - Nose reference → Eye movement has larger influence over earlobe reference
- Concern of hypothesis
  - Anterior fontal or lateral frontal regions
    - EOG artifacts are a concern
  - Midcentral electrodes
    - EOG artifacts are less of a concern
- Isolate potential EOG artifacts
  - Spatial filtering techniques
- For tasks where subjects should have been fixating on point
  - Eye movement indicates not fully engaged in task

8.4 Removing Trials Based on EMG in EEG Channels

- EMG → bursts of 20- to 40-Hz activity
    - Maximal in electrodes around face, neck, ears
    - Deleterious for EEG data analyzed above 15Hz
- If EMG has constant amplitude before / after trial onset
    - EMG removed during baseline normalization
- IF EMG present in all conditions
    - Subtracted out during condition comparisons

8.5 Removing Trials Based on Task Performance

- Cognitive noise (artifacts)
- If trial responses can be accurate or not
    - Remove (or separate) error trials
- Also remove trials where:
    - Subjects don't make a response if they were instructed to do so
    - Trials with more responses than were required
    - Trials with fast reaction times
    - Trials with very slow reaction times
        - Slower than 3 standard deviations from each subject's median reaction time
    - If rest break lasts several tens of seconds, subjects not fully engaged on first trial after break
    - Subjects perform a few tens of trials with one set of instructions, then switch to another set of instructions

8.6 Removing Trials Based on Response Hand EMG

- EMG from muscles subjects use to indicate response
- Helps identify partial errors
    - Subject twitches muscle of incorrect response
    - Brain activity looks like that of an error even though response was correct
- Algorithm for partial error identification
    - Z-transform of derivative of EMG Signal from each hand
    - Rectified (absolute value)
    - **Partial Error When:**
        - Z-derivative signal of hand not used exceeds 2 standard deviations in the time between stimulus onset and actual button press
        - Mag. Of EMG peak must be >2x larger than largest EMG peak from -300ms to stimulus onset
- Record EMG from thumb muscles
    - Have response button that takes some effort to press
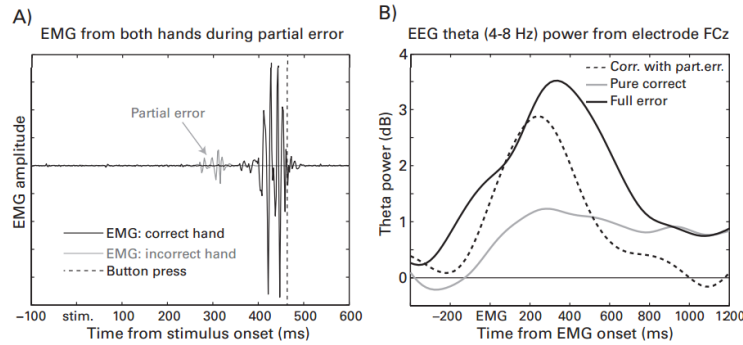   - Requires more muscle engagement → bigger, cleaner EMG response

Figure 8.4
Partial errors can be detected with EMG recordings and are useful to identify correct trials with error-like brain responses. Panel A shows example EMGs from right and left thumbs showing a partial error—a muscle twitch of the incorrect hand although only the correct button was pressed in this trial. Correct trials that contain partial errors elicit error-like brain processes, as can be seen in panel B. Theta-band power from electrode FCz is time-locked to EMG onset; each line is the average of all trials from one subject. Partial errors can be identified and removed from the dataset or separately analyzed.

8.7 Train Subjects to Minimize Artifacts

- Show subject EEG data in real time on computer
- Explain: EEG Has brain activity and noise from muscles
- Have them: blink, clench jaw, tense neck/shoulder muscles, talk, smile, wiggle ears, etc.
    - When subject knows behaviors that produce artifacts, they can reduce them

8.8 Minimize Artifacts during Data Collection

- Keep an eye on real-time EEG data
    - Check every 30 seconds to see if they look ok
    - If data doesn't look great, pause and do what you can to fix it

9 Overview of Time-Domain EEG Analysis

9.1 Event Related Potentials (ERPS)

- To create ERP
    - Align time-domain EEG to time = 0
    - Average across trials at each time point
    - Sum voltage at each time point across trials
        - Divide by number of trials
- ERP as quality inspection tool: ^^^ is all you need
- ERP to make inferences about cognitive processes
    - Learn about:
        - Component overlap

- Component quantification
- Appropriate Interpretation
- Statistical procedures

9.2 Filtering ERPs
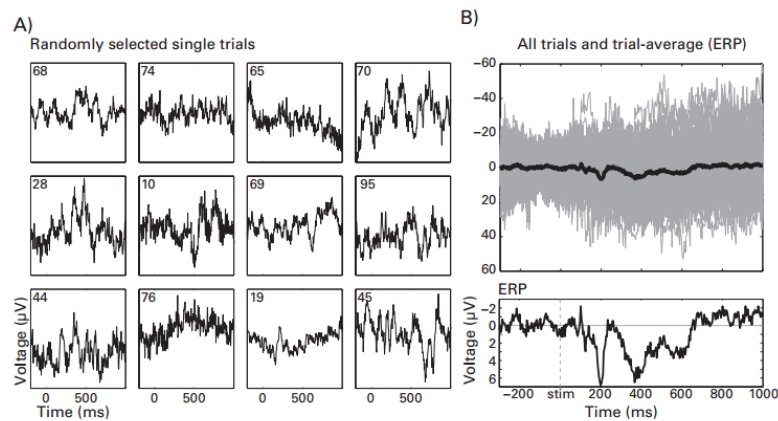
Figure 9.1
Panel A shows single-trial EEG traces from 12 randomly selected trials (number inside plot indicates trial number). Data are from electrode FCz. Panel B shows 99 single trials in gray and their average—the ERP—in black. Panel C shows the same ERP with focused *y*-axis scaling.

- Non-phase-locked activity doesn't survive time-domain averaging
  - Freq. above 15Hz usually not time locked
- Filtering ERPs:
  - Minimizes residual high frequency fluctuations
  - Makes ERPs look smoother
  - Facilitates peak-based component quantification
    - Reduces possibility that peak is a noise spike or non-representative outlier
- Poorly designed filters introduce ripples
  - Resulting from filters with narrow transition zones
- Applying low pass filter reduces temporal precision
  - Since voltage at each time point becomes a weighted average
  - Lower cutoff frequency of filter = more loss in temporal precision
- ERPs often filtered using frequency cutoff ~20 or 30Hz. Occasionally 5 or 10 Hz
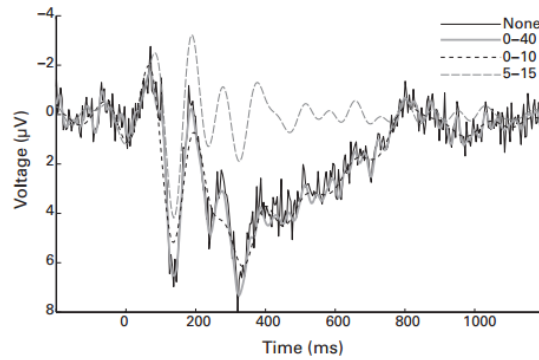
Figure 9.2
An ERP from electrode P7 with no filtering (black line) and with different filter settings (numbers in the legend indicate lower and upper frequency bounds in hertz). Note that some filter settings can have dramatic effects on the interpretation of specific ERP features. For example, the 5–15 Hz filter seems to have accentuated the first negative-going peak at around 100 ms and removed the later P3-type component, and the 0–10 Hz filter removed the negative-going peak at around 280 ms. The wide-band 0–40 Hz filter had the least effect on the larger ERP fluctuations while removing the high-frequency fluctuations. This plot is an illustration of why you should carefully consider the frequency range of the filter used for interpreting ERPs, particularly if you use a narrow frequency range.

9.3 Butterfly Plots and Global Field Power / Topographical Variance Plots

- ERP from all electrodes overlaid in same figure
- Global field power = standard deviation of activity over all electrodes at each point in time
- Useful as data quality indices + confirm timing of representations of task events in the data
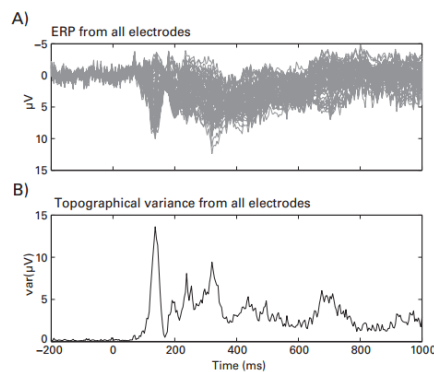
Figure 9.3
An example butterfly plot (panel A) and a topographical variance plot (panel B). Although they lack spatial information, these plots are useful for data inspection and provide an overview of the time periods with cortically diverse events, including approximately 180 ms, 220 ms, 320 ms, and 700 ms.

9.4 The Flicker Event

- Entrainment of brain activity to a rhythmic extrinsic driving factor
  - Ex. Looking at a strobe light flickering at 20 Hz → activity in visual cortex at 20Hz
- Also called:
  - Steady state evoked potential / Frequency Tagging / SSVEP / SSAEP
- Flicker effect allows you to "fake" high spatial resolution

- o   However, has poor temporal precision
- Stimulus flicker frequencies up to 100Hz can evoke a flicker effect
  - o   Lower frequencies generally elicit a stronger flicker effect
  - o   Size, luminance, and contract impact signal – to – noise ratio
- Generally: peak in frequency domain at flicker frequency should be readily observed in plot
  - o   Magnitude of frequency peak can be compared to same frequency before flickering stimulus began
  - o   Or compared to power of neighboring frequencies for which there was no flicker

9.5   Topographical Maps

- Show spatial distribution of EEG results
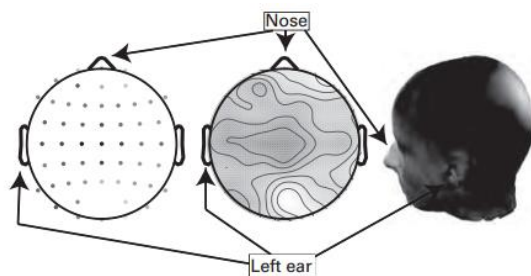
Figure 9.4
The same data shown by coloring dots on a topographical map (left) or interpolating those values over a surface (middle and right). Clearly, the interpolated maps are easier to interpret. Two-dimensional plots show data from all or most electrodes; 3-D plots (right) look more realistic but obscure data from part of the scalp. Locations of the nose and left ear are indicated to facilitate orientation and comparison.

- 2-D plots are less intuitive to interpret
  - o   Show activity simultaneously from all (or most) electrodes
- 3-D plots show activity from only a third of the head
  - o   Easier to interpret and look nicer in figures
- Allow you to:
  - o   Confirm timing of task events
  - o   Detect bad or noisy electrodes
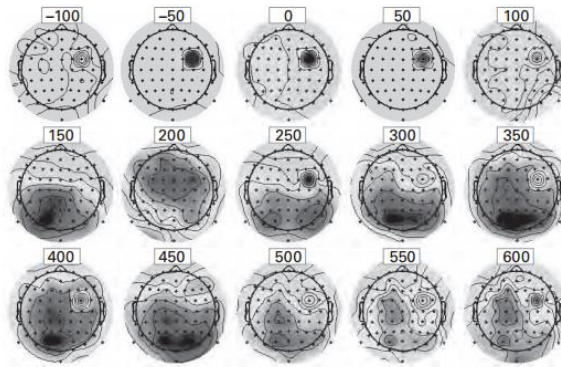
Figure 9.5 (plate 1)
Plotting topographical maps over time facilitates rapid data quality inspection. The numbers in white boxes indicate the latency at which the topographical data are plotted (in milliseconds) with respect to trial onset. These plots show, among other things, that there is one bad electrode. In this case the bad electrode was generated by replacing the true EEG activity at electrode FC4 with randomly generated numbers.

9.6   Microstates

- Temporal difference (global map dissimilarity) remains low for a period of time of tens to hundreds of milliseconds, then suddenly becomes relatively large
  - Sharp increase is a shift of microstates
- Stable maps used in hierarchical clustering analysis to identify small number of topographical maps during periods of stability (cluster maps)
  - Topography at each time point is labeled according to cluster map to which it is most similar
- Produces time course of map topographies to be used in task related and statistical analyses

9.7   ERP Images

- 2-D representation of EEG Data from single electrode
- Useful for:
  - Single-subject data inspection tool
  - Linking trial-varying task parameters or behaviors to the time-domain EEG signal
    - Sort EEG trials according to values of aligning event (reaction time or phase of a frequency-band-specific signal at certain time point)
- Images smoothed by convolving image with 2-D Gaussian
- ERP Also made for:
  - Frequency-band-specific activity (filtered signal, power, or phase)
  - Time-frequency plots because time is on x-axis and data are colored
    - Y-axis has no frequency information
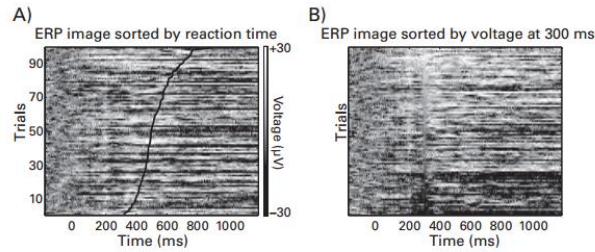
Figure 9.6
Example ERP images, both using data from electrode FCz. Time is on the *x*-axis, trials are on the *y*-axis, and the grayscale intensity (or color if you create this figure on your computer) corresponds to the EEG voltage values over time and trials. In panel A the trials are re-sorted according to the reaction time on each trial, and in panel B the trials are re-sorted according to the voltage value at 300 ms. The black line in panel A corresponds to the reaction time on each trial.

10 The Dot Product and Convolution

- Extracting time-varying, frequency-band-specific info from EEG Data

10.1 Dot Product

- Sum of elements in one vector weighted by elements of another vector
  - (signal-processing interpretation)
- Covariance or similarity between two vectors
  - (statistical interpretation)
- Mapping between vectors
  - Product of magnitudes of two vectors scaled by cosine of angle between them
  - (Geometric interpretation)
- To compute dot product:
  - Multiply each element in one vector by corresponding element in other vector
  - (First element in vector A x First element in vector B)
    - For all elements
  - All of the points are then summed
- Can conceptualize an EEG signal with 640 time points as a single point in a 640 dimensional space
  - Location of that point is defined by values of EEG signal at each time point
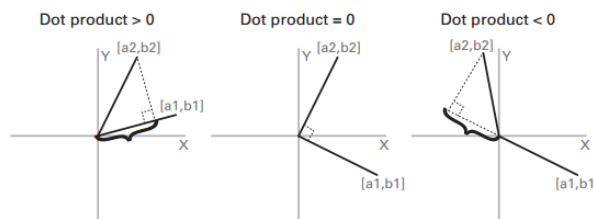
Figure 10.1
Graphical illustration of the geometric interpretation of the dot product between two two-element vectors. Curly brackets illustrate the magnitude of the projection of one vector onto the other (this is the dot product).

10.2 Convolution

- Dot product computed repeatedly over time or space
- Time series of one signal weighted by another signal that slides along the first signal
    - (Signal Processing Interpretation)
- Cross-covariance: similarity between two vectors over time
    - (Statistical Interpretation)
- Time series of mappings between two vectors
    - (Geometric Interpretation)
- Frequency Filter
- One vector is considered the signal (EEG Data) → other vector considered the kernel (Wavelet or sine wave)

10.3 How Does Convolution Work?
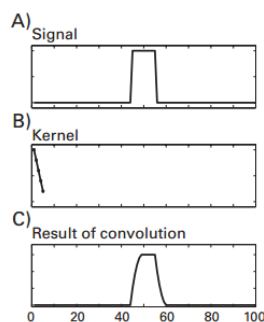
Figure 10.2
Example of convolution between two vectors, one labeled the signal (panel A) and one labeled the kernel (panel B). Panel C shows the result of the convolution between these two vectors.

- 1. Flip the kernel backward and line it up with the data
    - Compute dot product by multiplying each point in the kernel by each corresponding point in the data. Sum over all multiplications.
    - Place in a new vector in the position corresponding to the center of the kernel
        - Convenient to use kernel with odd number of data points so there's a center
- 2. Kernel is shifted to right by one time step. Signal does not move. Repeat dot product process and store result in position corresponding to center of kernel (now 1 time step to right)
    - Repeat until each end of data

- Begin convolution with time kernel to the far left. The right side of the (reversed) kernel overlaps with the leftmost points of the data
- (Length of kernel – 1) # of zeros added before kernel
    - After running convolution, trim results by removing ½ of the length of kernel from beginning and ½ from the end

Figure 10.4
Overlap of results of manual time-domain convolution and the Matlab frequency-domain convolution function conv. Also shown is the result of convolution after normalization, which is not necessary when the sum of the kernel points is zero, as is the case with wavelets.
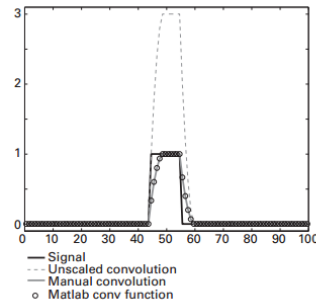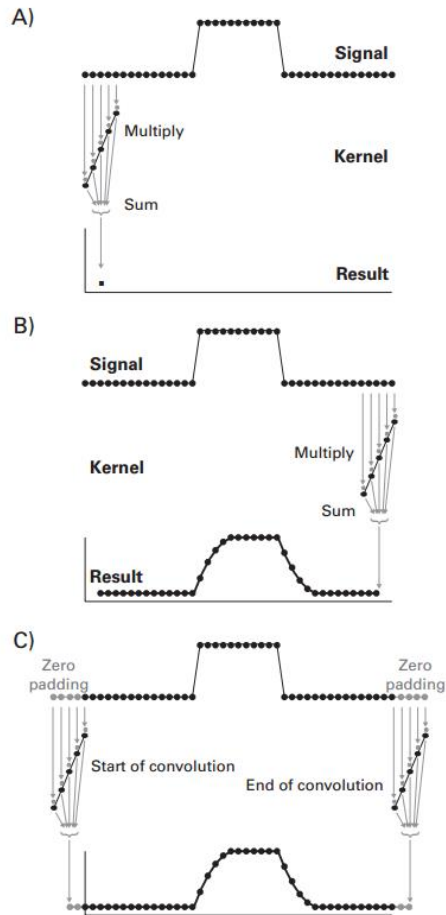
- To compare result of convolution with original signal → scale result of convolution by sum of kernel points
  - o Note: post-convolution scaling is not same as mean-centering kernel before convolution

10.4 Convolution vs. Cross-Covariance

- Cross-correlation = cross-covariance scaled by variances
- Convolution → Kernel is reversed
  - o Cross-covariance → Kernel kept in original orientation
- If convolution kernel is temporally symmetric
  - o Convolution and cross-covariance yield identical results

10.5 Purpose of Convolution for EEG Data Analyses

- Isolate frequency-band-specific-activity
- Localize frequency-band-specific activity in time
  - o As wavelet is dragged along EEG data, it reveals when and to what extent the EEG data contains features that look like the wavelet

- When convolution repeated on same data using wavelets of different frequencies, a time-frequency representation can be formed

11 Discrete Time Fourier Transform, the FFT, and the Convolution Theorem

- Computes dot product between the signal (EEG data) and sine waves of different frequencies (kernels)
- 3 Characteristics of Sine Waves:
  - Frequency – how fast (Cycles per second – Hz)
  - Power – amplitude squared
  - Phase – timing of the sine wave (radians or degrees)

11.1 Making Waves

- Sine waves are: A $\sin(2\pi ft + \theta)$
  - Amplitude
  - f = frequency of sine wave
  - t = time
  - θ = phase angle offset
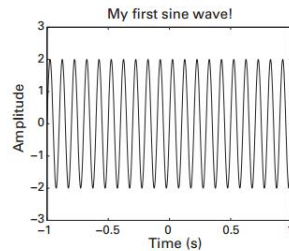
Figure 11.1
A sine wave that was created using the expression $2\sin(2\pi 10t + 0)$ (thus, amplitude of 2 and frequency of 10 Hz).
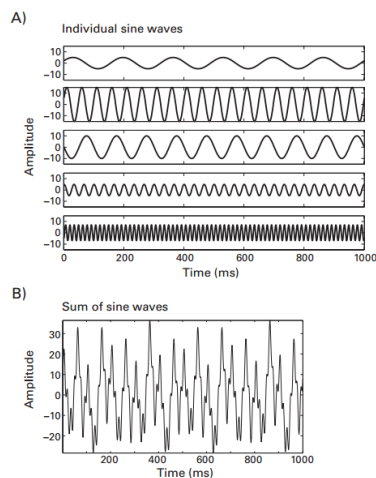
Figure 11.2
Several sine waves of differing amplitudes, frequencies, and phases, plotted separately (panel A) and after being added together (panel B).

11.2 Finding Waves in EEG Data with the Fourier Transform

- Have the time series

- o Want to know which sine waves with which frequencies, amplitudes, and phases will reconstruct the time series
- Fourier transform = computes dot product between sine waves of different frequencies and the EEG data
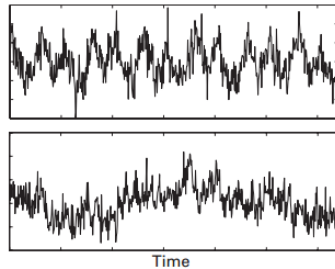
Figure 11.3
One of these time series was generated by summing different sine waves and adding noise; the other time series is real EEG data. Can you guess which is which? What evidence do you use to support your decision?

11.3 The Discrete Time Fourier Transform

```
N = 10; % length of sequence
data = rand(1,N); % random numbers
% initialize Fourier coefficients
fourier = zeros(size(data));
time = (0:N-1)/N; % time starts at 0; dividing by N normalizes to 1
% Fourier transform
for fi=1:N
    % create sine wave
    sine_wave = exp(-1i*2*pi*(fi-1).*time);
    % compute dot product between sine wave and data
    fourier(fi) = sum(sine_wave.*data);
end
```

- 1. Sine wave is created
- 2. Computes dot product between sine wave and the data
  - o Number of unique frequencies that can be extracted from a time-series is one half the number of data points in that time series, plus the 0 frequency
- Discrete Fourier Transform
  - o $X_f = \sum_{k=1}^{n} x(i)\, e^{-12\pi f(k-1)n^{-1}}$
    - n = # of data points in vector x
    - X_f = Fourier coefficient of time series variable x at frequency f

11.4 Visualizing the Results of a Fourier Transformation

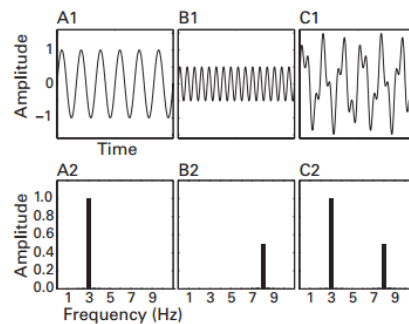- 2D plot with frequency on the x-axis and power (or amplitude) on the y-axis

Figure 11.4

Two sine waves of 3 Hz and 8 Hz (panels A1 and B1), and their sum (panel C1) have a Fourier power spectrum that can be represented in plots of frequency (x-axis) by amplitude (y-axis) (see plots in the bottom row). The frequency of the sine wave corresponds to the position on the x-axis, and the amplitude of the sine wave corresponds to the position on the y-axis.

- 3D plot representing frequency-power-phase space
  - Phase = position of sine wave at each frequency when it crosses time = 0
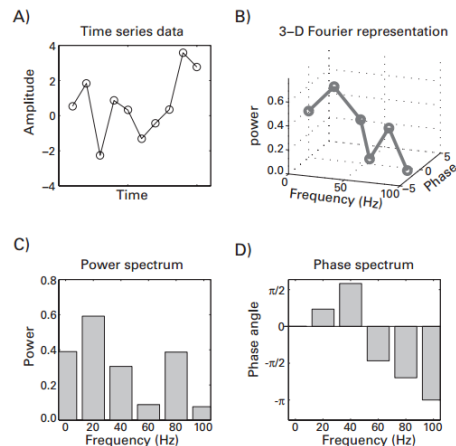
Figure 11.5

Panel A shows a time series of randomly generated data (sampled at 200 Hz). Panel B shows a frequency representation of the data as a line through 3-D Fourier space (frequency, power, and phase; see axis labels). Panels C and D show the projections of the 3-D Fourier result onto two dimensions at a time. Panel C shows only the frequency and power information from the Fourier transform, and panel D shows only the frequency and the phase information from the Fourier transform.

11.5 Complex Results and Negative Frequencies

- Half of the Fourier series has "positive" frequencies, and half has "negative"
  - Negative = capture sine waves that travel in reverse order around the complex plane compared to those that travel forward
- For a signal with only real numbers, the negative frequencies mirror the positive frequencies, so you can ignore them
  - Instead just double amplitude of positive frequencies
  - Don't remove negative frequencies (necessary for inverse Fourier transform)

11.6 Inverse Fourier transform

- Results of a Fourier transform in a series of coefficients that represent the dot product of each complex sine wave to the data
- Inverse Fourier transform:
    - Start with sine waves of different frequencies, amplitudes, and phases summed together to form a single time series
- To compute inverse Fourier transform:
    - Build sine waves of specific frequencies
    - Multiply them by respective Fourier coefficients at those frequencies
    - Sum all sine waves together
    - Divide by number of sine waves
- $X_f = \sum_{k=1}^{n} x(i) \, e^{i2\pi f(k-1)n^{-1}}$
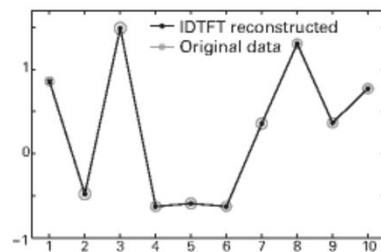
Figure 11.6
Time series data of randomly generated numbers and the reconstruction of the data using the inverse discrete time Fourier transform (IDTFT).

11.7 Fast Fourier Transform
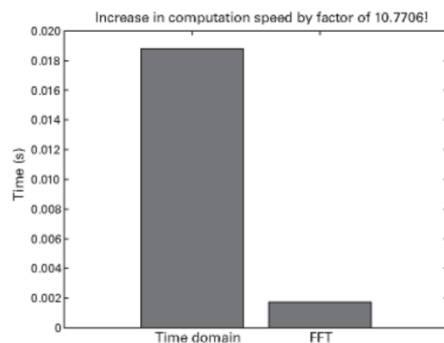
Figure 11.8
Computation times for the discrete time Fourier transform versus the FFT.
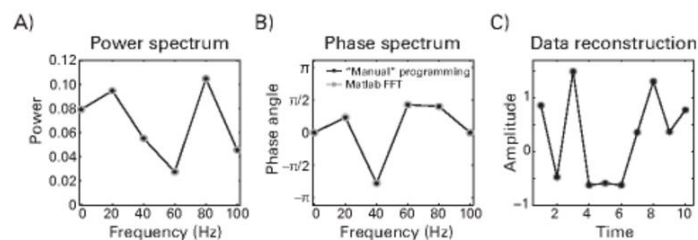

Figure 11.7
The fast Fourier transform (FFT) produces identical results as the discrete time Fourier transform and is much faster. (Note that due to very small computer rounding errors, you might see 0 or 2π, or $-\pi$ or $+\pi$, for the phase value at the Nyquist frequency.)

- With huge datasets, FFT is 1000x faster than FT

11.8 Stationarity and Fourier Transform

- Fourier transform assumption
    - Data is stationary – Data Elements Below Don't Change Over Time
        - Mean
        - Variance

- Frequency Structure
  - ^^^ Doesn't actually happen in real life so...
- Perform temporally localized frequency-decomposition
  - Wavelet Convolution
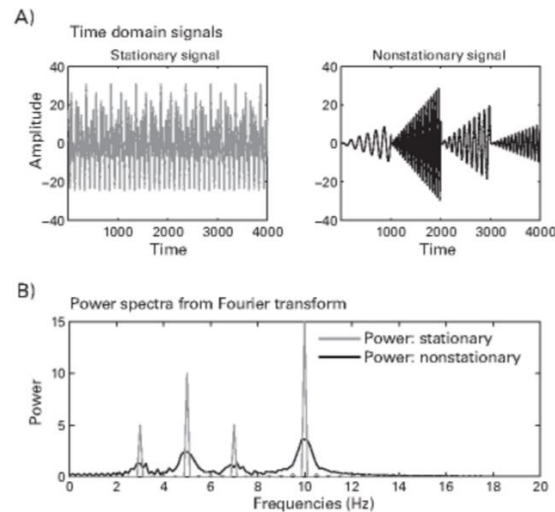 - Filter-Hilbert
  - Short-term FFT

Figure 11.9
Violations of stationarity (here introduced by changes in frequency and amplitude over time) result in energy at frequencies that were not explicitly generated when creating the data. Nonetheless, spectral peaks can clearly be observed for both the stationary and the nonstationary time series. Note that for the stationary time series, the power peaks appear to be carrot shaped, but this is due to using lines instead of bars to represent discretely sampled frequencies. If you look closely, you can also see small but non-zero power at many frequencies for the stationary time series; these are artifacts resulting from the sharp edges at the beginning and ends of the time series.

- Temporally localized methods:
  - Assume data are stationary within relatively brief periods of time (few hundred ms)
- Second reason to perform temporally localized frequency decomposition
  - Fourier transform doesn't show how dynamics change over time
    - Time varying changes in frequency structure cannot be observed directly in power or phase plots

11.9 Extracting More or Fewer Frequencies than Data Plots

- Number of frequencies from Fourier Transform
  - N/2 + 1.
  - N is the number of time points in the data
  - +1 is for DC or zero-frequency component
- Can increase frequency resolution by:
  - Adding zeros at end of time series → (Zero padding)
  - Increases quantities of frequency
  - Improves frequency resolution
  - DOES NOT improve frequency precision
- Zero padding can make Frequency-Domain Convolution more convenient + faster to perform

11.10 Convolution Theorem

- States that convolution in time domain is same as multiplication in frequency domain
  - o Therefore two means of performing convolution
- 1. Time-domain version of convolution
  - o Slow
  - o Flip kernel backward, slide it along signal, and compute dot product at each time step
- 2. Frequency-Domain Convolution
  - o Fast
  - o Fourier transforms of the signal and kernel
  - o Multiplying Fourier transforms together point-by-point (freq-by-freq)
  - o Taking inverse Fourier transform
  - o Result of multiplication (ie convolution) is frequency structure common to both kernel and signal → conceptualized as frequency-domain filter
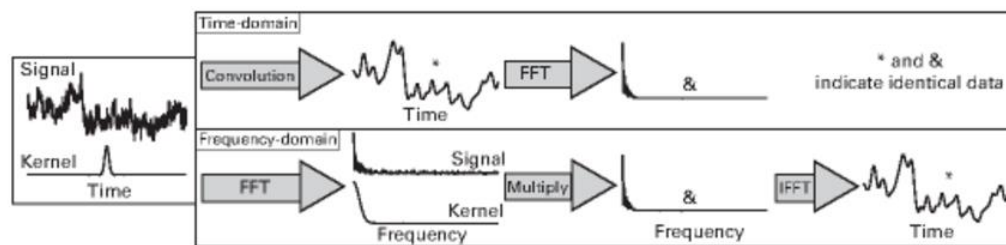
Figure 11.10

Illustration of the convolution theorem and the interchangeability of time-domain convolution and frequency-domain multiplication. The two time series with asterisks are identical, as are the two frequency spectra with ampersands.
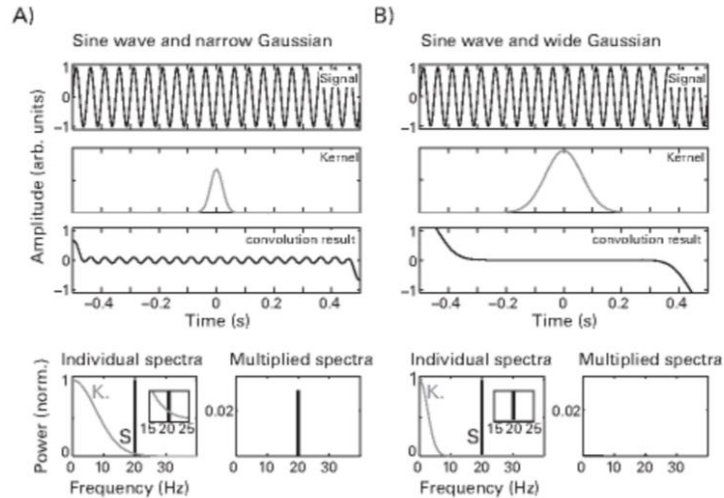
Figure 11.11

The convolution between a 20-Hz sine wave and a narrow Gaussian (panel A) dampens the sine wave, whereas the convolution between the same sine wave and a wide Gaussian (panel B) obliterates the sine wave. This is because the power spectrum of the sine wave (bottom row) overlaps slightly with the power spectrum of the narrow Gaussian at 20 Hz, but the power spectrum of the sine wave does not overlap with the power spectrum of the wide Gaussian. The gray line corresponds to the kernel (K), and the black bar corresponds to the signal (S). The insets in the power plots highlight the overlap (or lack thereof) between the frequency representations of the Gaussians and the frequency representation of the sine wave. The power spectra were normalized to 1 to facilitate visual comparison. The sharp rise and drop at the beginning and end of the result of convolution are edge artifacts.

- Power spectrum of narrow Gaussian has nonzero values that overlap with nonzero values of power spectrum of the 20Hz sine wave
- Power spectrum of wide Gaussian is zero at frequencies where power spectrum of sine wave is nonzero
  - Resulting freq.-spectrum multiplications produce:
    - Narrow Gaussian: amplitude-attenuated sine wave
    - Wide Gaussian: No sine wave
- BASIS OF WAVELET CONVOLUTION
  - Pass EEG data through set of filters (wavelets) tuned for specific frequencies
  - Result is frequency-band intersection between EEG data and wavelet
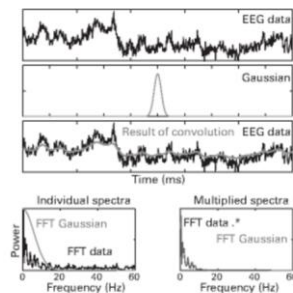

Figure 11.12

Convolving an EEG time series from one trial with a Gaussian low-pass-filters the data. This results from the frequency spectrum of the Gaussian tapering the higher frequencies in the EEG data. Note that this example here is meant for illustration of the convolution theorem and how convolution acts as a frequency filter; convolution with a Gaussian is not necessarily the best method for filtering EEG data. Chapter 14 contains more in-depth discussions of how to construct and apply bandpass filters to EEG data. The y-axis scaling of the power spectra is arbitrary to improve visibility.

- Gaussian kernel is a low-pass filter

11.11 Tips for Performing FFT-Based Convolution in Matlab

- Won't Produce Valid Convolution

```
result = ifft(fft(signal) .* fft(kernel));
```

- Result of convolution must be = length of signal + length of kernel – 1
- After computing inverse Fourier transformation →
    - Remove appropriate # of time points from beginning and end of time series
- Double-Check by Trying:

```
result = conv(signal,kernel,'same');
```

- Full instead of same option → length of the result of the convolution
    - Ie. Length of signal + length of kernel + 1

12 Morlet Wavelets and Wavelet Convolution

- Fourier transform
  - Gives frequency-domain representation of EEG data
  - Limitations:
    - Changes in frequency structure over time are difficult to visualize
    - EEG data violate the stationarity assumption of Fourier analysis
  - Because of limitations, apply:
- Time resolved frequency decomposition representations
  - (ie. Time-frequency representations)
  - Retains advantages of time and frequency domain
    - Sacrifices little in temporal and frequency precision
- Morlet wavelet (or Gabor wavelet)
  - A sine wave in middle, but tapers off to zero at both ends
  - Useful for localizing changes in frequency characteristics over time
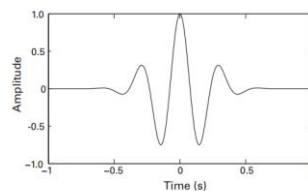
Figure 12.1
A Morlet wavelet, which is created by windowing a sine wave by a Gaussian.

12.1 Why Wavelets?

- Fourier transform doesn't show how frequency changes over time
  - Because kernel used has no temporal localization
  - (Amplitude of sine wave continues to fluctuate over its entire time series)
- To obtain temporally localized frequency information
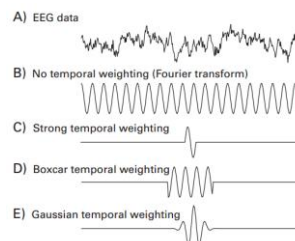 - Dot product should only be computed with part of sine wave in specific time window

Figure 12.2
In order to extract time-varying frequency-specific information from EEG data (A), the data must be convolved with a sine wave. Without tapering of the sine wave (B), the result reflects frequency-specific information from the entire time series. Use of only one cycle (C) maximizes temporal precision but at the expense of frequency precision. A uniform boxcar tapering (D) is suboptimal because of decreased temporal specificity and potential artifacts from sharp edges. A Gaussian tapering (E) (also known as a Morlet wavelet) provides an adequate balance between temporal and frequency precision without introducing edge artifacts.

- Best option is to use Gaussian taper to window the sine wave (Morlet Wavelet)
  - Has no sharp edges that produce artifacts
  - Dampen influence of surrounding time points on estimate of frequency characteristics at time point
  - Allow you to control trade-off between temporal and frequency precision
- Many types of wavelets. Criteria:
  - Very close to 0 (or 0) at both ends
  - Mean value of 0
  - Morlet wavelet well suited for localizing frequency information in time
- Assumption of wavelet convolution:
  - Signal is stationary only during time period in which wavelet looks like a sine wave
  - Also, violations of local stationarity do not invalidate wavelet convolution results
    - Will decrease accuracy of frequency information
- **All time-frequency decomposition methods characteristic:**
  - Activity at each time point is an estimate of instantaneous activity and is influenced by activity from neighboring time points

12.2 How to Make Wavelets

- 1. Create a sine wave (Same number of time points/sampling rate as Gaussian)
- 2. Create a Gaussian wave (Same number of time points/sampling rate as ^^ sine wave)
- 3. Multiply them point by point
  - Frequency of wavelet is frequency of sine wave
- Frequency of a Morlet wave is its *peak (or center) frequency*
  - Wavelets contain energy in a range of frequency bands for which the frequency of the sine wave is the peak
- Create Gaussian Window

$$GaussWin = ae^{-(t-m)^2/(2s^2)}$$

- a = amplitude (height of Gaussian)
- t = is time
- m = an x-axis offset
- s = standard deviation or width of Gaussian

$$s = \frac{n}{2\pi f}$$

- f is frequency in Hz
- n is number of wavelet cycles
  - Determines trade-off between temporal and frequency precision
  - **Super important**
- Fourier transform
  - Many sine waves of different frequencies
- Time-frequency decomposition via wavelet convolution

- o Many wavelets of different frequencies that can be specified by you (rather than by number of data points in time series)
- o Number of wavelets to be used is not constrained at all
- Group of wavelets that share properties but differ in frequency = family of wavelets
- **Practical and theoretical limits of constructing families of wavelets:**
  - o 1. Can't use frequencies slower than epochs
    - Generally: 1s of data, use wavelets 4Hz faster
  - o 2. Frequencies of wavelets cannot be above ½ of the sampling rate (Nyquist Frequency)
  - o 3. Due to frequency smoothing from time freq. presicion tradeoffs, frequencies close to eachother will likely provide similar or nearly identical results
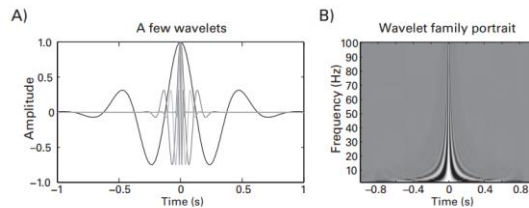   - Generally: 15 – 30 frequencies spanning 3Hz – 60Hz is sufficient

Figure 12.4
Different members of this wavelet family were created by changing the frequency of the sine wave while leaving other parameters unchanged.

- Figure 12.B
  - o Y-axis corresponds to peak frequency of members of wavelet family
  - o X – axis represents time
  - o Color of plot represents amplitude
    - Red = positive deflection of wavelet
    - Blue = negative

12.3 Wavelet Convolution as a Bandpass Filter

- Convolution
  - o A vector that expresses the time-varying mapping between a kernel and a signal
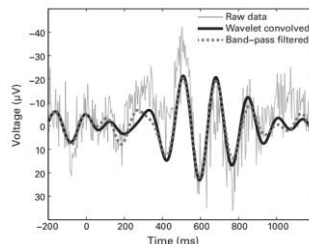
  - o Wavelet convolution is bandpass filtering
    - When freq. spectrum of wavelet is multiplied by frequency spectrum of EEG, inverse Fourier transform is computed
      - Resulting time series contains frequency characteristics of EEG data that are tapered by frequency characteristics of the wavelet (Gaussian around peak frequency)

12.4 Limitations of Wavelet Convolution as Discussed Thus Far

- 1. Some elements not apparent in bandpass filtered signal
    - Time-frequency analyses
    - Power Information
    - Phase information
        - Could use Hilbert transformation of band pass-filtered signal (but simper option in chap. 13)
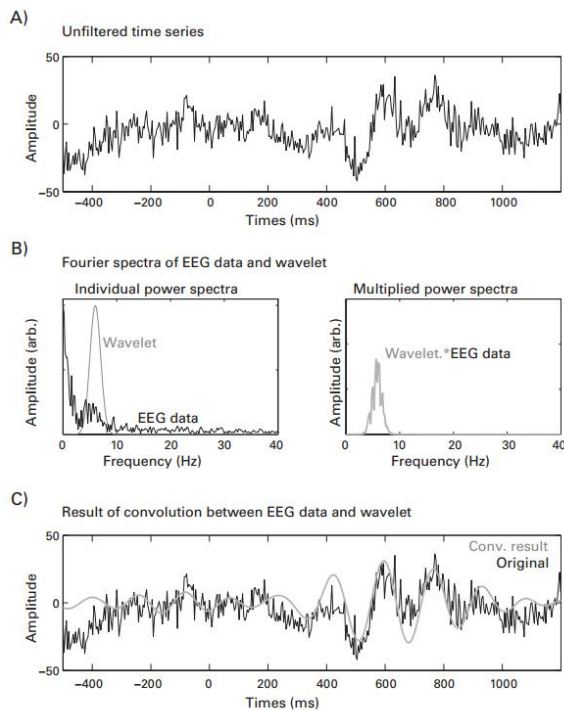- 2. Result of convolution with a Morlet wavelet depends on phase offsets between wavelet and data

Figure 12.6
Illustration of why wavelet convolution acts as a bandpass filter. Panel A shows raw EEG data from one trial from electrode FCz. Panel B shows the power spectra of the Morlet wavelet and the EEG data. Note that the wavelet has a Gaussian shape in the frequency domain. The left-hand plot in panel B shows the point-by-point multiplication of the frequency spectra of the wavelet and the EEG data. Panel C shows the result of convolution overlaid on top of the original EEG data. This figure also illustrates how the result of convolution reflects activity that is maximal at the peak frequency of the wavelet (here, 6 Hz), but also activity from a weighted combination of surrounding frequencies (here ranging from around 3 Hz to 9 Hz).
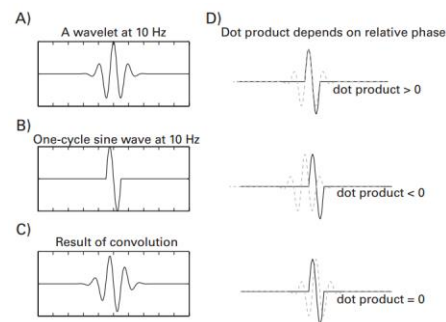

Figure 12.7
The result of each step of convolution (the dot product between the wavelet and the data) depends on the phase relationship between the kernel and the signal at that time point. This issue is resolved by using complex wavelets, as discussed in chapter 13.

- Same sine wave → different taperings
    - Convolution results show:
        - There are points where 2 vectors are orthogonal
            - When 90 degree phase lag between wavelet and one-cycle sine wave
        - Other points in time where two vectors have negative dot product
            - When 180 degree phase lag between two signals
- If looking for energy in EEG signal at specific frequency and at a specific time point
    - Align wavelet so it has 0 degree phase lag with EEG data at time point of interest
    - Then compute dot product
    - ^^^^ is not what you want

- Want relationship between wavelet and EEG data at all time points and all phase lags (not just some)
- Resolving Both Limitations of real-valued Morlet Wavelets
    - EEG data re convolved with complex Morlet wavelets
        - Wavelets with both real and imaginary components
        - Mapping does not depend on phase lags
        - INSTEAD, represented in a 2D space
            - Allows you to extract:
            - Bandpass-filtered signal
            - Time-frequency power
            - Phase information

13 Complex Morlet Wavelets and Extracting Power and Phase

13.1 The Wavelet Complex

- Complex Morlet wavelets can be used to extract estimates of time-varying frequency band-specific power and phase from EEG data
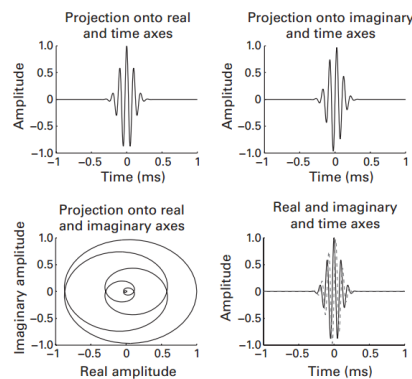   - Occupies 3D space:
        - Time / Real / Imaginary

Figure 13.1
A complex wavelet is a 3-D (time, real, imaginary) function. Plotted here are projections onto various pairs of those dimensions.
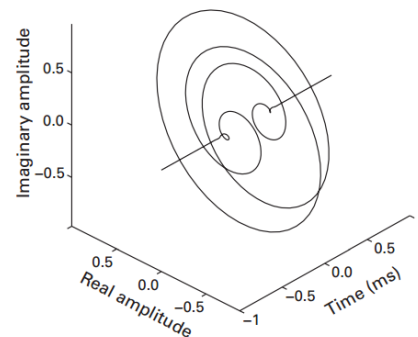

Figure 13.2
Three-dimensional view of a complex wavelet.

13.2 Imagining the Imaginary

- Imaginary numbers indicated with *i* or *j*

13.2 Rectangular and Polar Notation and the Complex Plane

- Polar coordinates useful for:
  - Describing properties of frequency-band-specific activity
    - Bandpass-filtered signal
    - Power
    - Phase
- Convert from polar to Cartesian (for imaginary component) using basic trig.
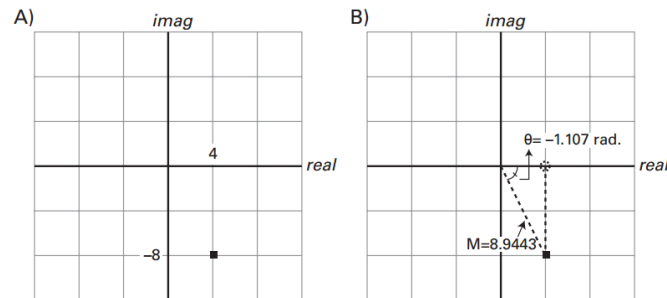

Figure 13.3
The same point in a complex space (with a real axis and an imaginary axis) can be represented using Cartesian (A) or polar (B) notations.

$$M = \sqrt{(real^2 + imag^2)} \qquad real = M\cos(\theta)$$

$$\theta = \arctan(imag\,/\,real) \qquad imag = M\sin(\theta)$$

$$real + imag = M\cos(\theta) + M\sin(\theta)$$

$$real + imag = M[\cos(\theta) + \sin(\theta)]$$

$$a + ib = M[\cos(\theta) + i\sin(\theta)]$$

13.3 Euler's Formula

- Allows you to represent complex numbers as points on a circle

$$Me^{i\theta} = M[\cos(\theta) + i\sin(\theta)]$$

  - e = base of natural log (2.718)
  - θ = any real number (angle in radians)
  - M = magnitude of vector
- Change cosine = change in real-axis position
- Change sine = change in imaginary-axis position
- Difference between real and imaginary parts of a complex wavelet is difference between sine/cosine
  - Sine / cosine related
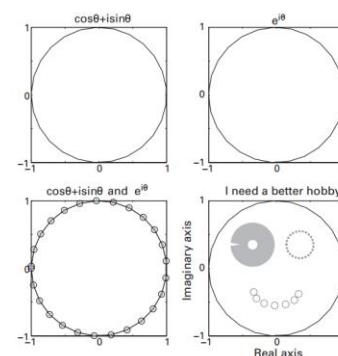   - ¼ counterclockwise rotation in complex space

Figure 13.4
Illustration of equality of Euler and trigonometric representations and another illustration (lower right panel) of how the real component corresponds to cosine while the imaginary component corresponds to sine.
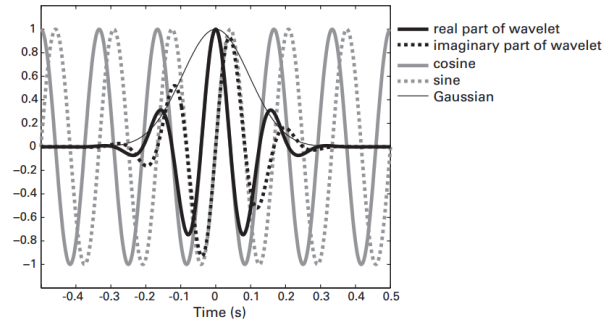
Figure 13.5
Overlaying the real and imaginary parts of a complex wavelet along with a cosine wave and a sine wave. This shows that the real part of the wavelet corresponds to a cosine wave, and the imaginary part of the wavelet corresponds to a sine wave.

- Complex Morlet Wavelet Equation

$$cmw = Ae^{-t^2/2s^2}e^{i2\pi ft}$$

$$A = \frac{1}{\left(s\sqrt{\pi}\right)^{1/2}}$$

- 1st part of equation: Gaussian
- 2nd part of equation: complex sine wave
  - Combination of Euler's method and 2*pi*f*t part of sine wave
- S = standard deviation of gaussian
- f = peak frequency
- A = frequency band-specific scaling factor
  - Not necessary if plan to apply:
    - Baseline normatization
      - Percentage change
      - Decibel
    - Complex wavelet convolution only to obtain phase angle time series
- Complex sine wave composed of:
  - Cosine (real) and sine (imaginary) component

13.4 Euler's Formula and the Result of Complex Wavelet Convolution

- Dot product between wavelet and one-cycle sine wave could be pos, neg, or zero
    - Depending on slight shifts in relative phase between signal and kernel
- Dot product between a complex wavelet and signal, result is complex number
    - Euler's formula applied to represent complex number as point in polar space
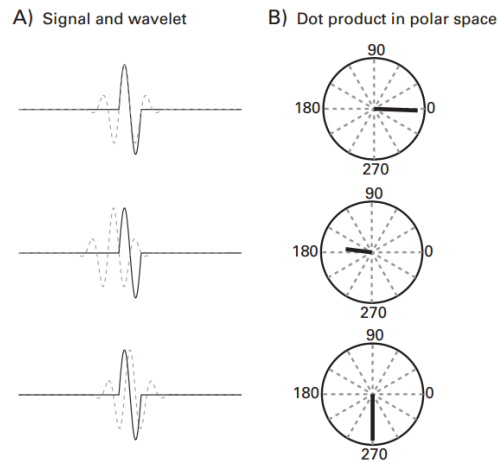   - (Endpoint of a vector from origin of polar space to that point)

A) Signal and wavelet   B) Dot product in polar space

Figure 13.6

The dot product (one step of convolution) between a complex Morlet wavelet (real part shown here using a dotted line) and a one-cycle sine wave produces a complex number, here represented as a vector in a polar plot using the magnitude and angle. Note that the more the wavelet and the one-cycle sine wave overlap, the longer the vector is in complex space, regardless of the phase angle of that vector. This figure can be compared with figure 12.7.

- Notice:
    - When dot product with real-valued wavelet was less than zero, dot product with complex wavelet has vector pointing to the left
    - … was zero, dot product with complex wavelet has vector pointing down
    - … was positive, dot product with complex wavelet has vector pointing right
- The result of dot products with real-valued wavelet maps onto the real axis in the result of dot products with complex wavelets
    - Imaginary axis is ignored in dot product with real-valued wavelet