



Digital Speech Processing

به نام خدا



گروه هوش مصنوعی، دانشکده مهندسی
کامپیوتر

گفتار پردازش رقمی

آنالیز فوری

حمیدرضا برادران کاشانی

پاییز ۱۴۰۱



سرفصل مطالب

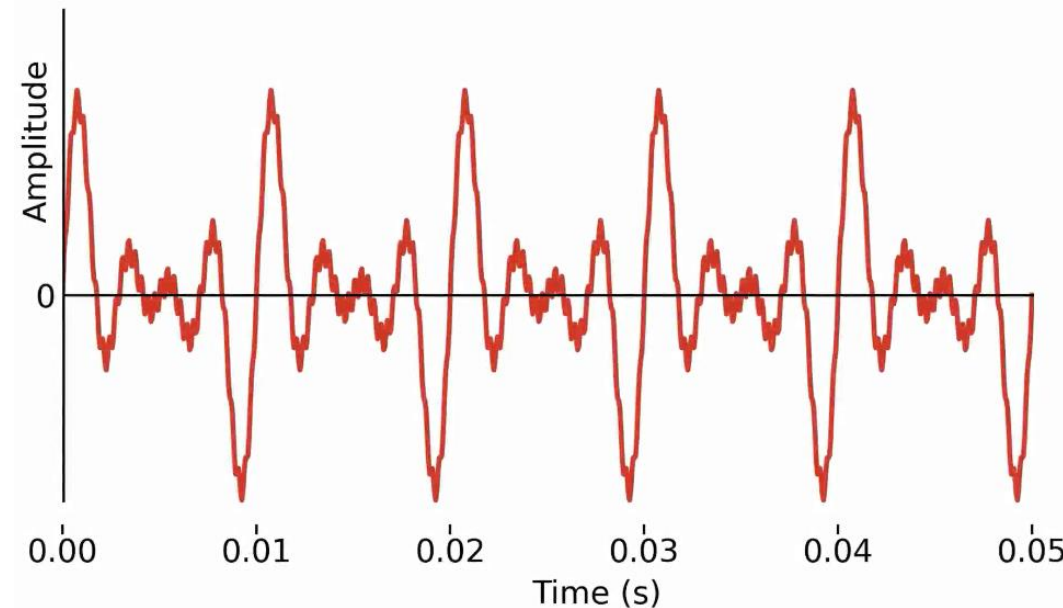
- ❖ بسط سری
- ❖ آنالیز فوریه
- ❖ تحلیل در حوزه فرکانس

Hamidreza Baradaran Kashani



گسترش سری (Series Expansion)

❖ یک سیگنال شبه گفتار را در نظر بگیرید:



❖ **آنالیز** یا **تحلیل** این سیگنال به چه معنا است؟

❖ می خواهیم بدانیم این سیگنال یا موج از چه سیگنال یا امواج دیگری ساخته شده است.

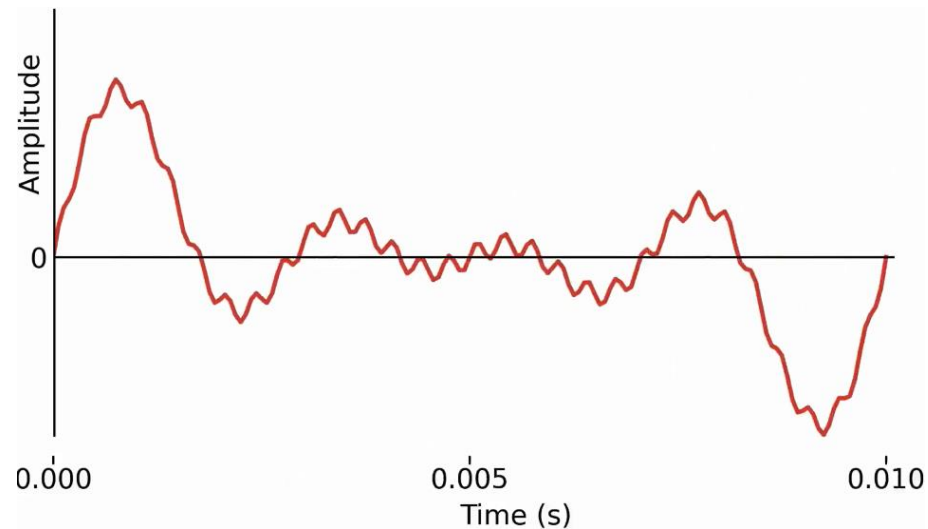
❖ روش **گسترش سری** به ما میگوید این سیگنال پیچیده از مجموع چه سیگنالهایی ساخته شده است.

Hamidreza Baradaran Kashani



گسترش سری (Series Expansion)

❖ یک دوره تناوب این سیگنال را در نظر بگیریم:



❖ حال می خواهیم این سیگنال را بصورت مجموع یک سری تابع یا سیگنال ساده نشان دهیم که به آنها توابع پایه (basis functions) می گوییم.

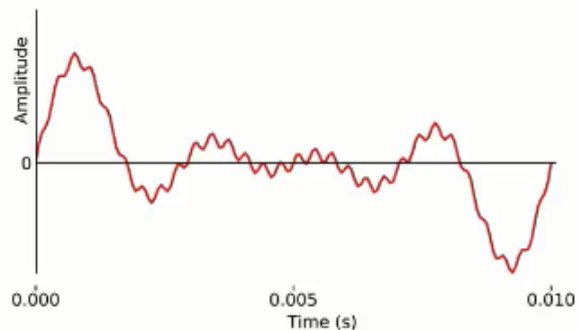
❖ ساده ترین نوع تابع پایه: **سیگنال سینوسی** که متناوب است.

Hamidreza Baradaran Kashani



گسترش سری (Series Expansion)

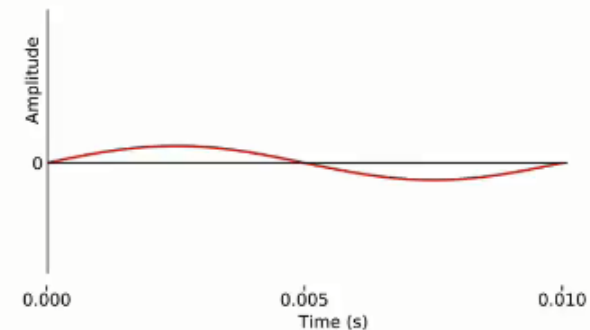
اولین تابع پایه بزرگترین دوره تناوب را دارد



\approx



$=$

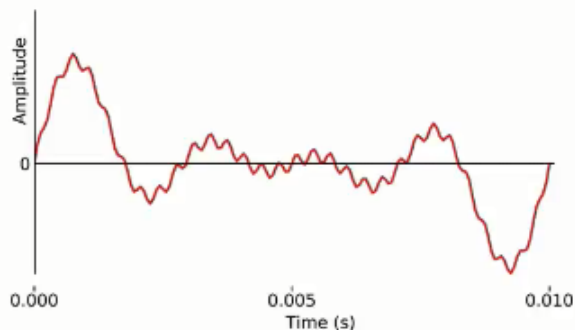


Idaran Kashani

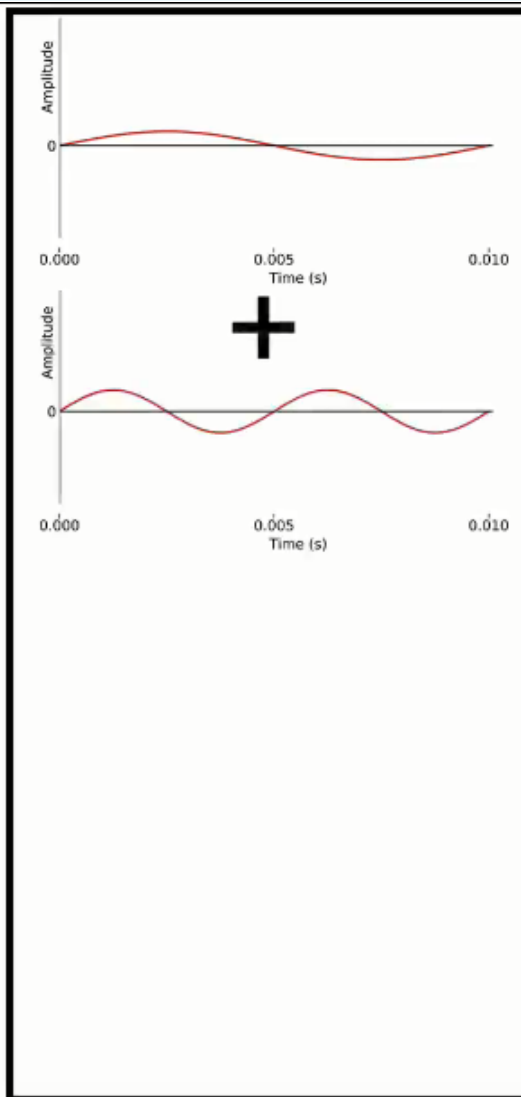


گسترش سری (Series Expansion)

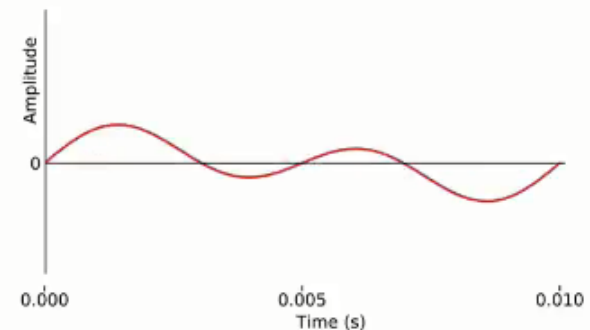
دومین تابع پایه نصف دوره تناوب
اولین را دارد



\approx



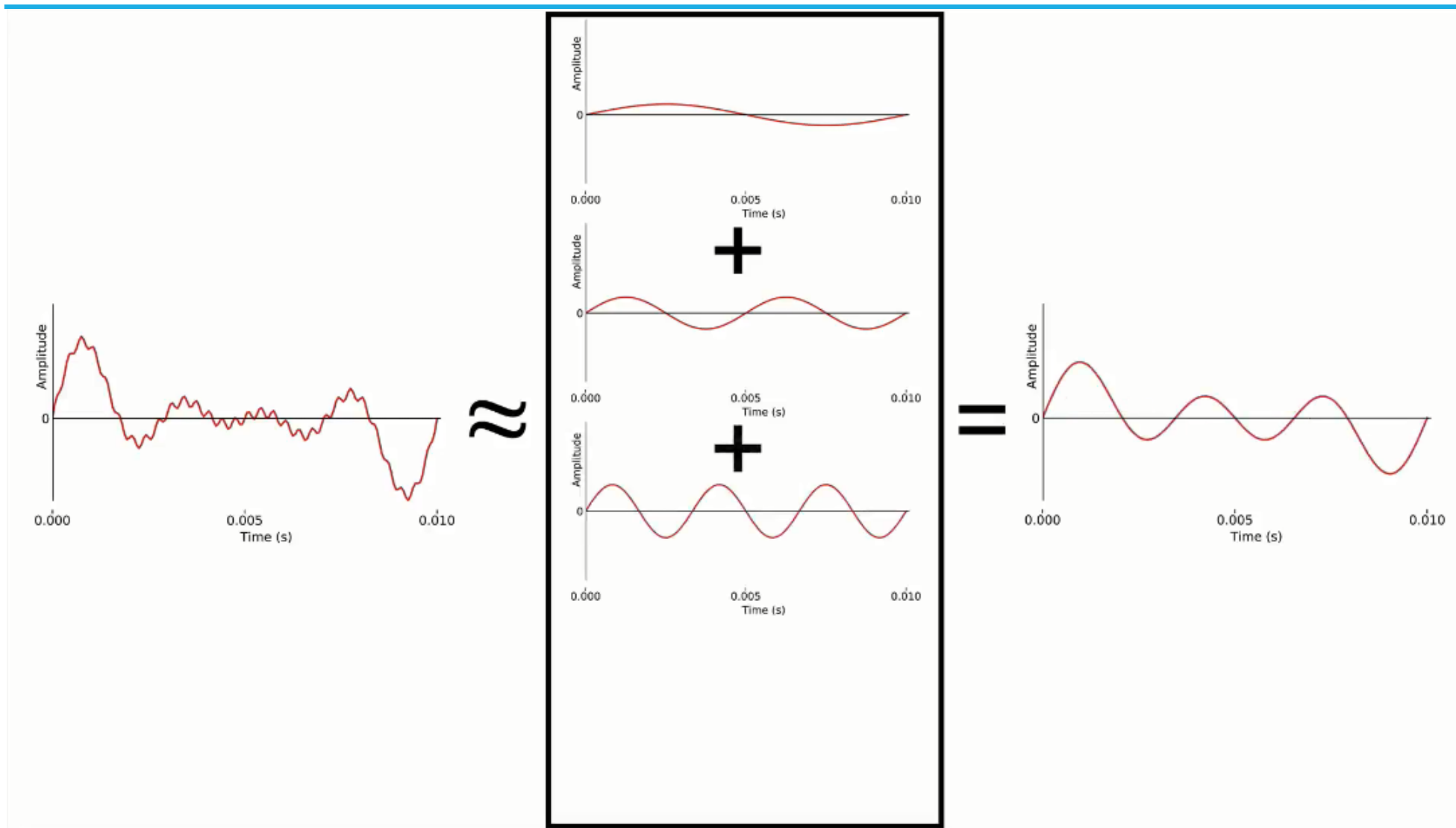
$=$



Iaran Kashani



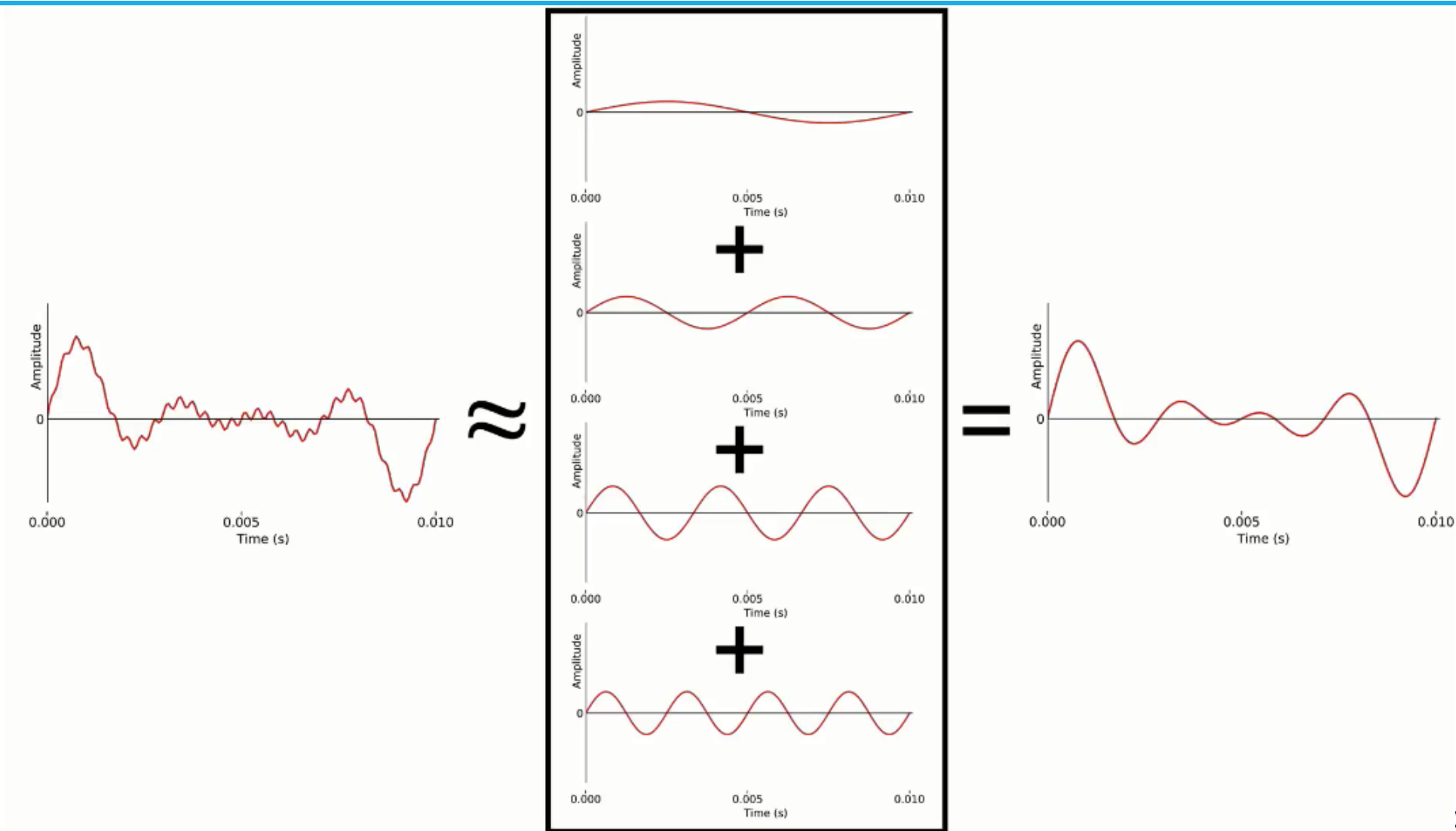
گسترش سری (Series Expansion)



Jaran Kashani



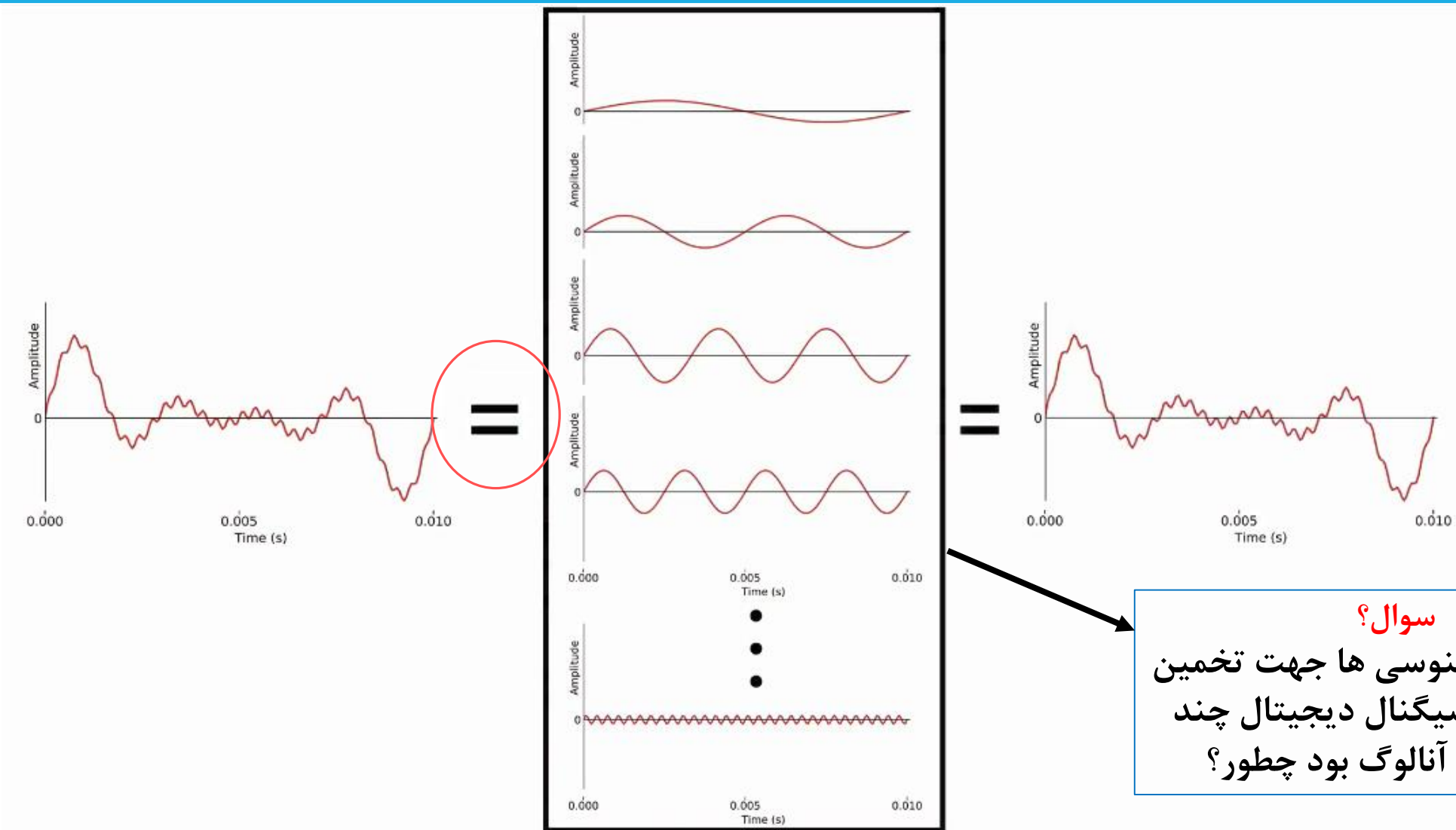
گسترش سری (Series Expansion)



rdaran Kashani



گسترش سری (Series Expansion)



Idaran Kashani



گسترش سری (Series Expansion)

❖ تعبیر دیگر:

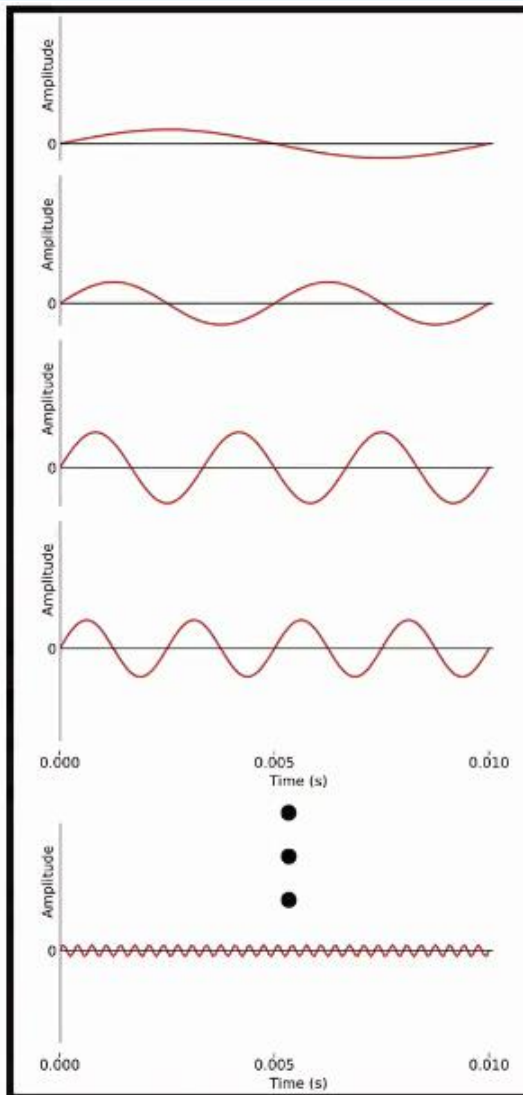
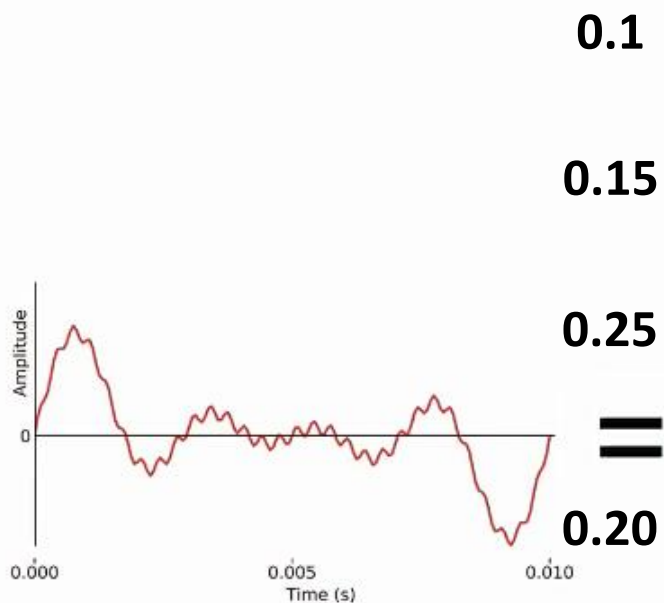
❖ توابع پایه خود نیز سیگنال های دیجیتال هستند.

❖ بالاترین فرکانس در این توابع پایه برابر با فرکانس نایکویست است.

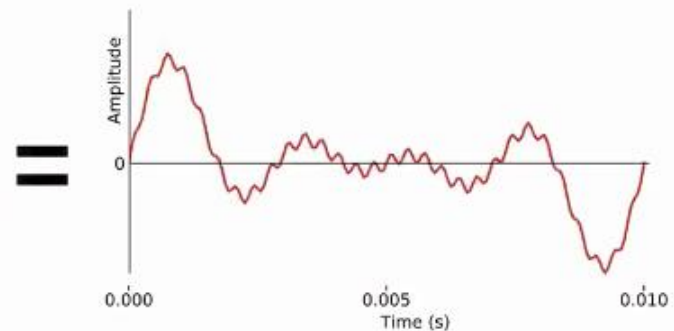
❖ پس اگر تعداد محدودی سیگنال دیجیتال که بیشترین فرکانس آنها برابر با فرکانس نایکویست هست را با یکدیگر جمع کنیم، دقیقاً به سیگنال دیجیتال اولیه می‌رسیم.



گسترش سری (Series Expansion)



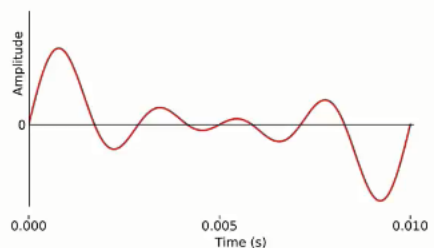
برای بازسازی سیگنال اصلی
توابع پایه بایستی با وزن های
مشخصی با یکدیگر جمع شوند.



در واقع یک موج پیچیده را بصورت
جمع وزن دار یک سری سینوسی با
فرکانس های مشخص نوشتیم (بسط
سری)

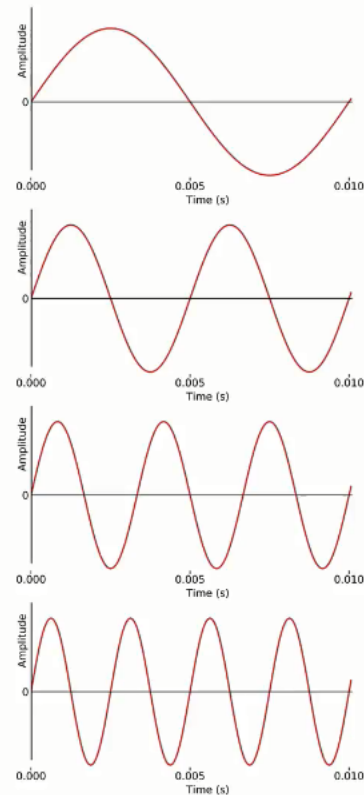


گسترش سری (Series Expansion)



=

$$\begin{aligned} &0.10 \times \\ &+ 0.15 \times \\ &+ 0.25 \times \\ &+ 0.20 \times \end{aligned}$$



آنالیز
فوریه

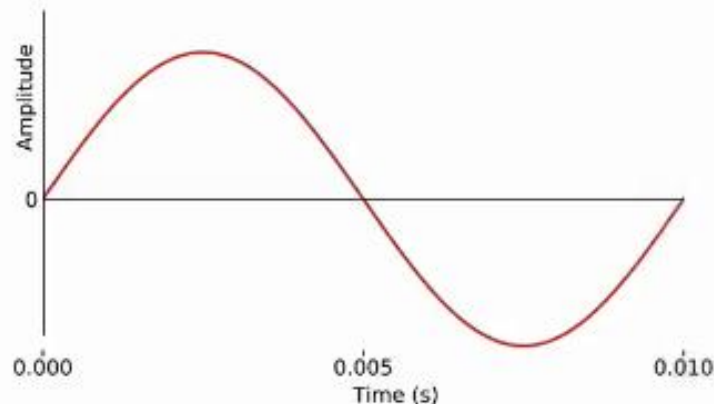
این ضرایب چگونه حاصل می شود؟ (با
فرض داشتن سیگنال اولیه و توابع پایه)

Hamidreza Baradaran Kashani

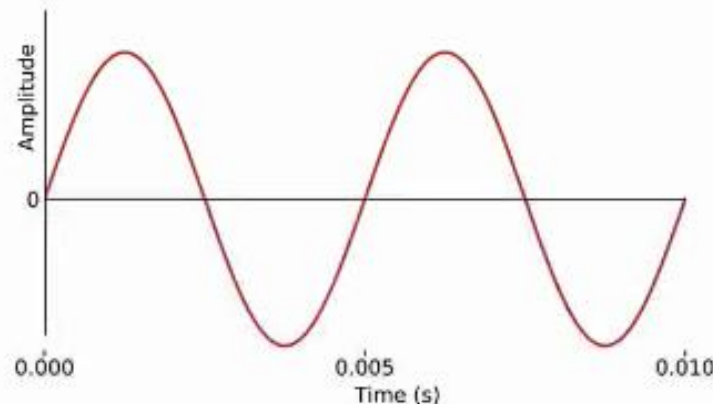


آنالیز فوریه

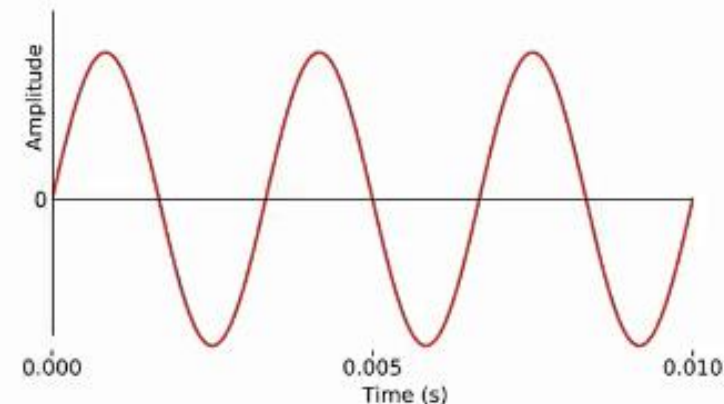
100 Hz



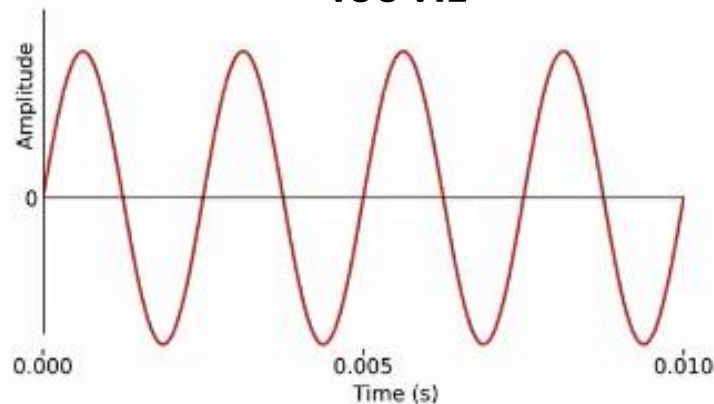
200 Hz



300 Hz



400 Hz



سوال؟

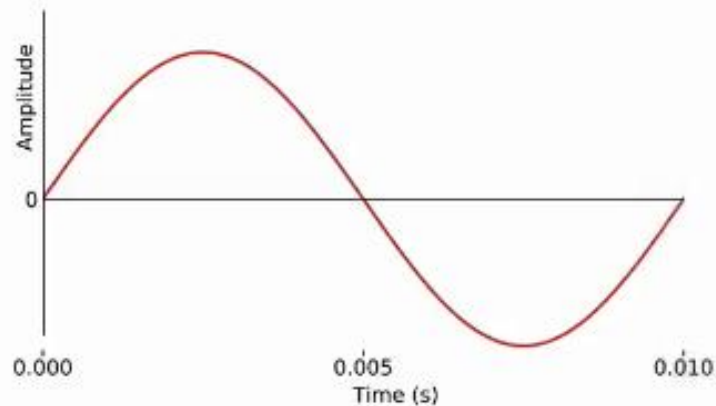
اگر فرکانس نمونه برداری 16kHz باشد،
فرکانس بیشترین تابع پایه چند
است؟

Hamidreza Baradaran Kashani

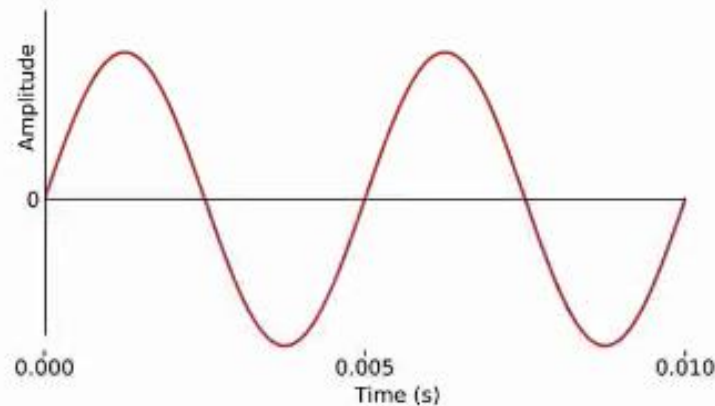


آنالیز فوریه

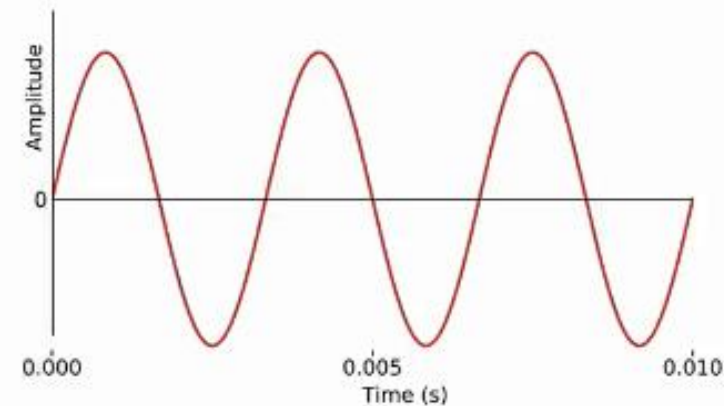
100 Hz



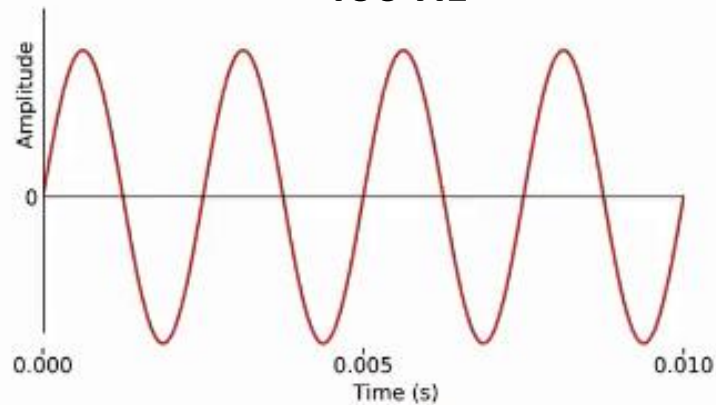
200 Hz



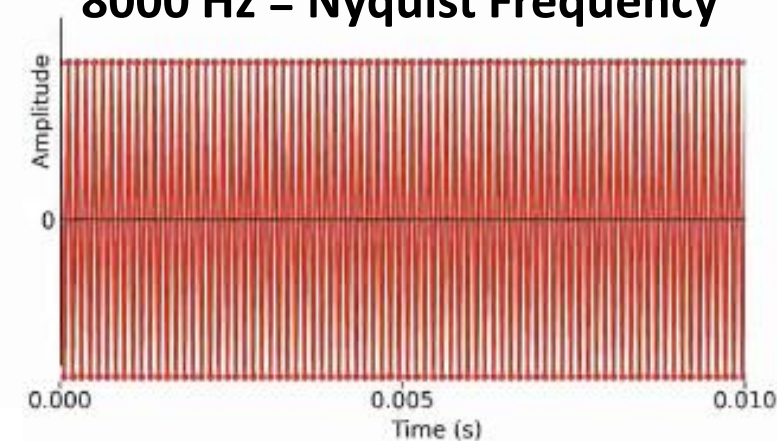
300 Hz



400 Hz



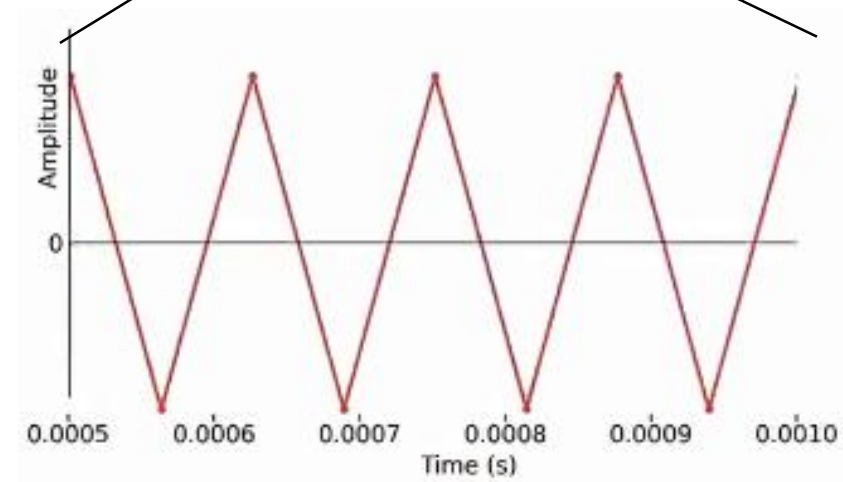
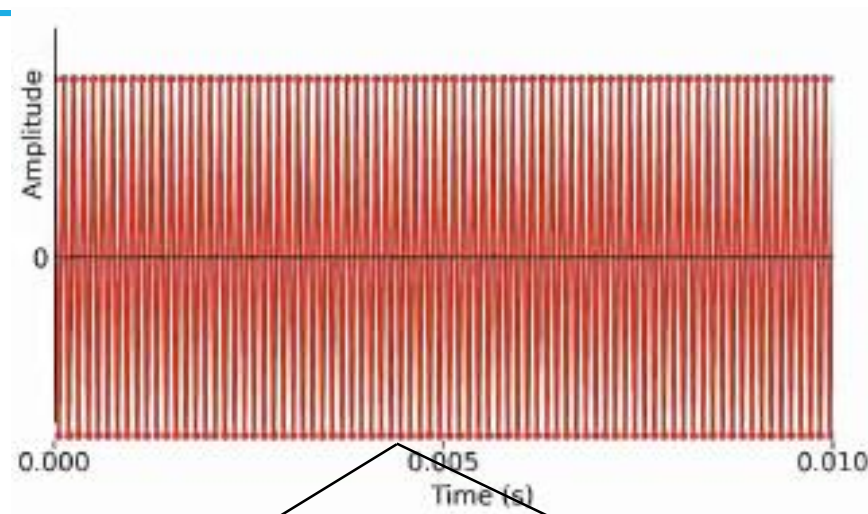
8000 Hz = Nyquist Frequency



Hamidreza Baradaran Kashani



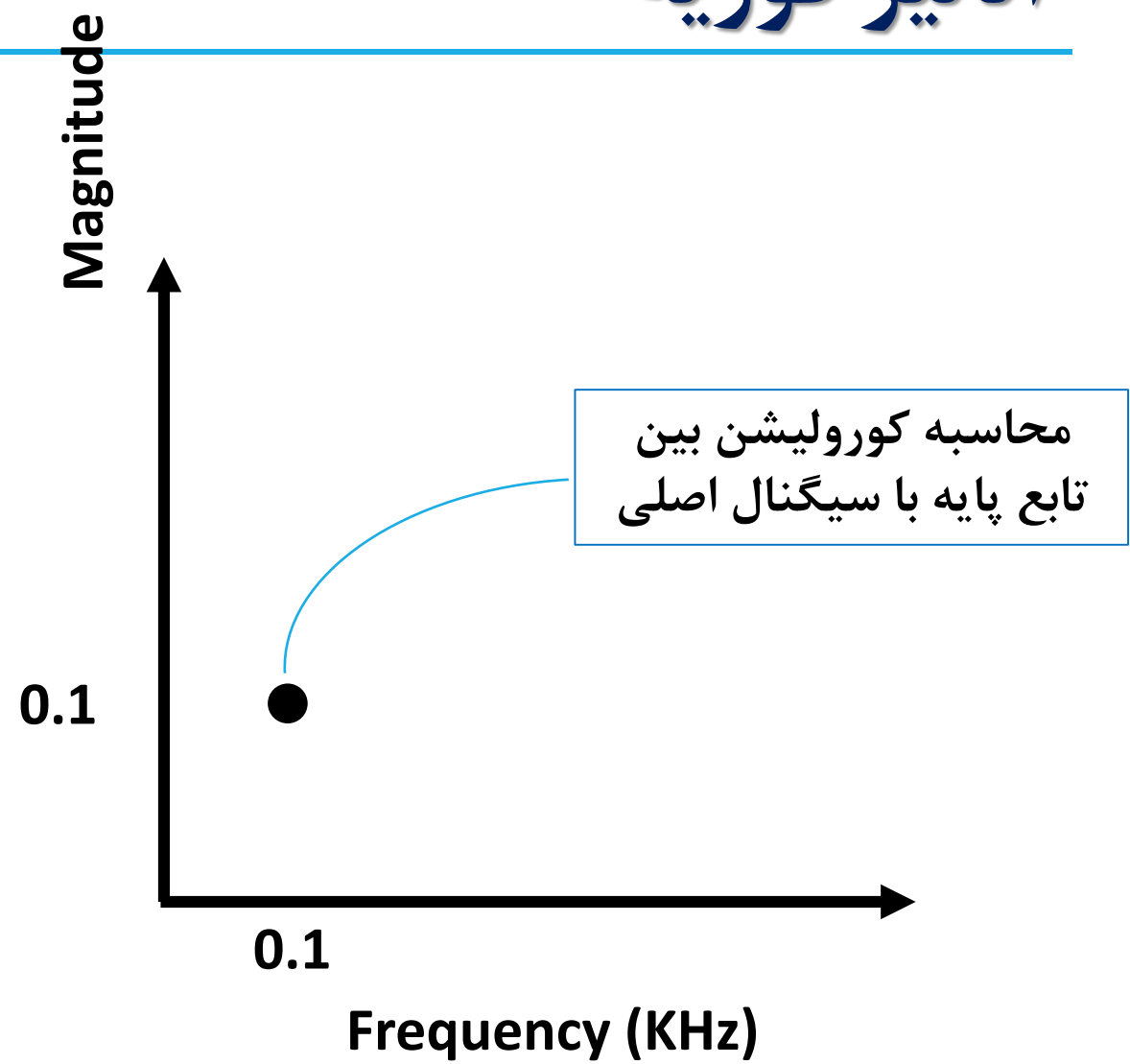
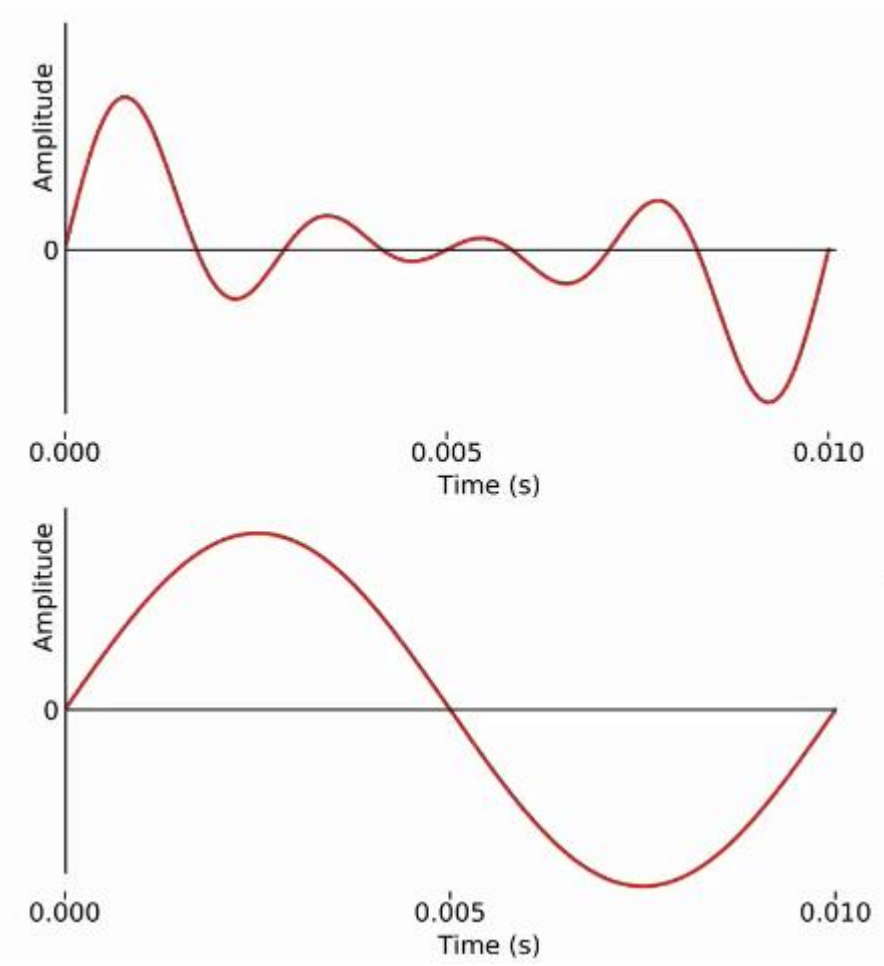
آنالیز فوریه



Hamidreza Baradaran Kashani



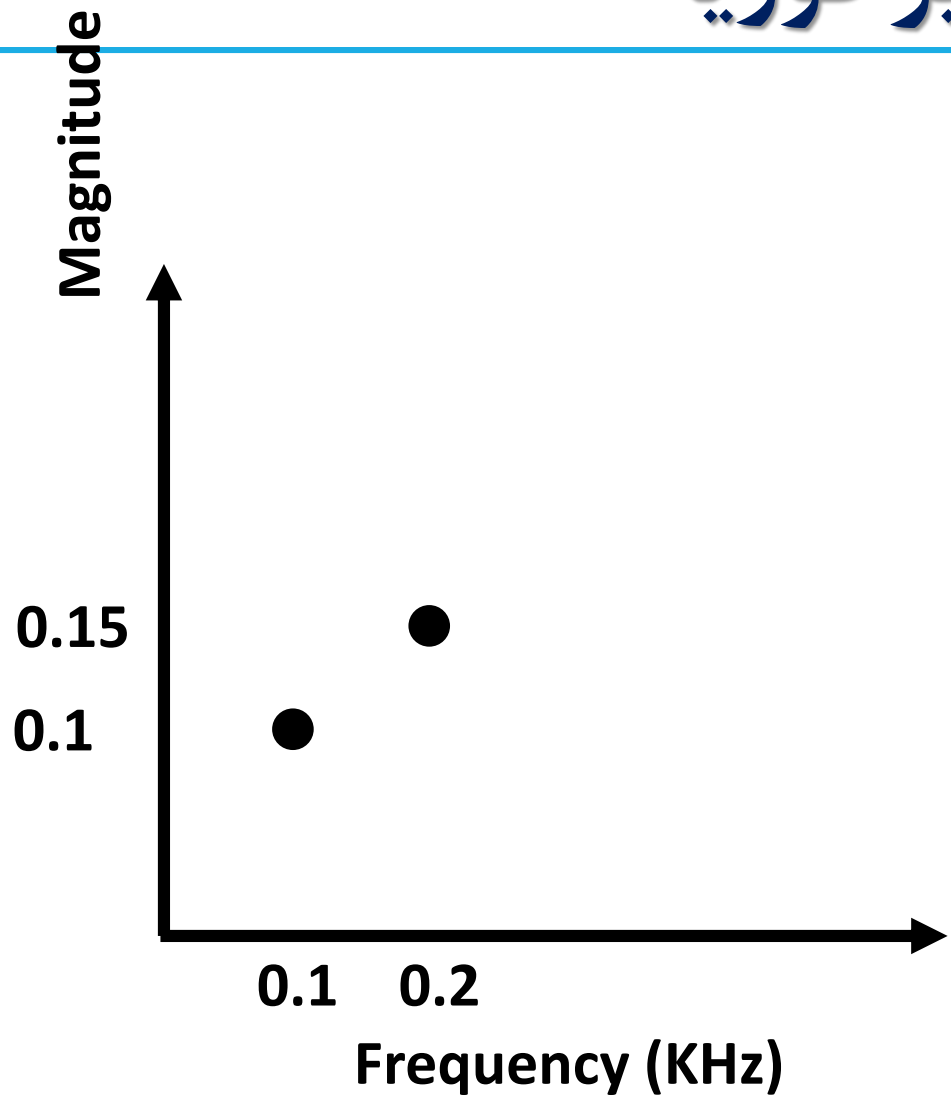
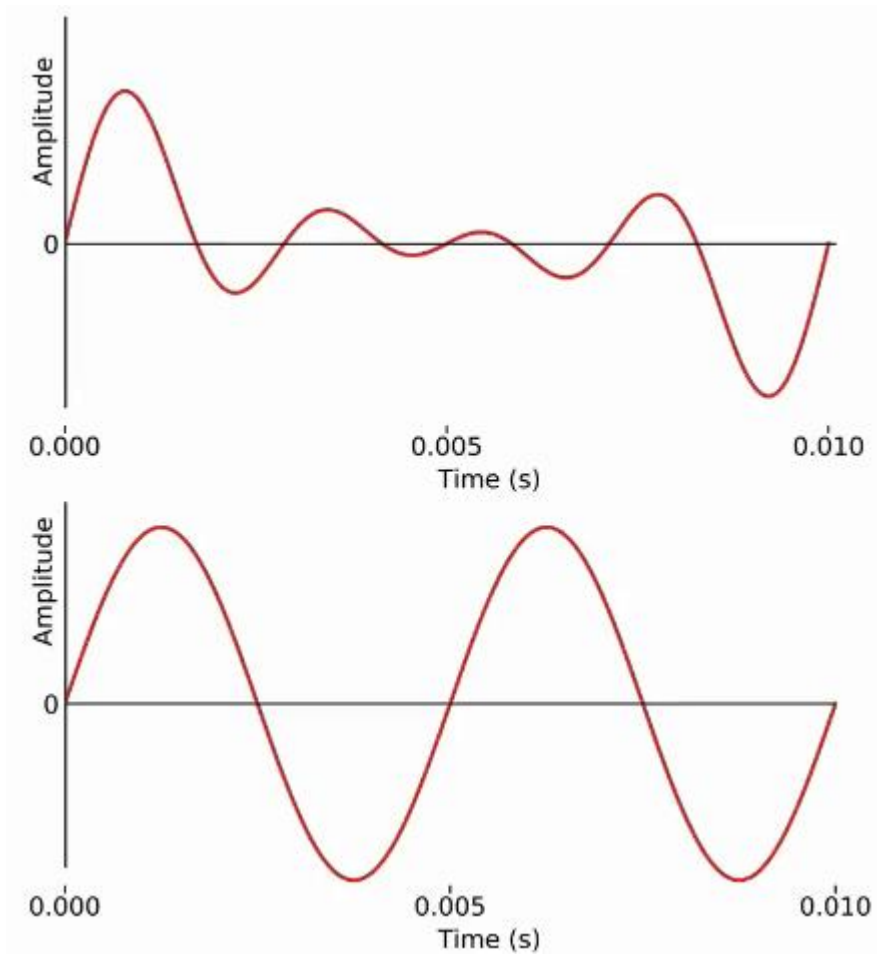
آنالیز فوریه



Hamidreza Baradaran Kashani



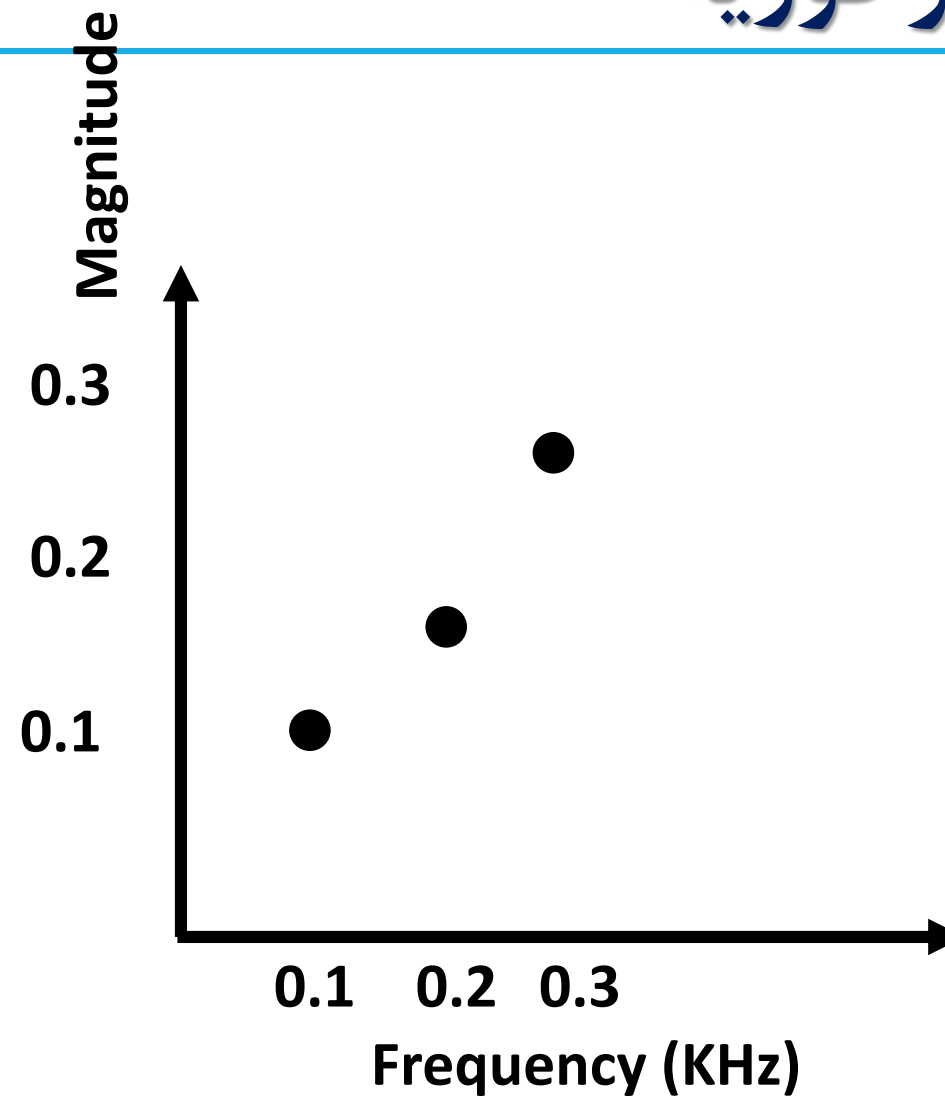
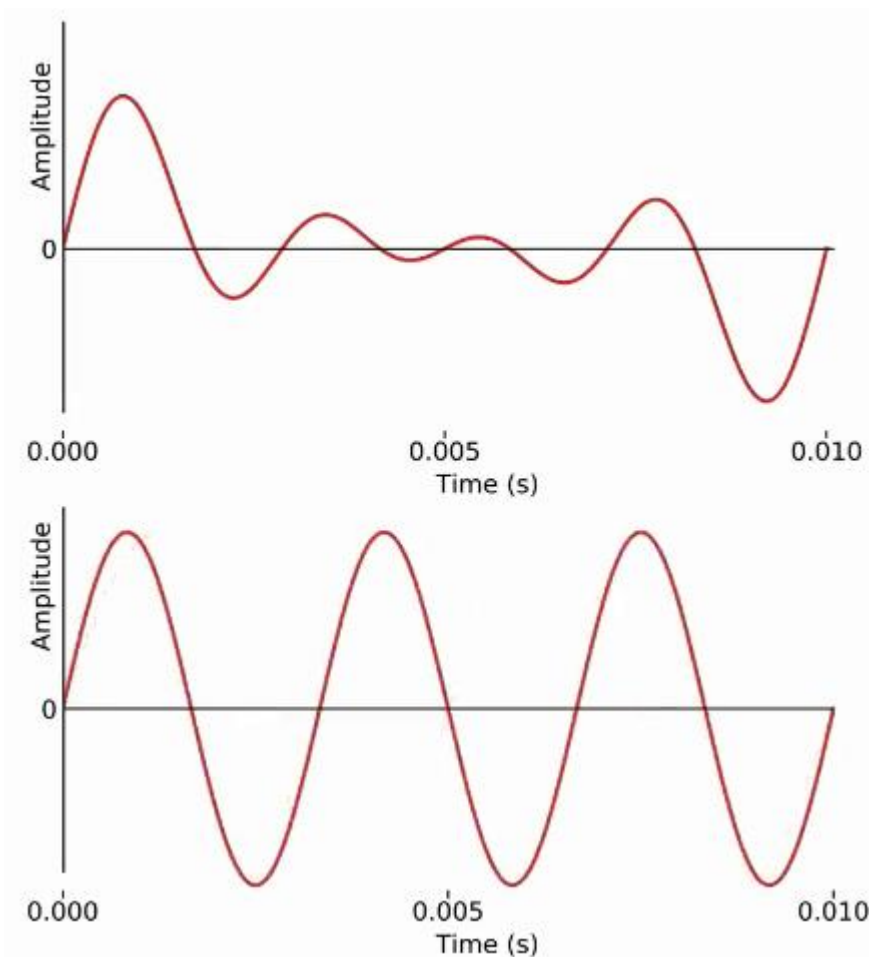
آنالیز فوریه



Hamidreza Baradaran Kashani



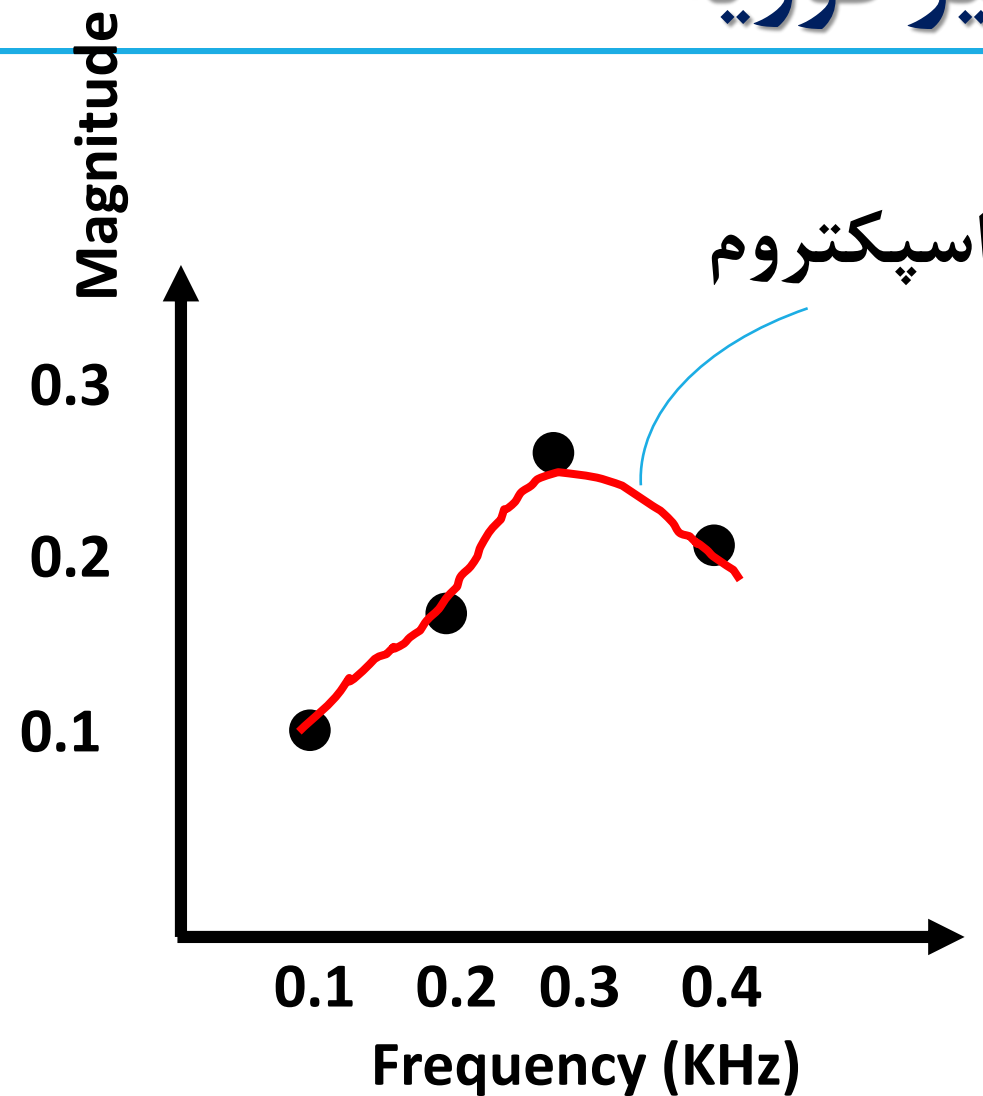
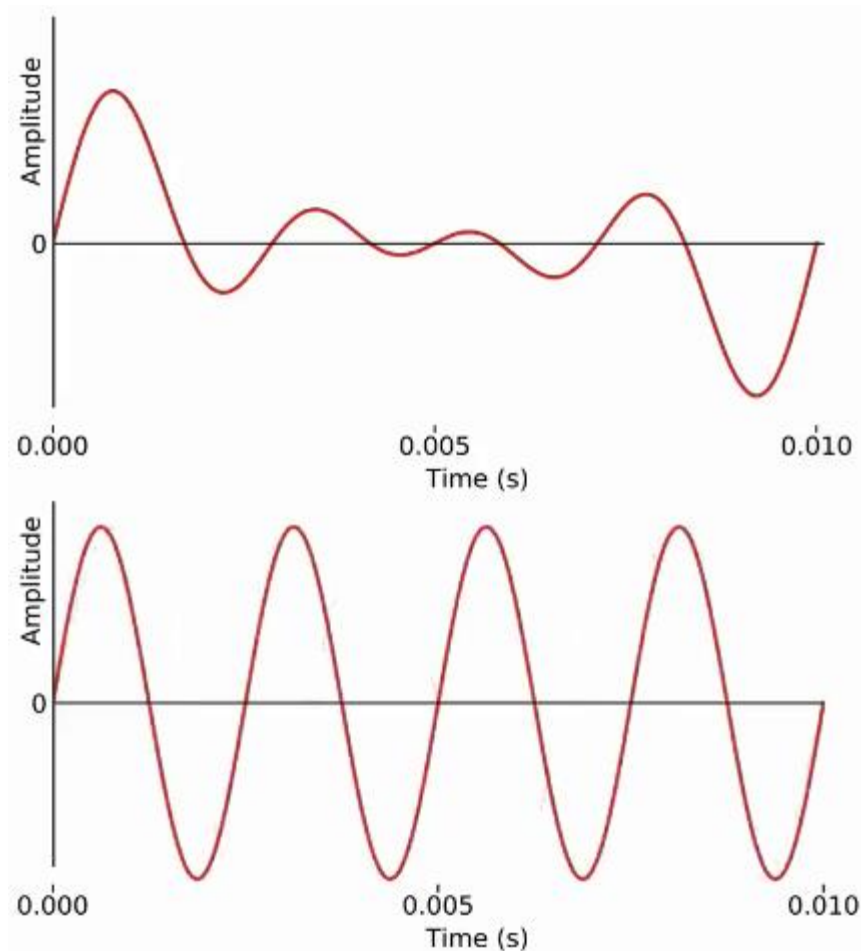
آنالیز فوریه



Hamidreza Baradaran Kashani



آنالیز فوریه

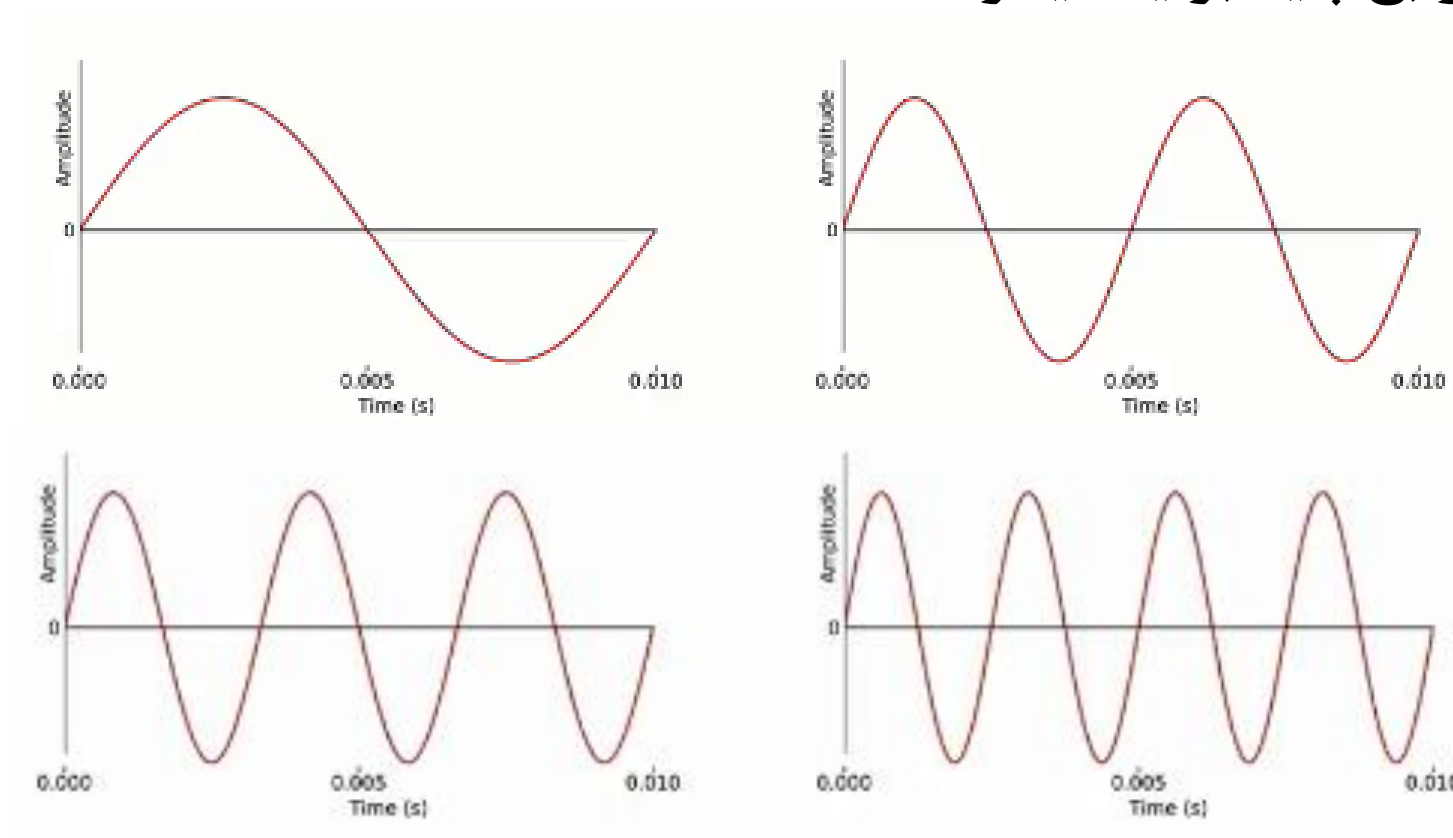


Hamidreza Baradaran Kashani



آنالیز فوریه

- ❖ خاصیت مهم آنالیز فوریه:
- ❖ متعامد بودن توابع پایه بر یکدیگر



Hamidreza Baradaran Kashani



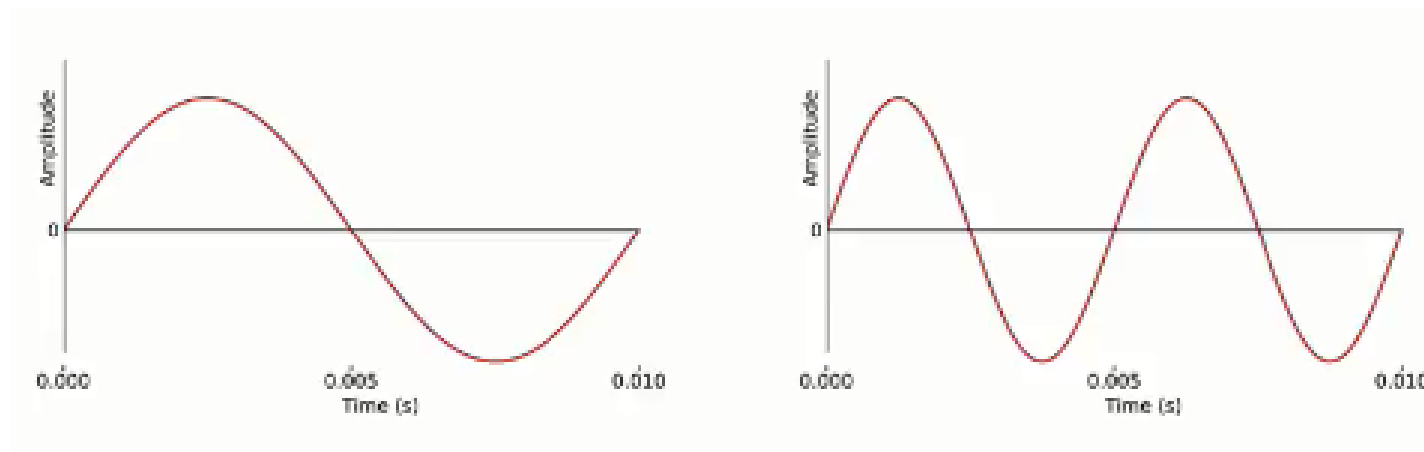
آنالیز فوریه

❖ خاصیت مهم آنالیز فوریه:

❖ متعامد بودن توابع پایه بر یکدیگر

❖ ضرب دورنی هر جفت از این توابع در یکدیگر برابر با صفر است.

❖ به عبارت دیگر، از فرکانس سیگنال پایه اول هیچ مقدار انرژی در سیگنال پایه دوم نیست و بالعکس.



Hamidreza Baradaran Kashani



❖ متعامد بودن توابع پایه بر یکدیگر به معنای آن است:

❖ می توان هر سیگنال را بصورت جمع وزن دار یک سری توابع پایه نوشت که مقادیر وزن منحصر بفرد هستند (unique solution).

❖ به عبارتی برای هر سیگنال فقط یک مجموعه ضریب وجود دارد (unique solution).

❖ این مجموعه ضریب همان اطلاعات موجود در سیگنال اصلی است.

❖ از طرفی معکوس این تحلیل یعنی سنتز سیگنال اصلی از روی توابع پایه و مقادیر وزنشان براحتی امکان پذیر است.

❖ تجزیه و تحلیل فوریه را تبدیل فوریه نیز می گویند: تبدیل از حوزه زمان به حوزه فرکانس



تحلیل حوزه فرکانس

❖ در تبدیل فوریه سیگنال (طیف سیگنال) دو نوع اطلاعات وجود دارند:

❖ دامنه طیف (magnitude spectrum)

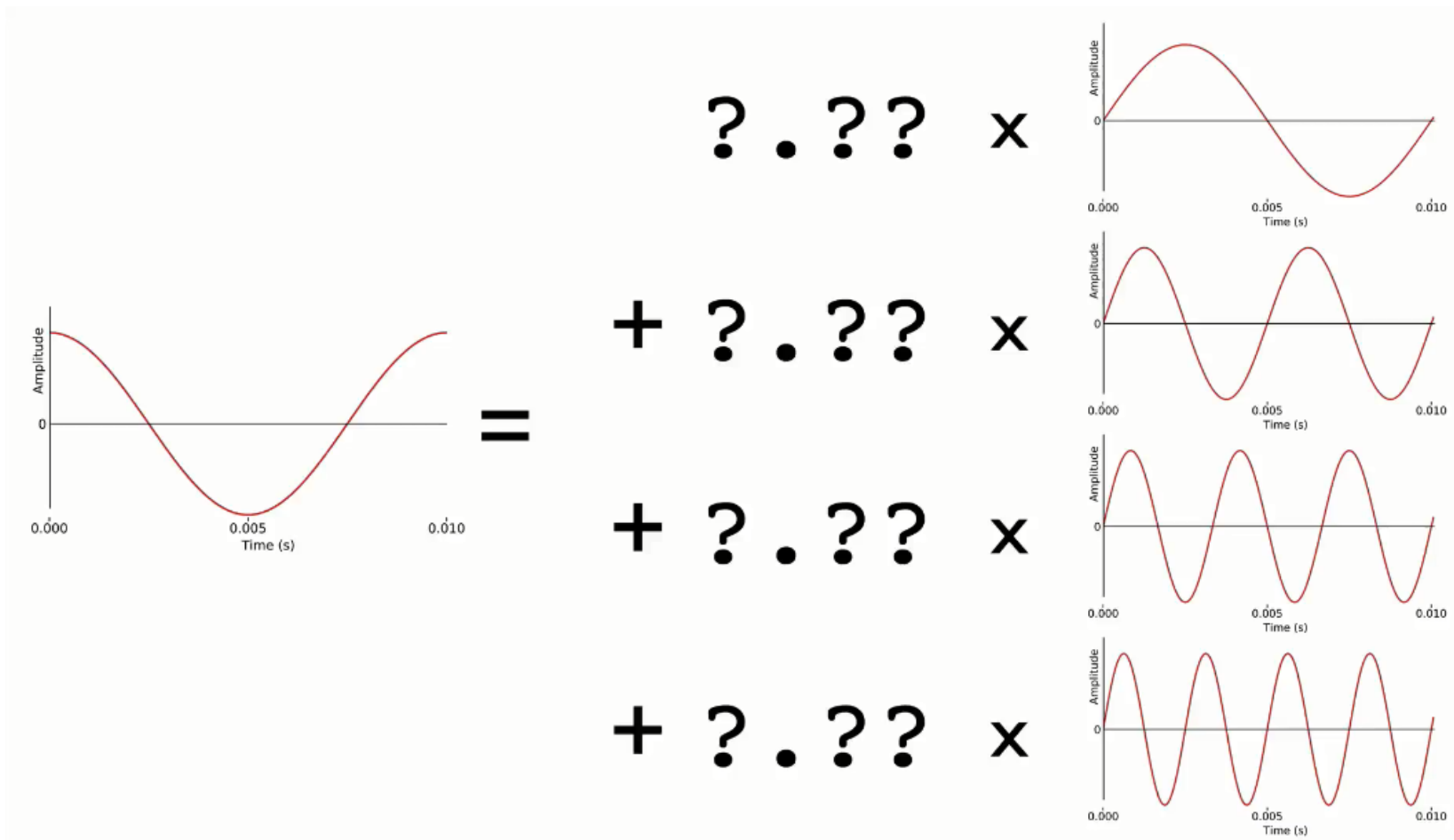
❖ فاز طیف (phase spectrum)

❖ سوال؟

فاز چقدر در پردازش گفتار اهمیت دارد؟



تحلیل حوزه فرکانس



سوال؟

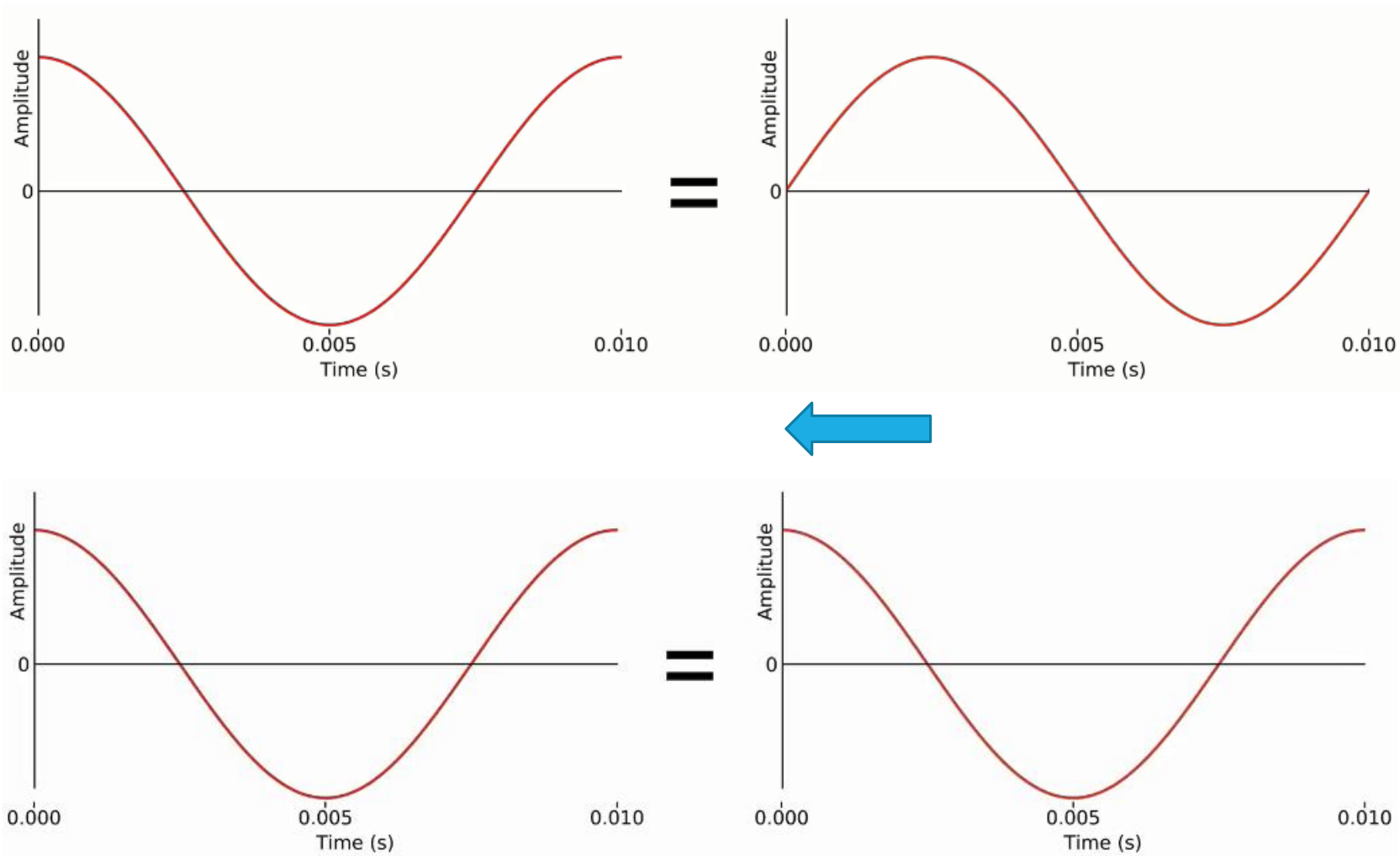
آیا می توان مجموعه ضرایبی
یافت که از روی توابع پایه
شکل فوق به سیگنال اصلی
برسیم؟ چرا؟

Hamidreza Baradaran Kashani



تحليل حوزه فرکانس - اثر فاز

راه حل؟
تغییر فاز توابع پایه

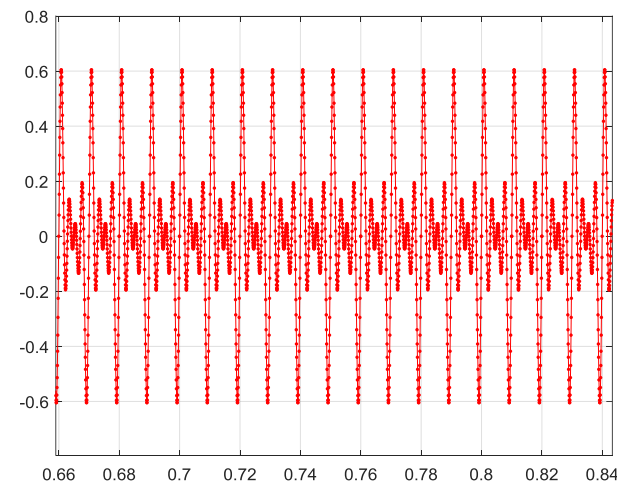
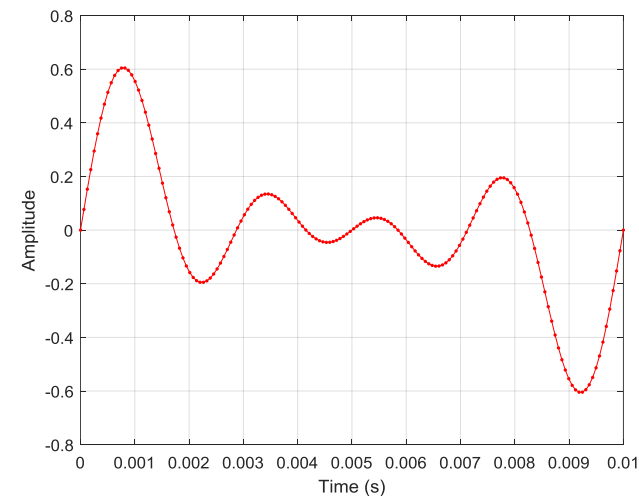
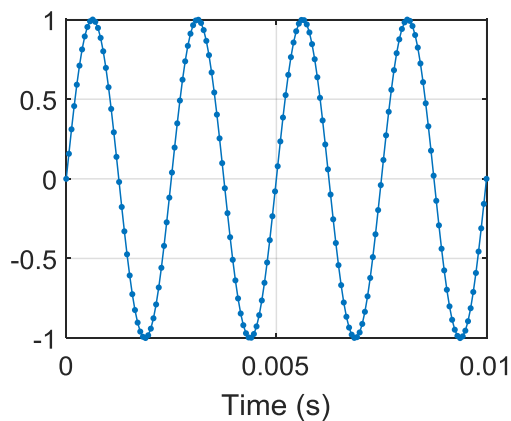
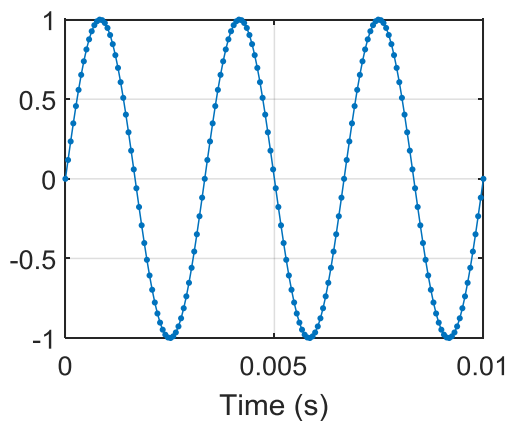
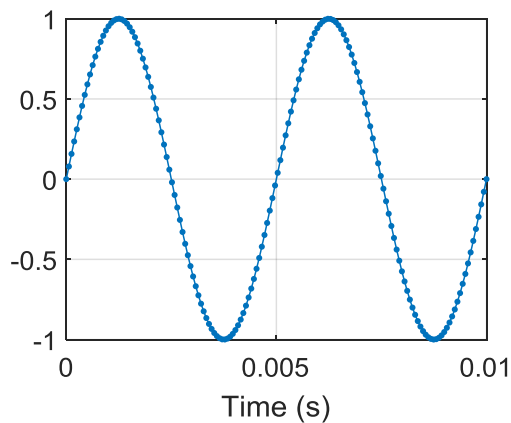
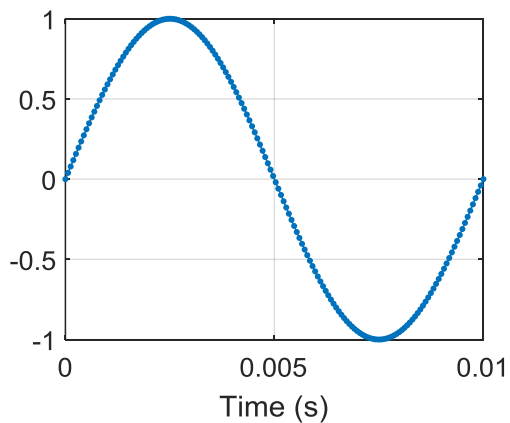


شیفت تابع پایه به
اندازه $1/4$ موج آن به
سمت چپ

Hamidreza Baradaran Kashani



تحليل حوزه فرکانس - اثر فاز

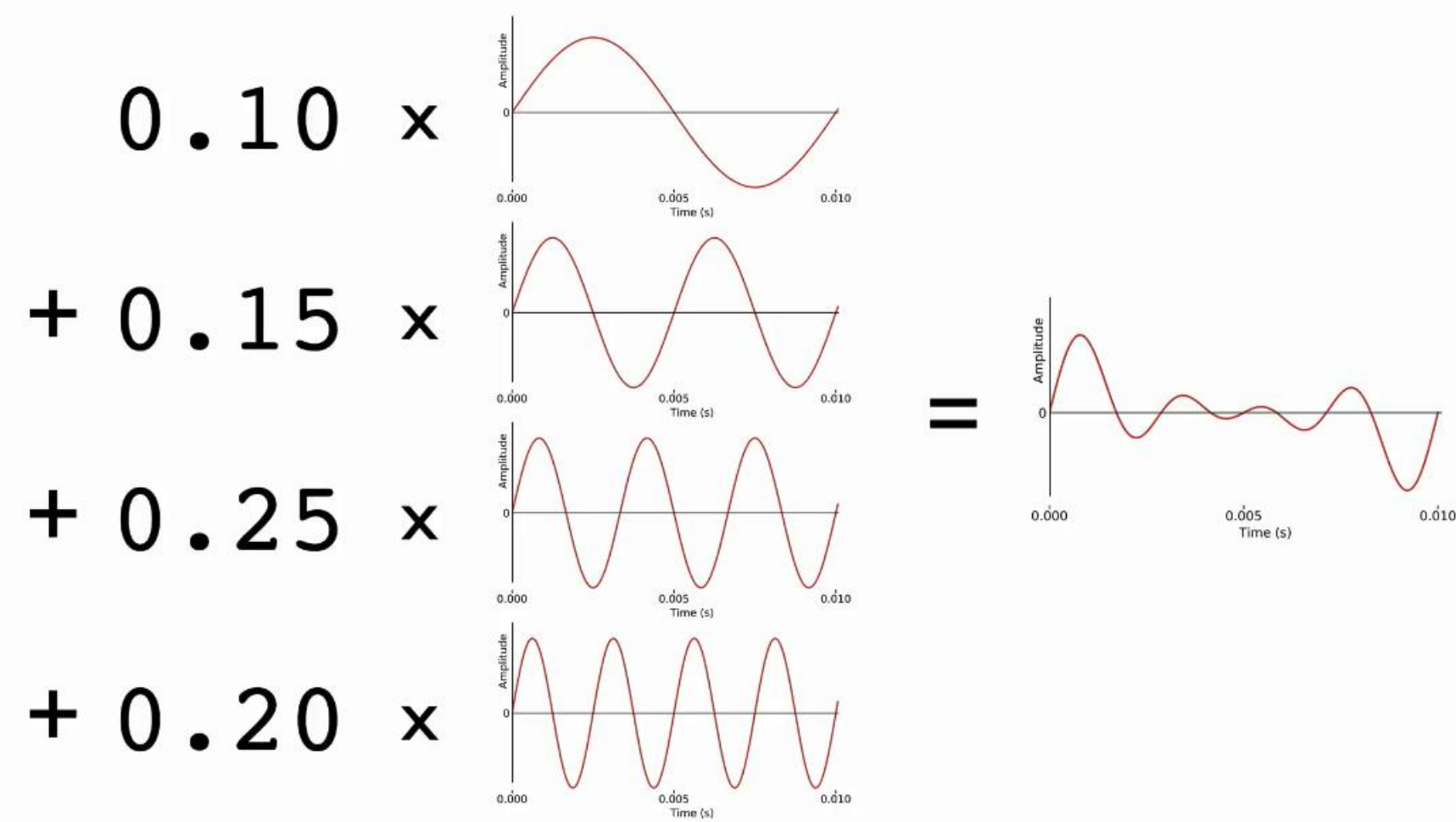


SumSineWaves.wav

Hamidreza Baradaran Kashani



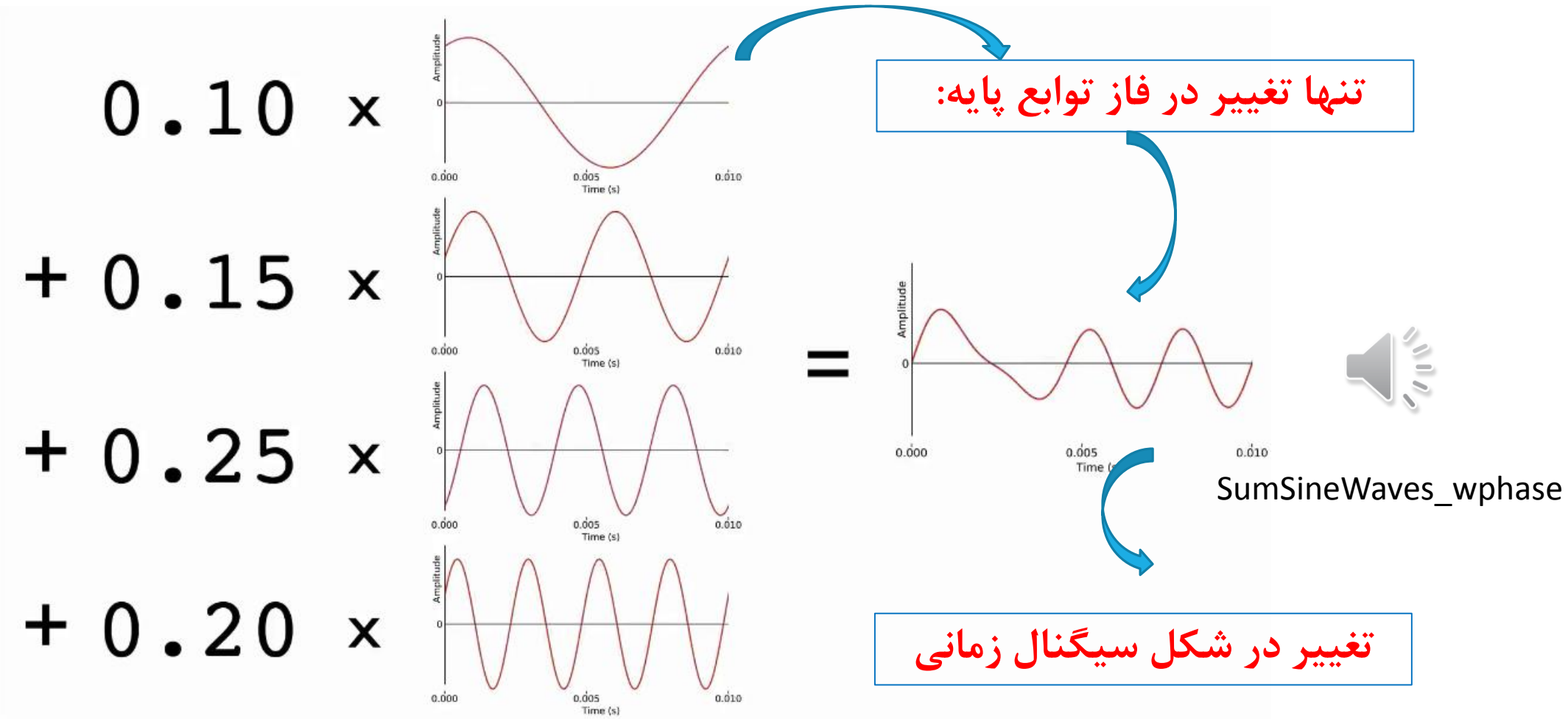
تحليل حوزه فرکانس - اثر فاز



Hamidreza Baradaran Kashani



تحلیل حوزه فرکانس - اثر فاز



Hamidreza Baradaran Kashani



تحلیل حوزه فرکانس - اثر فاز

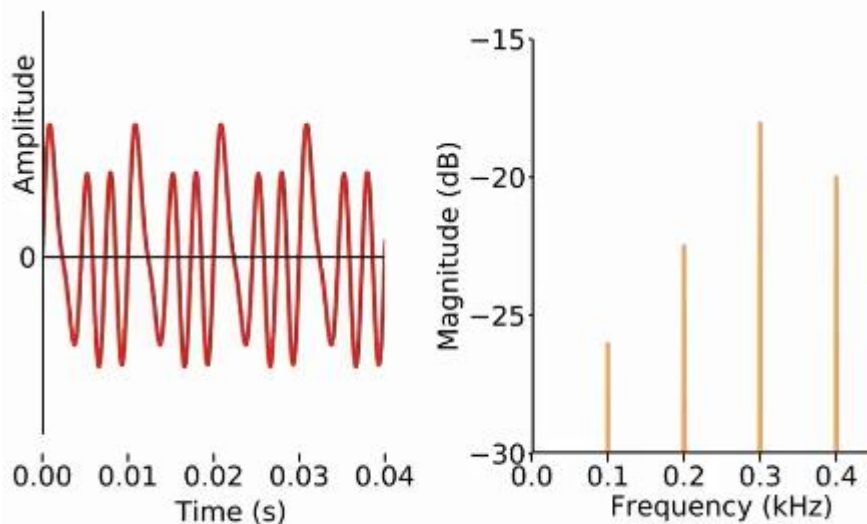
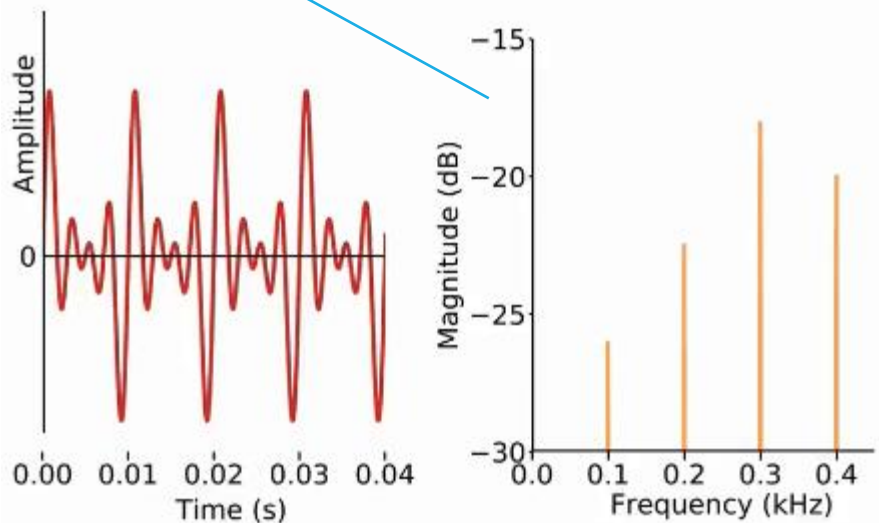
❖ در واقع با تغییر فاز توابع پایه:

❖ شکل سیگنال زمانی کاملاً متفاوت شده است. البته از لحاظ شنیداری تغییر چندانی حس نمی شود، چرا که سیستم شنوایی ما به این مقدار تغییرات فاز خیلی حساس نیست.

استفاده از مقیاس Log
روی دامنه طیف

❖ از این پس تمرکز ما روی دامنه طیف هست (نه فاز طیف).

❖ دامنه طیف (spectrum magnitude) دو شکل موج زیر یکسان است.



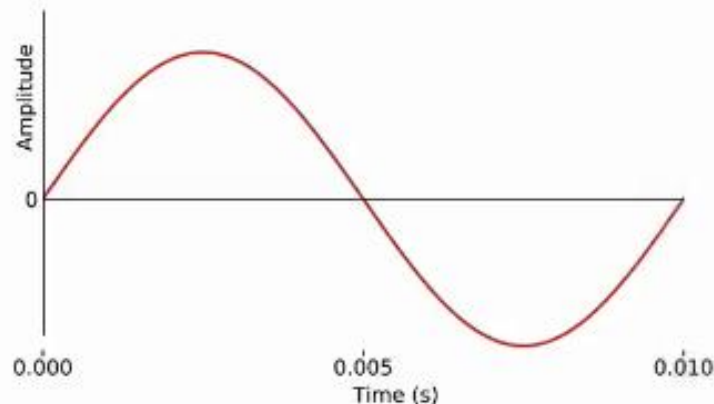
به نظر می رسد حوزه
زمانی حوزه مناسبی برای
آنالیز سیگنال گفتار
نیست و حوزه فرکانس و
بخصوص دامنه طیف
حاوی اطلاعات بیشتری
است

Hamidreza Baradaran Kashani

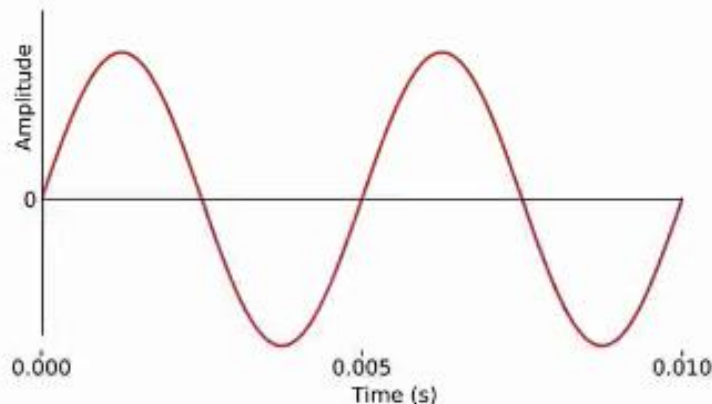


تحلیل حوزه فرکانس - اثر افزایش طول پنجره آنالیز

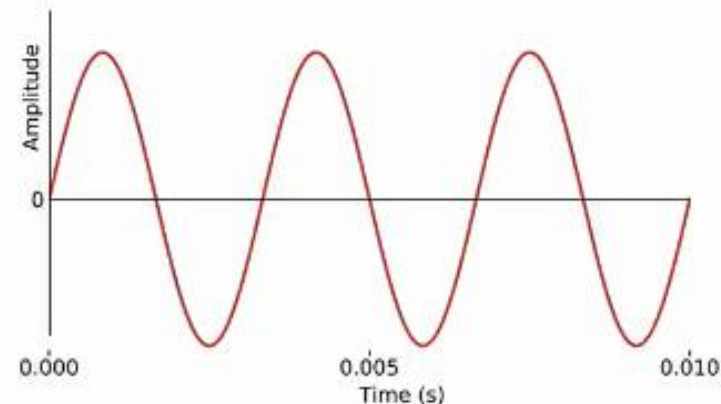
100 Hz



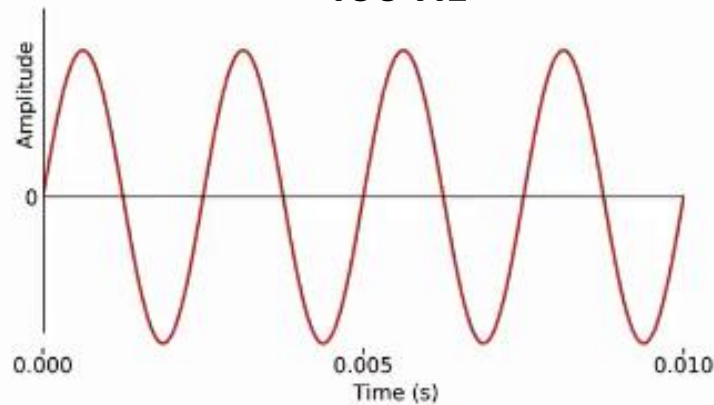
200 Hz



300 Hz

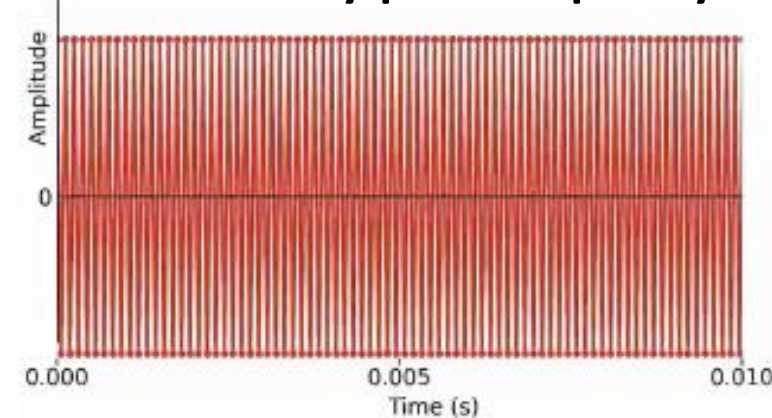


400 Hz



افزایش طول پنجره آنالیز:
- کمترین فرکانس سیگنال پایه
کاهش می یابد.
- سیگنال های پایه در فرکانس
یکسان از یکدیگر هستند

8000 Hz = Nyquist Frequency



Hamidreza Baradaran Kashani



تحلیل حوزه فرکانس - اثر افزایش طول پنجره آنالیز

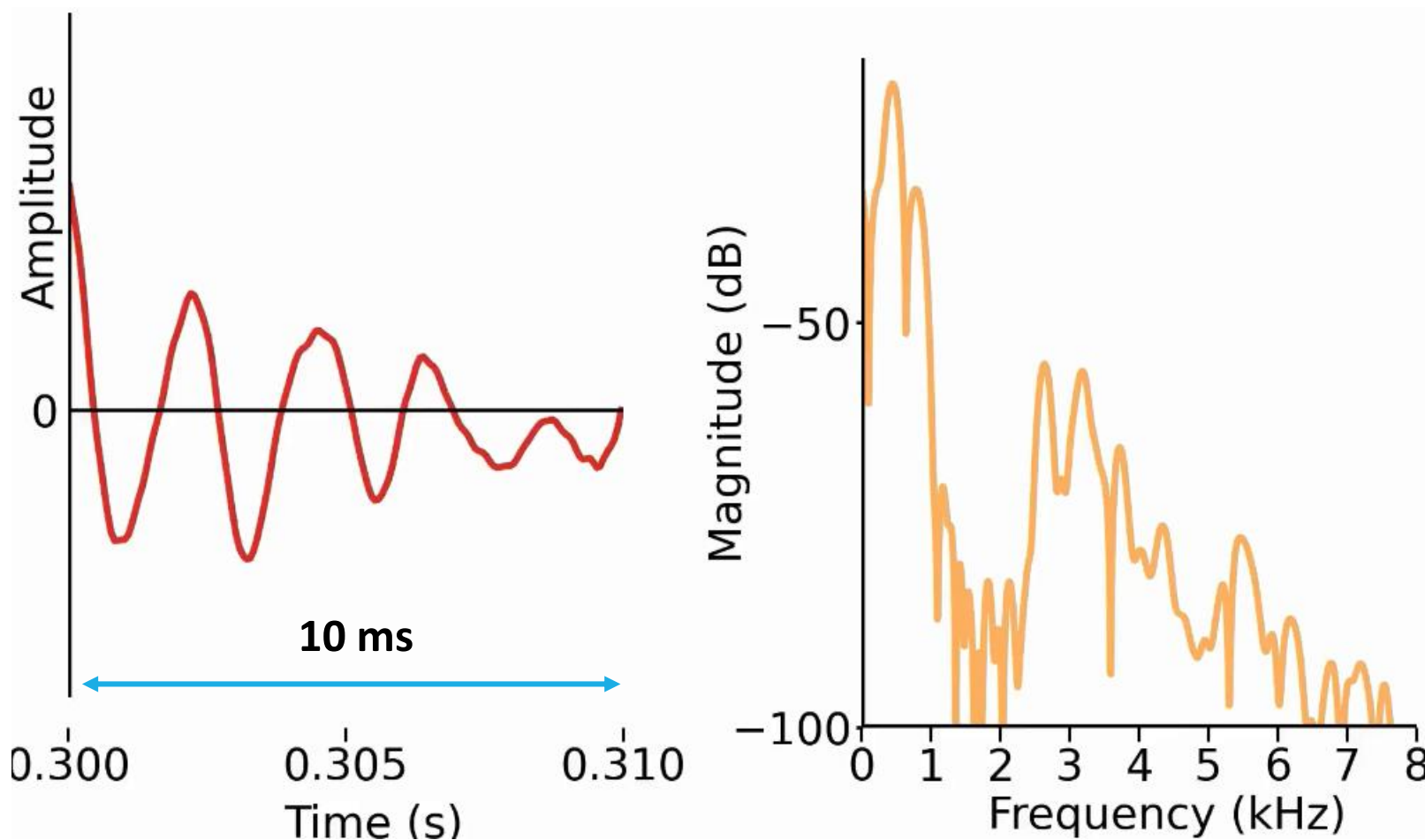
❖ افزایش طول پنجره آنالیز:

- ❖ فرضا در مثال قبل، طول پنجره آنالیز از ۰.۰۱ به ۰.۰۲ ثانیه افزایش یابد، یعنی ۲ برابر:
- ❖ کمترین فرکانس سینوسی که دقیقا در ۰.۰۲ ثانیه قرار گیرد، فرکانسش ۵۰ هرتز است.
- ❖ از طرفی سیگنال های پایه در فاصله فرکانسی یکسان قرار گرفته اند، یعنی دومین سیگنال پایه فرکانسش ۱۰۰ هرتز است و به همین ترتیب ۱۵۰ هرتز برای سومی و الی آخر!
- ❖ البته حداکثر فرکانس می تواند فرکانس نایکویست باشد (در اینجا مثلا ۸۰۰۰ هرتز)

با این توصیف تعداد سیگنال های پایه با دو برابر شدن طول پنجره چه تغییری کرد؟



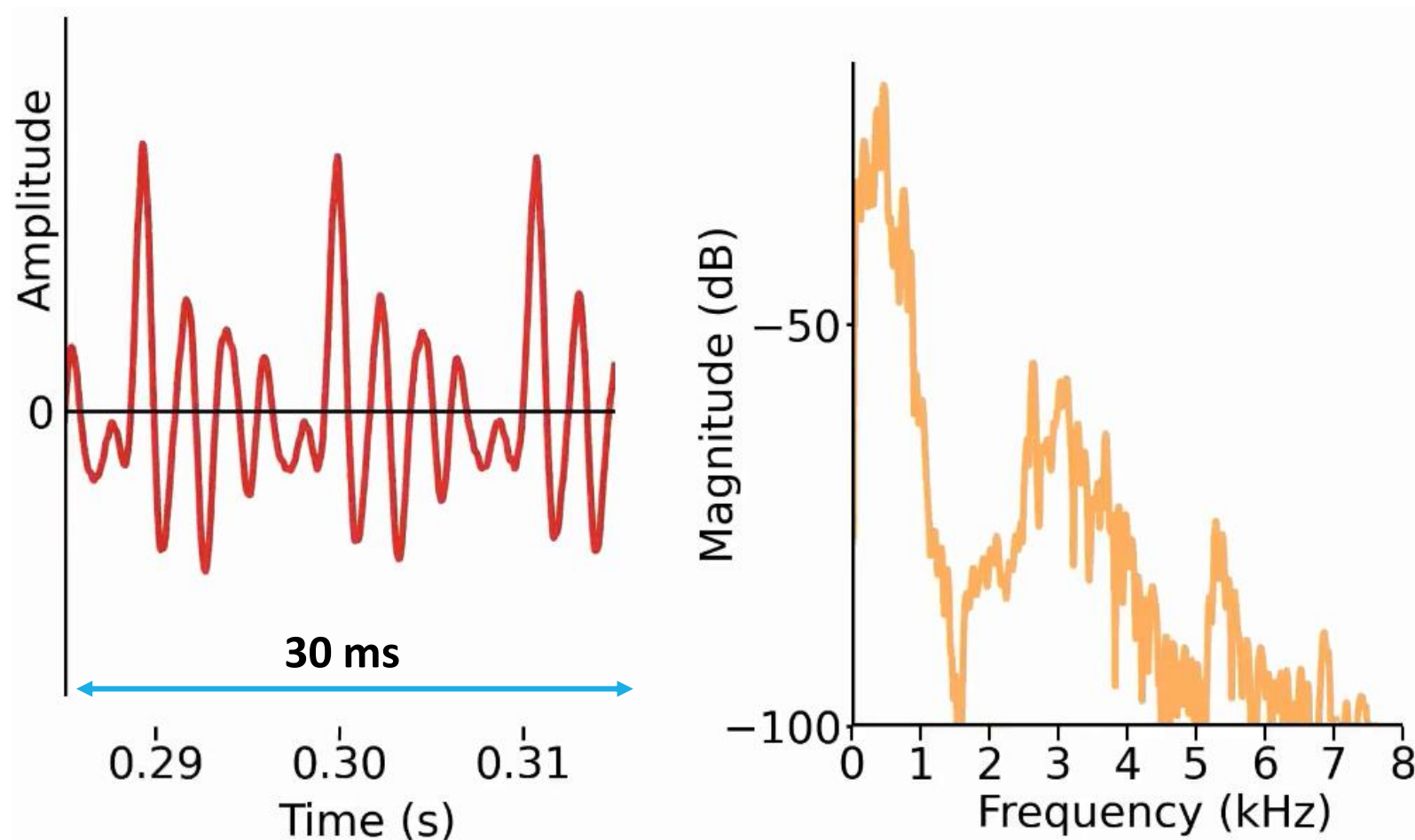
تحليل حوزه فرکانس - اثر افزایش طول پنجره آنالیز



Hamidreza Baradaran Kashani



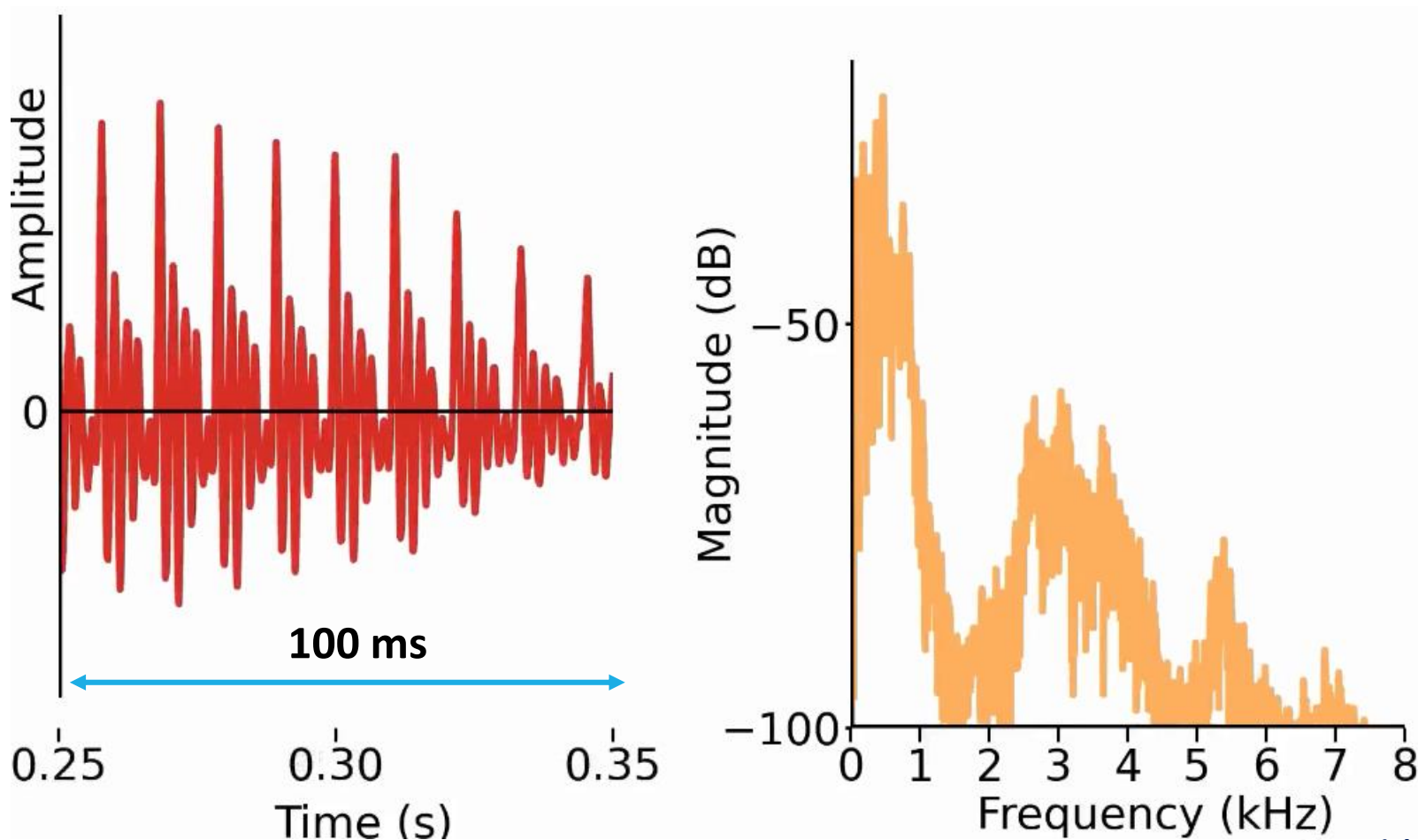
تحليل حوزه فرکانس - اثر افزایش طول پنجره آنالیز



Hamidreza Baradaran Kashani



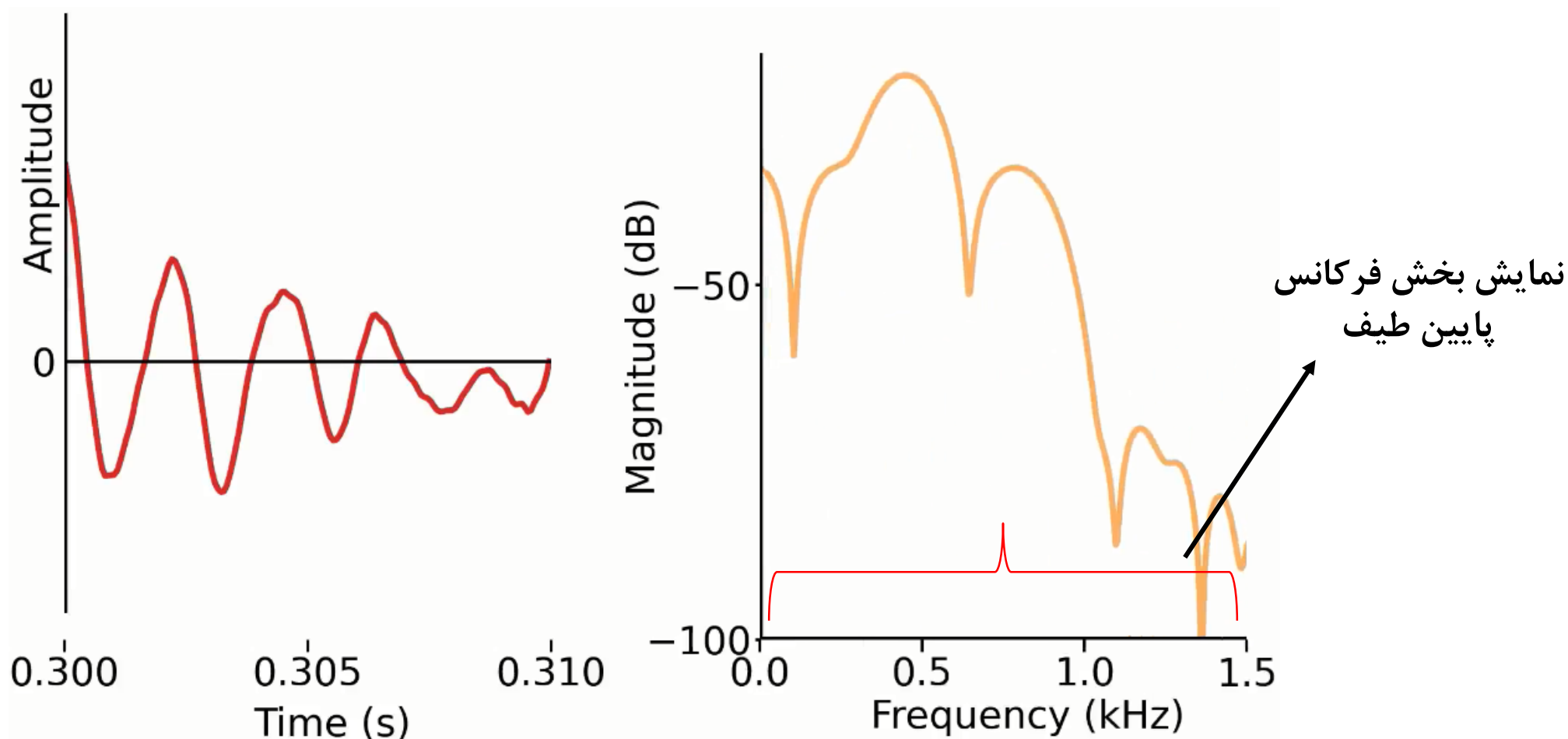
تحلیل حوزه فرکانس - اثر افزایش طول پنجره آنالیز



Hamidreza Baradaran Kashani



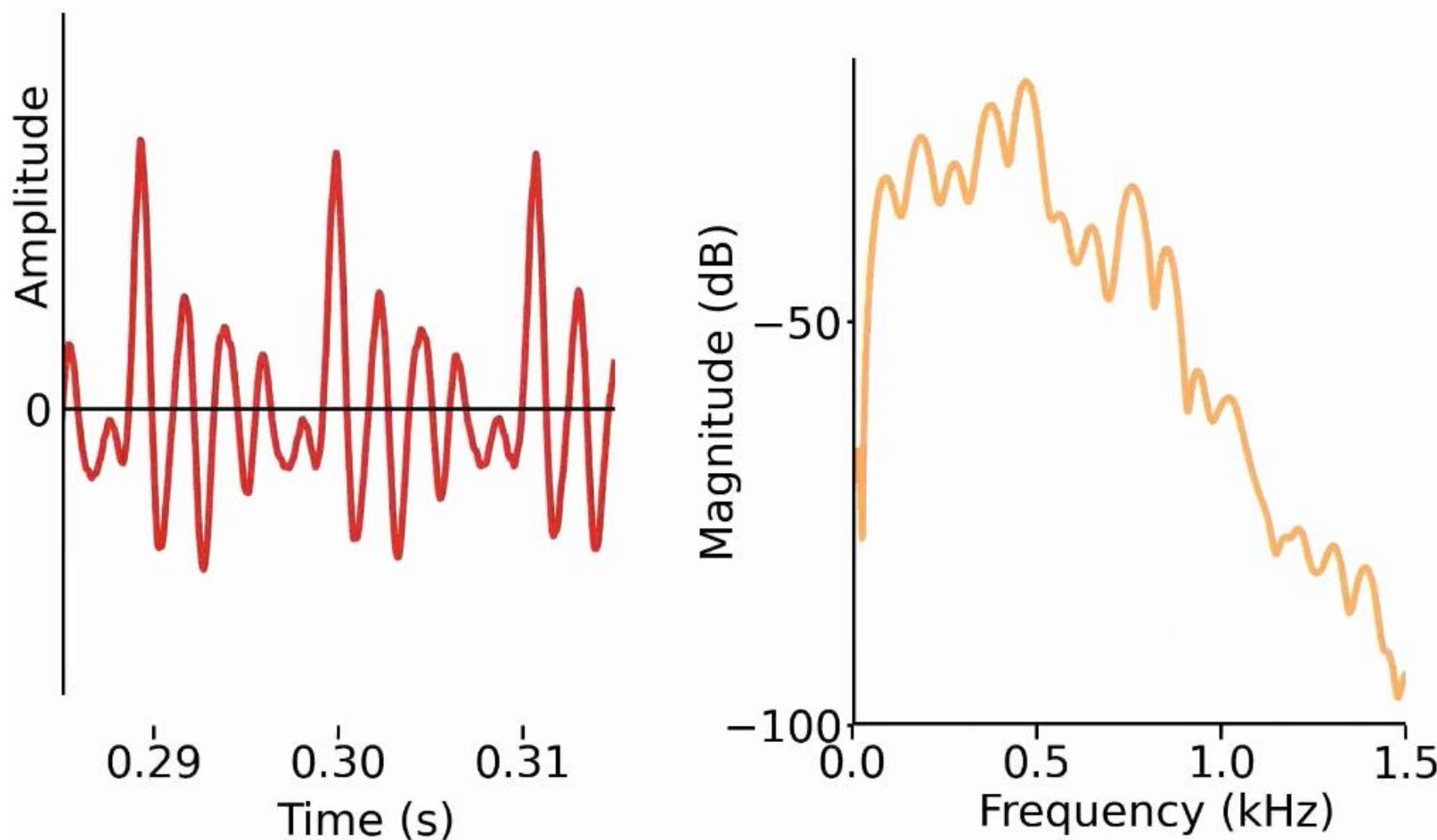
تحلیل حوزه فرکانس - اثر افزایش طول پنجره آنالیز



Hamidreza Baradaran Kashani



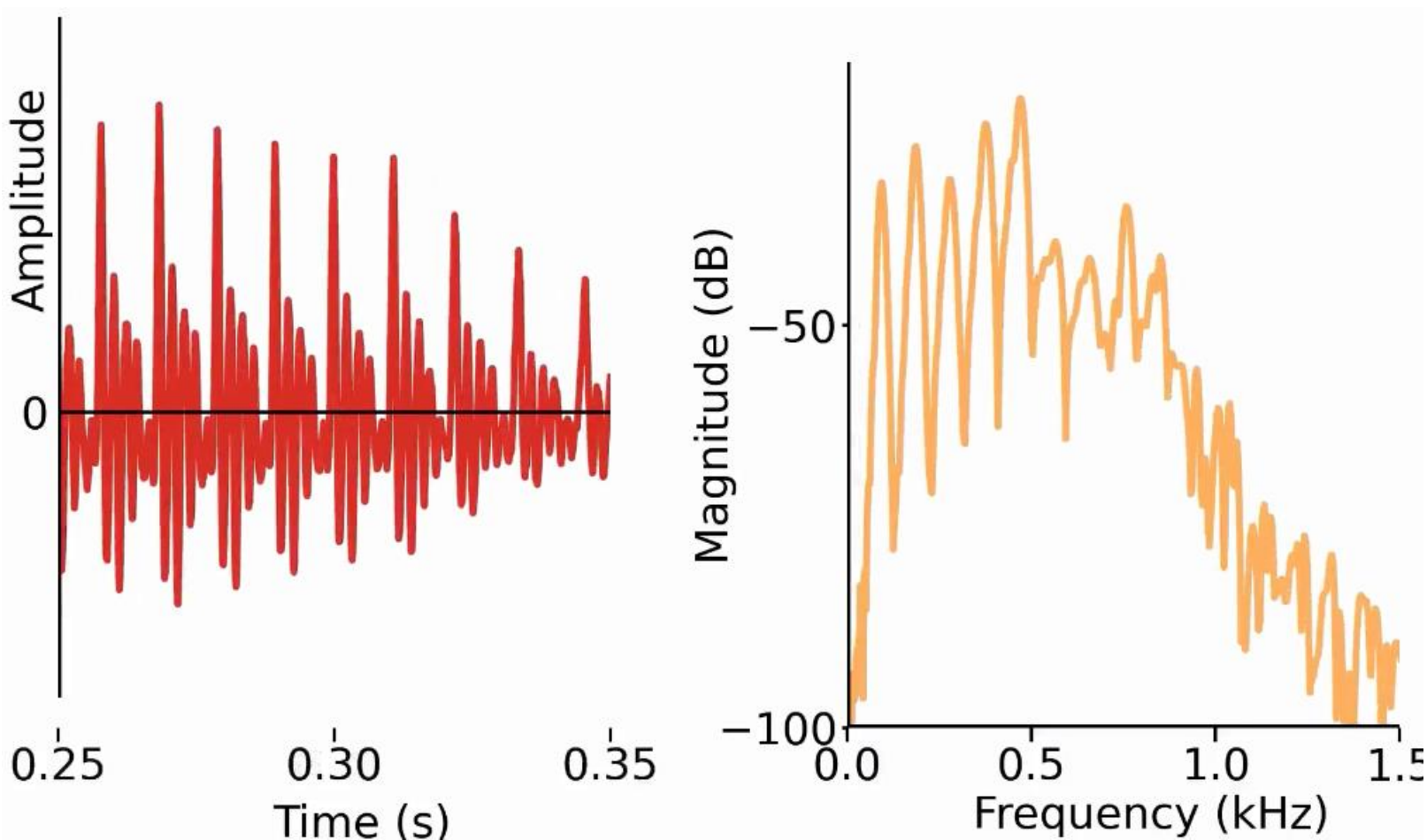
تحلیل حوزه فرکانس - اثر افزایش طول پنجره آنالیز



Hamidreza Baradaran Kashani



تحلیل حوزه فرکانس - اثر افزایش طول پنجره آنالیز

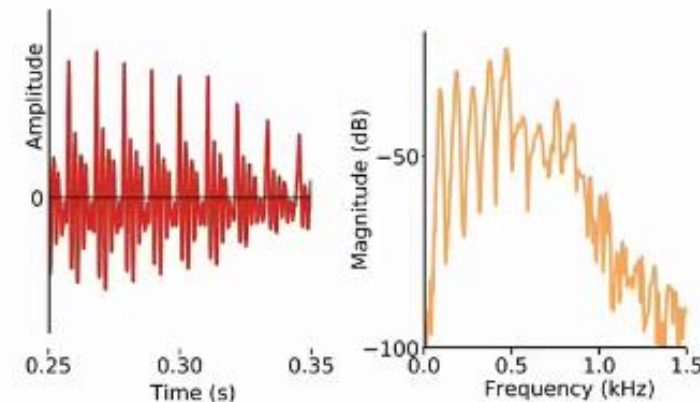
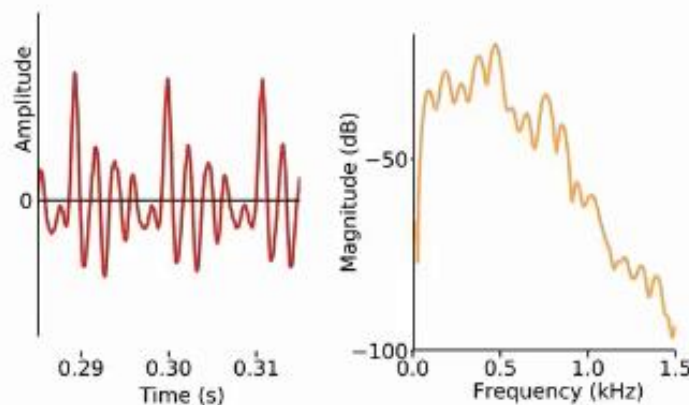
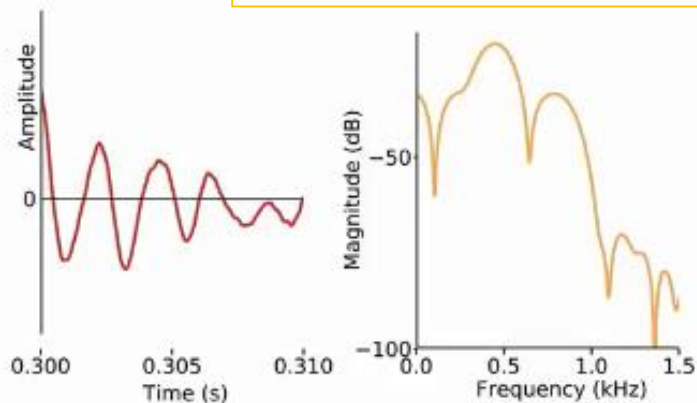


Hamidreza Baradaran Kashani



تحلیل حوزه فرکانس - اثر افزایش طول پنجره آنالیز

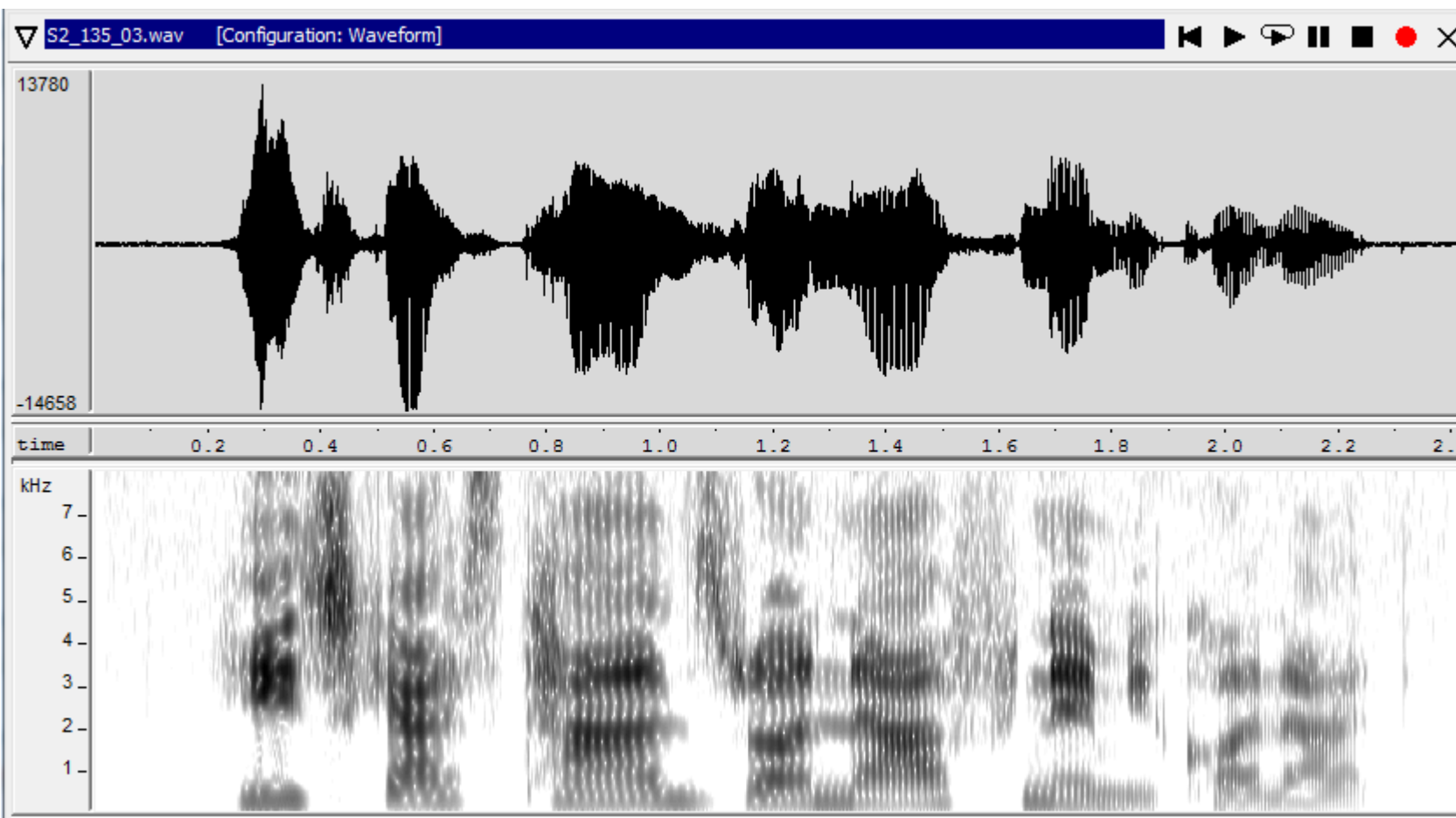
فریم آنالیز بزرگتر = تعداد سیگنال های پایه بیشتر = رزولوشن فرکانسی بیشتر



Hamidreza Baradaran Kashani



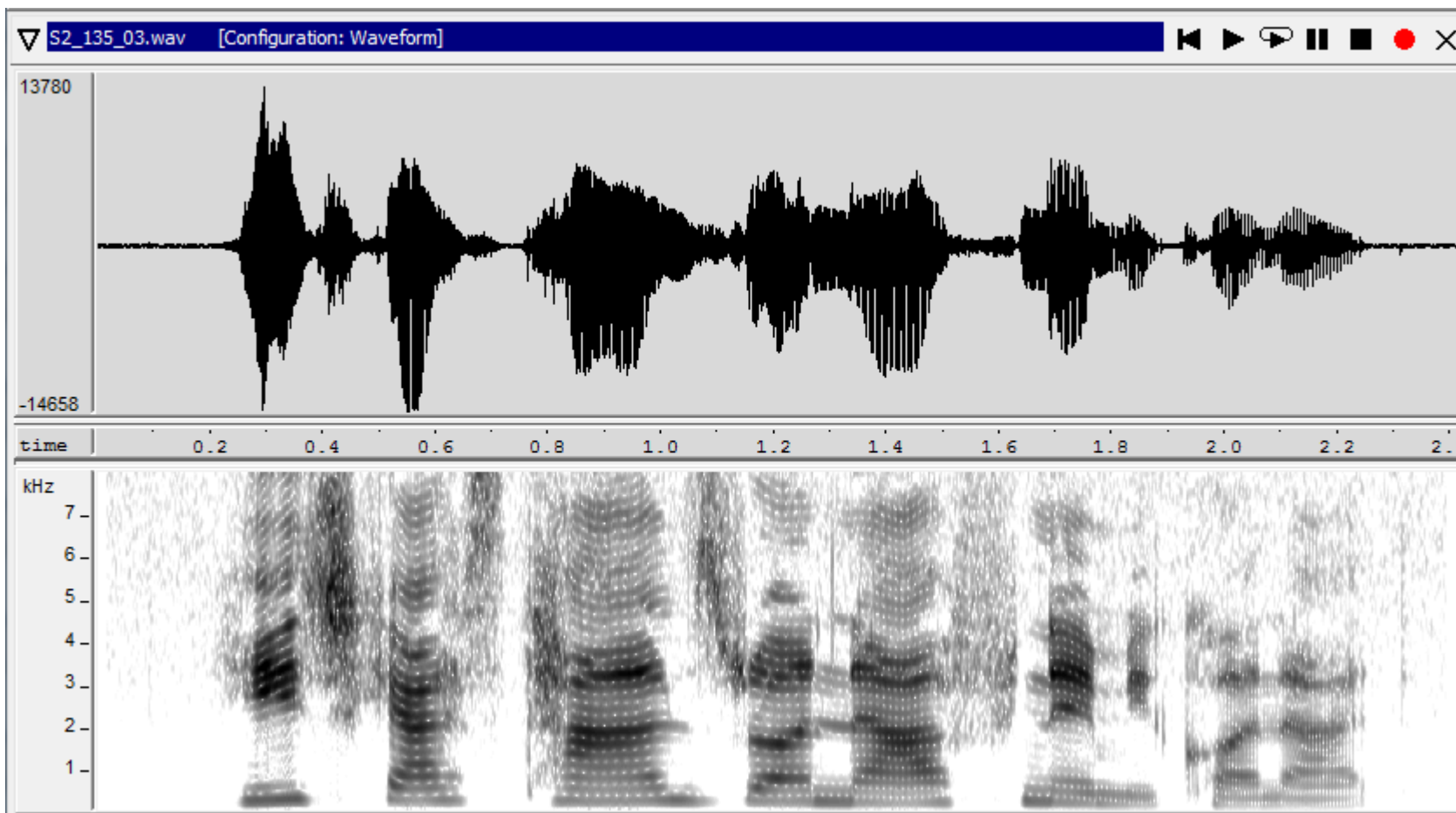
تحليل حوزه فرکانس - اثر افزایش طول پنجره آنالیز



Hamidreza Baradaran Kashani



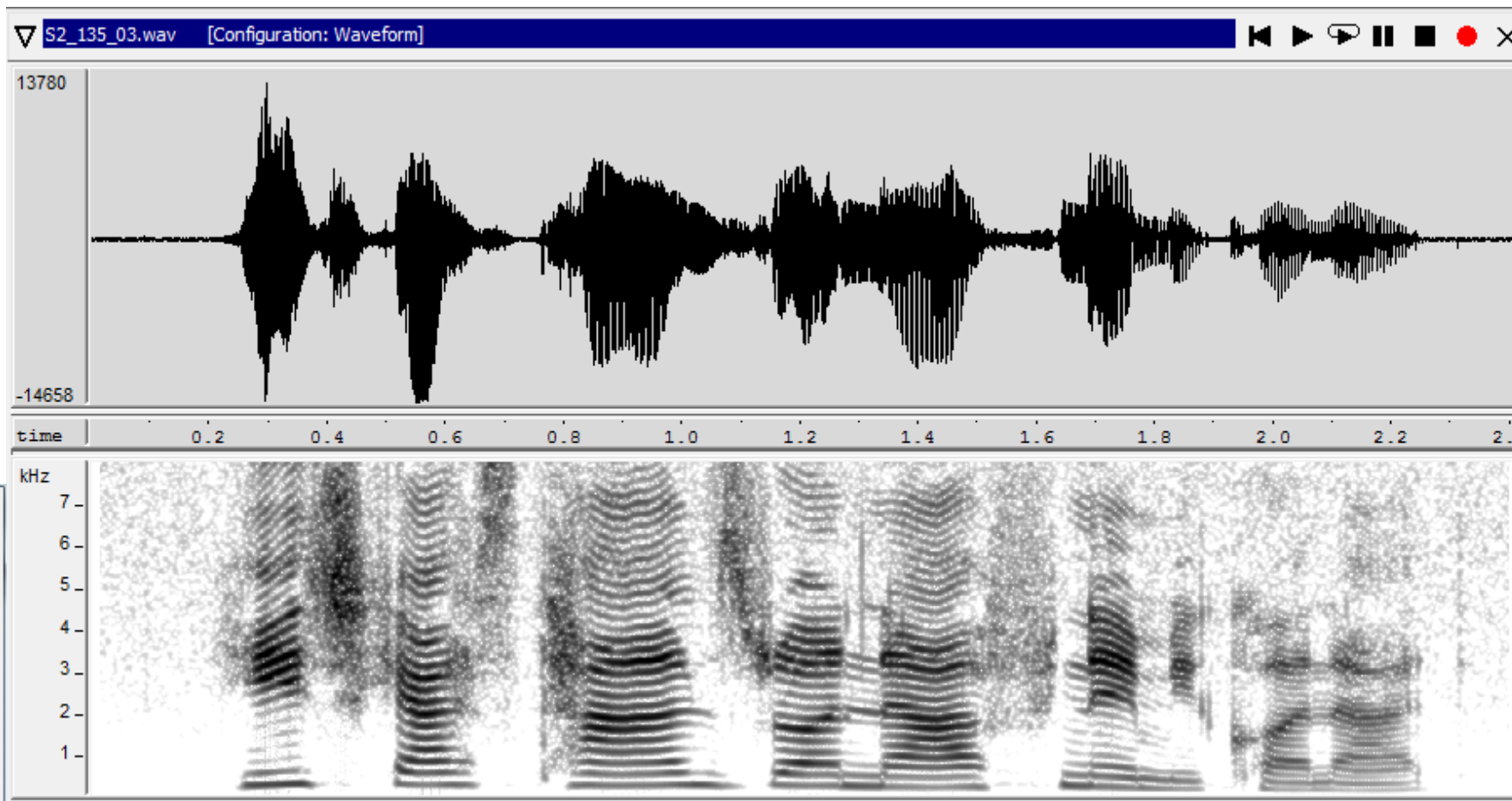
تحلیل حوزه فرکانس - اثر افزایش طول پنجره آنالیز



Hamidreza Baradaran Kashani



تحليل حوزه فرکانس - اثر افزایش طول پنجره آنالیز



Hamidreza Baradaran Kashani



با تشکر از اساتید و همکاران گرامی:
آقای دکتر همایونیپور
آقای دکتر کبودیان