

强震记录的采样与插值研究

谢礼立 胡成祥

(中国科学院工程力学研究所)

摘 要

本文从频域和时域两方面对强震记录的采样和插值过程进行了研究。研究中将采样与插值当作一个信号转换系统,通过数值计算,求得和各种常用的采样和插值过程相对应的传递函数。通过对传递函数的进一步研究表明:采样和插值方式对强震记录数据处理的结果有重要的影响,发现采样过程相当于一个低通滤波器,可以滤除信号中的某些高频信息,而插值过程犹如一个高频噪声源,会在数字记录中引入某些伪高频分量。分析结果还表明:在相同的采样密度下,不等距采样方式具有较高的精度,而等间距采样方式能给出更宽的平坦的频率特性曲线;抛物线插值给出的传递函数结果要比线性插值的结果更精确,这对恢复数字信号中的高频成分具有重要的意义。

一、引 言

目前世界上已积累约2000余条重要的强地震加速度记录。随着强震台阵的迅速扩大,不难预见,今后强震记录积累的数量和速度将不断增加。现今各国使用的强震仪虽然型号繁多,种类不一,但所提供的记录无非是二种形式,即模拟图形记录和数字记录,前者经过人工读数或在半自动化、自动化的读数仪上经数字化可变成具有各种间距的离散数字记录,后者则可直接提供等间距的数字记录。人们在从这些等间距或不等间距的数字化记录中提取所需要的地震或抗震的重要信息时,一般总还须作进一步的数据处理,例如记录的调整(包括记录零线的调整,时标不均匀的调整,光滑化等)、校正(零线校正和仪器频响特性的校正)以及有关的常规计算(如积分计算求速度和位移、反应谱与频谱计算)。在所有这些计算过程中都要反复地运用插值运算。例如在现行加速度记录的数字分析处理方法中,一般较多地采用线性插值计算的方法^[1],它具有运算简单,精度可加控制等优点,除此以外也有人建议采用抛物线插值,带加权函数的非线性插值等方法^[8]。

插值方法是一种比较成熟的古典近似计算方法,人们对它已有较完美的研究^[2],但值得指出的是,最近在地震工程和地球物理学的研究中,对插值问题的研究又出现了新的活跃趋势^{[1]、[3]、[4]、[5]、[6]、[7]}。着重研究和讨论的问题集中于:1.从频率域上来研究并估计插值方法对真实信息产生的干扰影响,以及这种影响的传播规律;2.设计出一种新的插值

方法或插值函数,使其所产生的干扰尽量减小。人们对地震动过程的研究已不能只满足于单纯地在时间域上的了解,许多新的课题,例如地震动的反演,结构系统的验证,以及震源过程的反演等问题都希望能了解地震动信息的精确结构,因而也就要求搞清一切干扰(包括插值过程)在频率域上的特征。实践经验证明,插值效果的好坏还往往和原始数据的采样方式有密切的关系,一种插值方法和一定的采样方式相配合往往能取得更佳的效果。

本文利用数字计算方法,对一个已知的函数图形,按等间距和不等间距两种采样方法得到几份具有不同采样密度的离散数字记录,分别按线性和抛物线方法插值后,得到了对应不同插值和采样方式的四组传递函数。从这些传递函数的特征可清楚地表明不同插值和采样方法的频率特性,从而可进一步估计来自插值方法的误差。此外,还对人工模拟的地震记录按插值近似计算和精确解求得的积分位移以及反应谱曲线进行了对比。研究结果表明:不同插值和采样方法对地震信号的数据处理有重要影响。一般说来,采样过程能滤去信号中的高频成分,插值过程却如高频噪声源,会在信号中引入伪高频分量。分析结果表明,在一般情况下,不等间距采样方法具有较高的计算精度,而等间距采样方式能给出更宽的频率特性。

二、数值计算简介

为了比较和估计线性与非线性插值计算在不同的采样方式和采样密度下,所具有的复现记录信号的能力,可以把一定的采样和插值计算过程作为一个信号转换系统(见图1)来研究,即将原始记录波形看作该系统的输入函数 $f_0(t)$,而把经采样和插值计算后得到的离散数字或连续图形看作系统的输出函数 $f(t_i)$ 或 $f(t)$ 。因此插值计算过程的特性,特别是在频率域上的特性可以通过求出该转换系统的频率特性或传递函数来获得。为此可以分别计算输入信号和输出波形的Fourier谱,进而求出后者对前者的比值,即可求得系统的传递函数。

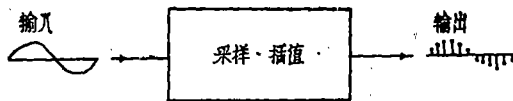


图 1

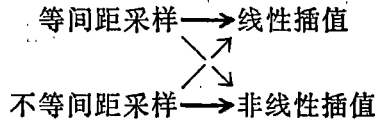
Fig 1

考虑到大多数地震加速度记录的波形都可以近似地看作是由许多不同振幅和频率的半正弦波组合而成。为了简单起见,在研究插值、采样方法的影响时,可以先从一个简单的波形出发,把输入函数取作一个振幅为1,周期为 T_0 的单个完整周期的正弦波形,即取输入函数 $f_0(t)$ 为:

$$f_0(t) = \begin{cases} \sin \frac{2\pi}{T_0}t & 0 \leq t \leq T_0 \\ 0 & t < 0, t > T_0 \end{cases} \quad (1)$$

对输入波形的不等间距采样完全由主观决定，认为当采用直线或抛物线来连接诸相邻的采样点时，能够最佳地逼近输入波形。为了防止读数误差，不使它对所考虑的问题产生任何干扰，采样点的座标值不是由半自动化的读数机读出，而是直接根据图形上选定的点，确定其横坐标，再由计算机的内部标准函数给出采样点的纵标值。因此这样得到的采样点数据，除了包括影响极小的计算机的舍入误差外，不包含任何数字化误差。

在等距采样时，采样间隔 Δt 分别取为 $T_0/3$ 、 $T_0/4$ 、 $T_0/6$ 、 $T_0/7$ 、 $T_0/10$ 、 $T_0/20$ 和 $T_0/40$ 七种，对于不等距采样，采用相同的采样密度，即 $N = 3, 4, 6, 7, 10, 20$ 点/周，图2中给出了几种典型的采样图。根据采样得到的离散座标值，分别采用线性和抛物线插值方法进行插值，得到以下四种组合的不同采样方式和插值方法的离散数字记录：



分别计算了它们的富里叶振幅谱 $F(\omega)$ 。

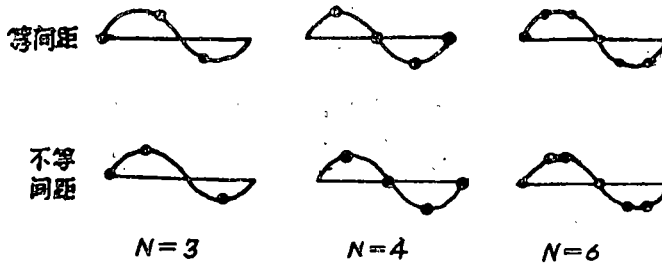


图2 几种典型的采样方式

Fig. 2 Some typical sampling methods.

三、富里埃谱和传递函数

对于离散的数字化记录，按下式计算其富里埃变换：

$$F(kf_0) = \frac{T_0}{6M} \left\{ 4 \sum_{j=0}^{M-1} \sin \left[(2j+1) \frac{\pi}{M} \right] e^{-i \frac{(2j+1)K\pi}{M}} + 2 \sum_{j=1}^{M-1} \sin \left(\frac{2j\pi}{M} \right) \cdot e^{-i \frac{2jK\pi}{M}} \right\}$$

式中：

$f_0 = \frac{1}{T_0}$ 为所研究的正弦波的周期。

$M = \frac{T_0}{2H}$ ，式中H为计算步长。

$i = \sqrt{-1}$

图3、图4给出了一部分计算结果，即当采样密度 $N = 3$ ， $N = 6$ 点/周时，按不同插

值方法算得的部分富里埃谱图。图中还同时画出了所研究函数(见式1)的精确富里埃谱:

$$F_0(f) = \left| \frac{\sin \pi f T_0}{\pi T_0 (f^2 - f_0^2)} \right| = \left| \frac{\sin \pi f / f_0}{\pi (f^2 / f_0^2 - 1)} \cdot T_0 \right| \quad (2)$$

图中的横坐标是以无量纲参数 f/f_0 表示的。

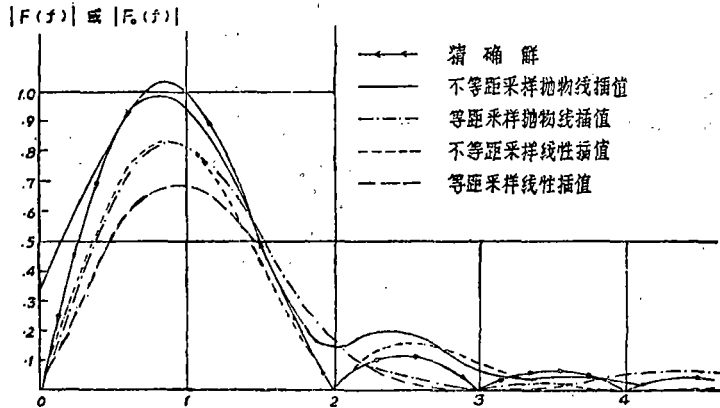


图3 $N = 3$ 点/周时的富里埃振幅谱

Fig 3 Fourier amplitude spectrum ($N = 3$).

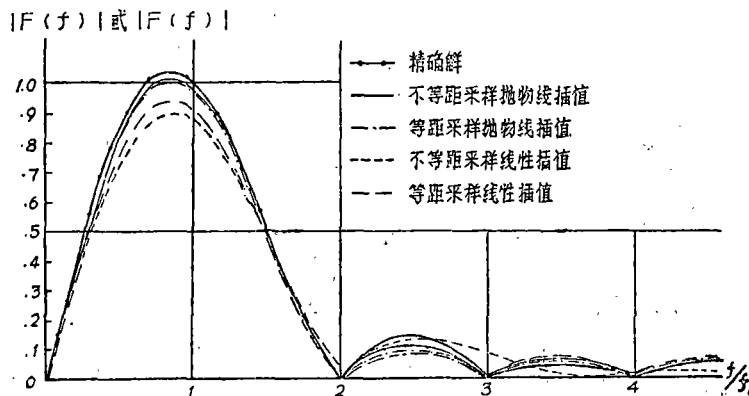


图4 $N = 6$ 点/周时的富里埃振幅谱

Fig 4 Fourier amplitude spectrum ($N = 6$).

根据图3、图4给出的曲线不难发现,当 f/f_0 大致等于0.8, 2.4, 3.4及4.5时,谱图中都有一个峰值与之对应。现将这些峰值与精确的峰值之比分别按插值方式和采样密度值列于表1——表4中,在表中还同时给出了两种平均值:平均值I,是对同一种插值方法,在不同采样密度下的误差平均值,它可以给出不同插值计算之间的精度比较;平均值II,表示不同插值方法在同一种采样密度下的平均误差,它表示不同采样密度之间的精度比较。从图3、图4可见,由于采样和插值的方式不同,所得信息在频率域上对原始波形的复原程度是不相同的,总的来说;

(1)比较频谱曲线上峰值处的精度可知有如下特点:就插值方法而言,抛物线方法给出

的精度总要比线性插值高；就采样方法而言，情况就比较复杂。当采样率低时，不等距采样方法的精度要明显地比等距采样方法高，这一点对线性插值来说更为明显。对此，图3可不难看出。但在中等采样密度时，情况恰恰相反，等距采样给出的精度要比不等距采样高（图4）。在更高的采样率的情况下，两种采样方法给出的精度，虽然互有高低，但总的来说是比较接近的（图6）。由此可见，传统观点认为在相同采样密度下，不等距采样方法总比等距采样方法的精度高，只有在低采样率时才是正确的。至于在中等或高采样率时，这种观点就不确切了。表1——表4给出了定量的结果。

(2) 不论采用哪种插值和采样方式，都有如下特点：(i) 频谱曲线在低频段的精度要高于高频段；(ii) 采样密度越高，相应的精度也就越高。

表1

$$\frac{F_0(0.8) - F(0.8)}{F_0(0.8)}$$

N	等距线性%	不等距线性%	等距抛物线%	不等距抛物线%	平均值 I %
8	71.2	-18.6	54.0	-50.8	48.65
4	18.6		14.2		16.4
6	4.7	0.96	0.3	-12.9	4.7
7	-4.0		0.4		2.2
10	1.8	-4.2	-0.9	-1.7	2.1
20	-0.4	-3.1	-1.3	-1.5	1.5
平均值 I	16.78	6.71	11.85	16.73	

表2

$$\frac{F_0(2.4) - F(2.4)}{F_0(2.4)}$$

N	等距线性%	不等距线性%	等距抛物线%	不等距抛物线%	平均值 I %
8	71.2	-18.6	54.0	-50.8	48.65
4	18.6		14.2		16.4
6	4.7	-0.96	0.3	-12.9	4.7
7	-4.0		0.4		2.2
10	1.8	-4.2	-0.9	-1.7	2.1
20	-0.4	-3.1	-1.3	-1.5	1.5
平均值 I	16.78	6.71	11.85	16.73	

表3

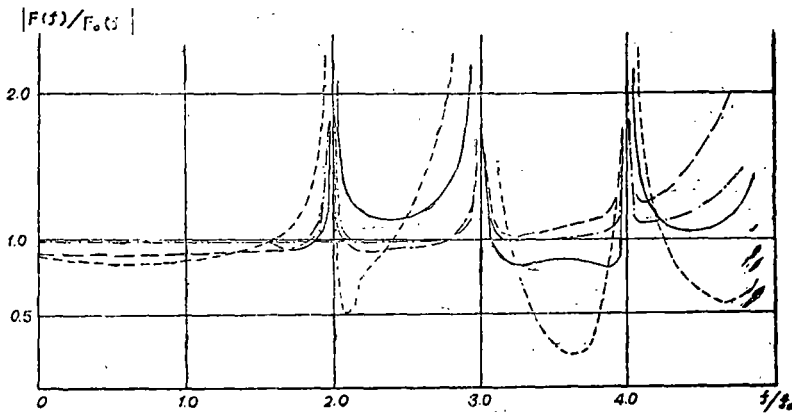
$$\frac{F_0(3.4) - F(3.4)}{F_0(3.4)}$$

N	等距线性%	不等距线性%	等距抛物线%	不等距抛物线%	平均值 I %
8	90.3	60.8	69.6	37.3	64.5
4	61.0		28.3		44.7
6	-4.3	57.7	-2.1	15.8	20.0
7	-8.7		-4.0		6.4
10	0.2	14.5	-1.9	0.8	4.3
20	-1.6	5.8	-2.3	-1.4	2.7
平均值 I	27.68	34.7	18.0	13.83	

表4

$$\frac{F_0(4.5) - F(4.5)}{F_0(4.5)}$$

N	等距线性%	不等距线性%	等距抛物线%	不等距抛物线%	平均值 I %
8	-58.5	78.0	60.9	48.3	61.5
4	78.0		49.4		63.7
6	-60.7	40.0	-19.5	-0.8	30.25
7	-28.0		-8.0		-18.0
10	-2.7	-1.8	-3.8	-3.8	3.03
20	-3.7	0.73	-4.4	-4.0	3.2
平均值 I	38.6	30.13	24.33	14.2	

图5 $N = 6$ 点/周时的传递函数Fig 5 Transmit function ($N = 6$).

为了对不同采样、插值方法进行比较,计算并比较了它们的传递函数 $F(f)/F_0(f)$ 。为了示例,仅在图5中给出了不同插值方法在同一种采样密度下的传递函数曲线,这些曲线给出如下特征:

(1) 给出的传递函数或频率特性曲线,一般来说在低频段(指 $f/f_0 = 0 \sim 2.0$ 的范围)比较平坦,而在高频段(当 $f/f_0 > 2.0$ 时)频率特性的幅值有较大的起伏,但随着采样密度的增加,平坦部分延伸得就越长。

(2) 当 $f/f_0 = 2, 3, 4, \dots$ 等整数(1除外)时,传递函数出现了奇异点,在这些位置上,传递函数趋向于无穷大,这是由于正弦波形的精确富里叶谱在这些座标点的函数值为零的缘故。由图可见,若传递函数曲线在奇异点处的宽度较小,它的平坦部分就较长,经插值计算后的失真就较小。反之,当奇异点处的宽度越大时,传递函数曲线的平坦段就较短,经过插值计算以后的失真度就较大。

(3) 就频率特性来说,不论采用何种插值方式,等距采样的频率特性要比不等距采样的有更宽的平坦部分,这一点从图5可见,例如在 $N = 6$ 情形下,具有平坦段的频率范围前者约为后者的2倍(奇异点处例外)在同样的采样方式情况下,则抛物线插值给出的频率特性要较直线插值更好。

(4) 不论采用什么样的插值方法或采样方法,他们的频率特性曲线都出现:(i)奇点;

(ii)在高频段的传递函数值要远大于 1。由此可见，插值方法会在奇异频率处或高频处引入某种虚假的振幅成分。在许多文献中已经指出，由于采样密度的限制，采用一定插值方法后难免会滤掉信号中的高频成分，使采样的作用犹如一个低通滤波器。本文所要强调的另一个侧面是，插值过程的作用却相当一个高频噪声源，在高频处会引入具有相当大振幅的伪高频分量。

四、 $F(T_0)/F_0(T_0) \sim N$ 曲线

这里将根据计算结果给出一个具有实用意义的 $F(T_0)/F_0(T_0) \sim N$ 曲线（见图 6）。图中的横座标是采样密度率，纵座标是由插值近似计算得的、与正弦波周期 T_0 相对应的富里埃振幅谱的谱值 $F(T_0)$ 与精确谱值 $F_0(T_0)$ 之比。图中对应 $N = 2$ 处的比值是根据理论计算得到的一个平均值^[8]。图中清楚地表明，在同样的采样密度下，抛物线插值要比线性插值方法具有更精确的结果，特别在较小的采样密度下，如 $N = 3$ 或 $N = 4$ 时，两者差别可达 30% 这一点对强震记录中的高频信息尤为重要。在一般情况下，由于数字化设备能力的限制，难以使高频地震信号获得足够的数字化采样密度，而导致高频数字信号的失真。但是研究结果表明，当采样密度即使只满足 $N = 3$ 点/周的情况下，采用抛物线插值方法也能使富氏谱值的精度达 93%。如采用线性插值，这时的精度只有 60%。举例来说，当强震加速度记录的平均数字化密度为 100 点/秒时，对于 35 赫兹的高频地震分量来说，平均采样密度大致为 3 点/周。如采用抛物线插值的话，离散数字记录仍具有很精确的信号复现能力。

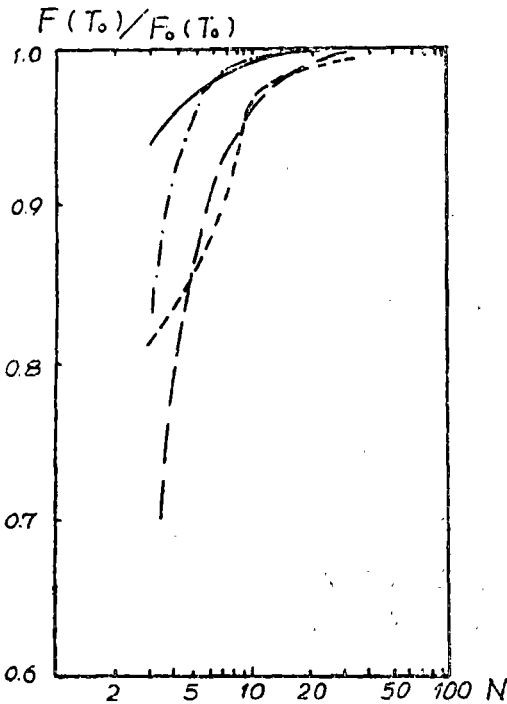


图 6 $F(T_0)/F_0(T_0) \sim N$ 曲线
Fig. 6

五、插值方法对积分位移与反应谱计算的影响

计算中采用了 Bogdanoff, Goldberg 和 Bernared^[8] 建议的一个随机加速度函数，它取如下表达式：

$$\begin{cases} \ddot{u}_g(t) = 0 & t < 0 & (3) \\ \ddot{u}_g(t) = \sum_{j=1}^J \xi_j t e^{-\theta_j t} \cos(\omega_j t + \varphi_j) & t \geq 0 & (4) \end{cases}$$

式中， ξ_j, θ_j 是正实数，并有 $\omega_1 < \omega_2 < \omega_3 < \dots < \omega_j$ ， $\varphi_1, \varphi_2, \dots, \varphi_j$ 都是 J 个相互独立的均匀地分布在 $(0, \pi)$ 区间中的随机实变量。

只要使级数 (4) 中项数尽量增多，便可获得最接近实际地震的加速度记录。Parmelee 等人建议^[9] 只要取十项便可足够精确地表示一个地震加速度记录，于是文中采用了如下的地震加速度记录：

用了如下的地震加速度记录：

式中的 ω_j 与 φ_j 的值列于表5中:

$$\begin{cases} \dot{u}_g(t) = 0 & t < 0 \quad (5) \\ \dot{u}_g(t) = 0.50te^{-0.333t} \sum_{j=1}^{10} \cos(\omega_j t + \varphi_j) & t \geq 0 \quad (6) \end{cases}$$

表5 地震加速度函数的频率及相角
Table 5 Frequency and phase angle of the seismic acceleration function

j	ω_j (弧度/秒)	ψ_j (弧度)
1	6.0	3.7663
2	8.0	1.3422
8	10.0	4.8253
4	11.15	0.2528
5	12.30	4.5204
6	13.25	1.8834
7	14.15	1.3320
8	16.20	1.7852
9	17.35	0.1517
10	19.15	2.4881

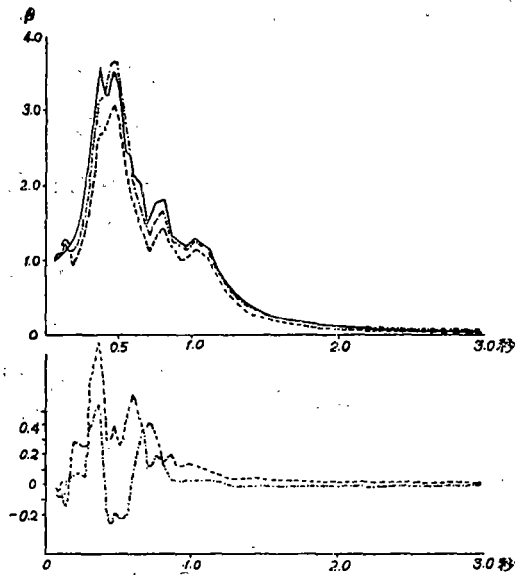


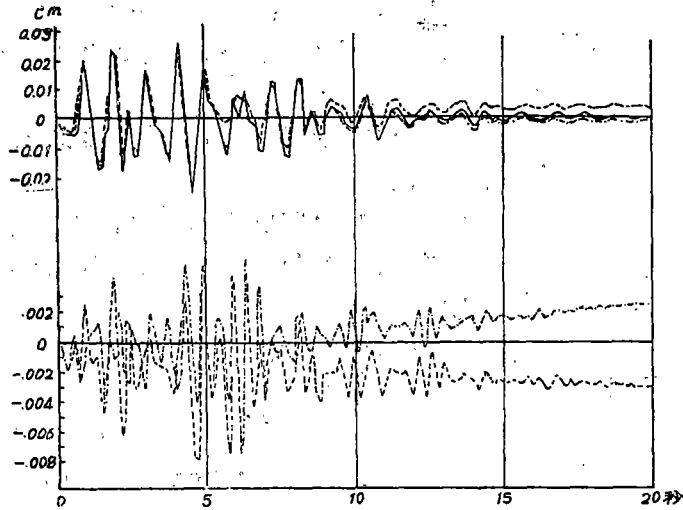
图7 $\Delta t = 0.2$ 秒采样时的反应谱曲线 ($\theta = 0.05$)

Fig 7 Response spectrum at sampling time ($\Delta t = 0.2$ sec.).

图中,

- 精确解
-线性插值反应谱
- 抛物线插值反应谱

在对人为地震记录计算时,先按等间隔 $\Delta t = 0.005$ 秒对函数(5)和(6)进行采样。以这样间隔按精确法^[10]计算的反应谱和按梯形求和公式算得的两次积分可作为反应谱与积

图8 $\Delta t = 0.2$ 秒采样时积分位移Fig 8. Integral displacement at sampling time ($\Delta t = 0.2$ sec.).

分位移曲线的精确解。然后再按间隔 $\Delta_1 = 0.1$ 秒, $\Delta_2 = 0.2$ 秒和 $\Delta_3 = 0.3$ 秒的数据, 分别用抛物线插值和线性插值方法算得的反应谱曲线作为插值近似计算的反应谱, 用同样插值方法求得的两次积分作为插值近似计算的位移。比较这些结果的精确解和近似解, 便可了解插值计算的影响。在计算反应谱时一律采用阻尼比 $= 0.05$ 。

图7给出了对应一种采样间隔的计算结果, 图中同时列出了精确解。由此可见, 精确的反应谱曲线在 $T = 0.3 \sim 0.5$ 秒之间, 包含两个明显的峰值。当采样间隔 $\Delta = \Delta_1 = 0.1$ 秒, 两种插值计算都能给出相应的峰值位置, 但按抛物线插值计算的结果要比线性插值方法有更高的精度, 当采样间隔 $\Delta = \Delta_2 = 0.2$ 秒时(图7), 两种近似计算都只能给出一个峰值点了, 但抛物线法给出的结果为精确解的93%, 而线性插值法给出的峰值只为精确解的86%, 在这些图的下方还同时画出了按这种方法算得的反应谱与精确解的差。

图8中给出了采样间隔为0.2秒时的位移积分的计算结果, 结果仍然表明, 采用抛物线插值方法仍有较高的精度, 当 $t = 20$ 秒时, 精确解的位移值等于零, 按抛物线法和线性插值法计算的位移分别为峰值位移的8%和14%, 后者造成的漂移位移约为前者的二倍左右。一方面对峰值位移来说, 抛物线方法可达精确解的93%, 线性插值方法只及精确解的80%; 另一方面, 由于采用了插值方法还引起积分位移曲线的漂移。从这个角度来说, 抛物线方法引起的漂移量也要小于线性方法造成的漂移量。但需要指出的是, 对目前通常采用的数字加速度记录的间距0.02秒和0.01秒来说, 两种插值方法对反应谱或积分位移计算的差别都是很小的, 最大的误差约为1%左右。

六、结 论

经过上述数字计算比较, 可得如下看法:

(1) 就富里埃振幅谱峰值处的精度来说, 抛物线插值要比线性插值给出的精度更高。而

不等距采样与等距采样相比只是在低采样率时才有较高的精度, 而在中等或高采样率时, 两

者给出的精度互有高低,总的来说是接近的。

(2)就传递函数给出的频率特性曲线平直段的范围来说,等间距采样方法给出的平坦部分要比不等间距采样方法更宽一些。

(3)由于采用插值计算,对地震信号来说,犹如增加了一个高频噪声源,会在地震信号中引入高频噪声,导致高频段的信噪比降低。

(4)在同样的采样方式,采样密度下,不论是富里埃谱的峰值精度,频率特性曲线平坦段的长度,反应谱曲线的精度以及位移积分中峰值位移的精度和长周期的漂移程度,抛物线插值方法给出的结果都要比线性插值的结果为佳。

(5)文中给出的 $F(T_0)/F_0(T_0) \sim N$ 曲线,可为常规计算中选择采样方式和采样密度以及插值计算时,提供参考。

(1981年6月收到初稿)

(1982年7月10日收到修改稿)

参 考 文 献

- [1] Trifunac, M.D. and Lee, V., "Computer processing of strong-motion accelerograms" EERL73-03, oct, 1973, california institute of technology, pasadena, california.
- [2] 冯康等, "数值计算方法" 第一章, 国防工业出版社, 北京, 1978.
- [3] Arturo Arias and Horacio Sandoval, "Sampling and interpolation of strong motion accelerograms" 7th WCEE, Vol, II, Part II, 25-32, sept, 8-13, 1980, istanbul, turkey.
- [4] Wiggins, R.A. "Interpolation of Digitized Curves" Bull. Seism. Soc. Am. Vol. 66, No. 6, 1976.
- [5] Griffiths, L.J. and R. prieto-Diaz, "Spectral Analysis of Natural events using autoregressive technique", IEEE. Trans. Geoscience Electronics Vol. GE15, No. 1, 1977.
- [6] Bath, M. "Spectral Analysis in Geophysics", Elsevier Scientific publishing Compang Amsterdam-Oxford-Newyoxh, 1974.
- [7] Trifunac, M.D., "Zero Baseline Correction of Strong Motion Accelerograms", Bull, Seism, Soc, Am., Vol, 61, No. 5, 1971.
- [8] Bogdanoff, J.L., Goldberg, J.E., and Bernard, M.C., "Response of a Simple structure to a Random Earthquake-type Disturbance" Bull. Saism. Soc. Am. Vol, 51, No: 2, Apr. 1961. pp. 293-310.
- [9] P.A. parmelee, D.S. perelman etc., "Seismic Response of Structure-Foundation Systems", Journal of The Engineering Mechanics Division, ASCE, Vol. 94, NO. EM6, PP, 1295-1315. 1968.
- [10] Nigam, N.C. and Jennings, P.C.: "Digital Calculation of Response Spectra From Strong Motion Earthquake Records" EERL, California Institute of Technology, Pasadena, California, 1968.

STUDY ON SAMPLING AND INTERPOLATION OF STRONG MOTION DATA

Xie Lili Hu Chengxiang

(*Institute of Engineering Mechanics, Academia sinica*)

Abstract

The effects of the sampling and the interpolation of strong motion data on the corrected accelerograms and integrated displacements are investigated in both frequency and time domains, It is found that the sampling of a record can be considered as a low-pass filter which eliminates the high frequency contents in the record, but the interpolation of data does play a role of high frequency noise source which introduces high frequency noise into digital records. It is also found that an unequally spaced sampling model will give a more accurate information than an equally spaced sampling model at the same sampling rate. But the latter will result in a wider range of flat frequency response than the former. In comparison with the results given by the linear interpolation method, the parabolic one can generally provide with much better results, especially when a lower sampling rate is used.