PROBLEM-1: SOLUTION SUBMITTED BY GITA AYU GIVEN BELOW.

Gita Ayu
ISYE 6413
HW #6
Gt8651b@mail.gatech.edu

Problem 1

3^{5-2} design with D = ABC and E = ABC^2

a. Defining contrast subgroup: I = $ABCD^2$ = ABC^2E^2 = ABDE = CDE

The 3^{5-2} design has 27 runs which has total of 13 sets of aliased effects, shown below:

	I	= $ABCD^2$	= ABC^2E^2	= ABDE	= CDE^2
A^2	= A	= BCD^2	= BC^2E^2	= BDE	= AC^2D^2E
		= AB^2C^2D	= AB^2CE	= $AB^2D^2E^2$	= $ACDE^2$
B^2	= B	= ACD^2	= AC^2E^2	= ADE	= BC^2D^2E
		= AB^2CD^2	= $AB^2C^2E^2$	= AB^2DE	= $BCDE^2$
C^2	= C	= ABD^2	= $ABCE^2$	= ABC^2DE	= DE^2
		= ABC^2D^2	= ABE^2	= ABCDE	= CD^2E
D^2	= D	= ABCD	= $ABC^2D^2E^2$	= ABE	= CE^2
		= ABC	= ABC^2DE^2	= ABD^2E	= CD^2E^2
E^2	= E	= $ABCD^2E^2$	= ABC^2E	= ABD	= CDE
		= $ABCD^2E$	= ABC^2	= $ABDE^2$	= CD
A^2B^2	= AB	= CD^2	= CE	= DE	= ABC^2D^2E
		= ABC^2D	= ABCE	= ABD^2E^2	= $ABCDE^2$
A^2B	= AB^2	= BC^2D	= BCE	= BD^2E^2	= $AB^2C^2D^2E$
		= AC^2D	= ACE	= AD^2E^2	= AB^2CDE^2
A^2C^2	= AC	= BD^2	= BCE^2	= BC^2DE	= AD^2E
		= AB^2CD	= AB^2E	= $AB^2C^2D^2E^2$	= AC^2DE^2
A^2C	= AC^2	= BC^2D^2	= BE^2	= BCDE	= ACD^2E
		= AB^2D	= AB^2C^2E	= $AB^2CD^2E^2$	= ADE^2
A^2D^2	= AD	= BCD	= $BC^2D^2E^2$	= BE	= AC^2E
		= AB^2C^2	= AB^2CD^2E	= AB^2DE^2	= ACD^2E^2
A^2D	= AD^2	= BC	= BC^2DE^2	= BD^2E	= C^2DE
		= $AB^2C^2D^2$	= AB^2CDE	= AB^2E^2	= ACE^2
A^2E^2	= AE	= BCD^2E^2	= BC^2E	= BD	= $AC^2D^2E^2$
		= $AB^2C^2DE^2$	= AB^2C	= AB^2D^2E	= ACD
A^2E	= AE^2	= BCD^2E	= BC^2	= BDE^2	= AC^2D^2
		= AB^2C^2DE	= AB^2CE^2	= AB^2D^2	= ACDE

b. Clear main effects are: A, B

Clear two-factor interactions are: AB^2

c. Word length pattern is, W = (1,3,0)

Compare the 3^{5-2} design above (D = ABC and E = ABC^2) with the minimum aberration design of 3^{5-2} design with D = AB and E = AB^2C available on page 250. From page 216, the defining relations and the number of words length i in the subgroup are calculated. By using the information available on page 216, we obtained that the current design and the minimum aberration design have the same word length of W = (1,3,0). Because they are identical, we can conclude that the design above is another minimum aberration design besides 3^{5-2} design with D = AB and E = AB^2C.

PROBLEM-2

Clearly, B=CE, D=ACE and F=A2CE.

Thus, the defining contrast subgroup is

$I = BC^2E^2 = ACD^2E = AC^2E^2F = ABD^2 = AB^2C^2D^2E^2 = ABCEF = AB^2F = ADF^2 = CDEF = AB^2CDEF^2 = BCD^2EF^2 = BDF = ABC^2DE^2F^2$.

It can easily be seen that $AC = DE^2$, $AC^2 = EF^2$, $AE = CD^2$, $AE^2 = CF^2$ and thus none of them are clear.

The aliasing relations of main effects and two-factor interactions are:

$A = BD^2 = BF^2 = DF^2$
$B = AD^2 = AF = CE = DF$
$C = BE^2$
$D = AB = AF^2 = BF$
$E = BC^2$
$F = AB^2 = AD = BD$
$AC = DE^2$
$AC^2 = EF^2$
$AE = CD^2$
$AE^2 = CF^2$
$CD = EF$
$CE^2 = BC = BE$
$CF = DE$

Thus, in the ANOVA, besides A,B,C,D,E,F, AC,AC^2,AE,AE^2, one can choose one from (CD, EF), one from (CE^2,BC,BE) and one from (CF, DE)

METHOD-1: SIMPLE ANALYSIS

ANALYSIS OF VARIANCE TABLE

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
factor(A)	2	47.620	23.810	2.8297	0.0766119	.
factor(B)	2	11.822	5.911	0.7025	0.5041863	
factor(C)	2	39.586	19.793	2.3523	0.1143593	
factor(D)	2	188.037	94.019	11.1736	0.0002913	***
factor(E)	2	80.433	40.216	4.7795	0.0167104	*
factor(F)	2	9.933	4.966	0.5902	0.5611903	
factor(AC)	2	17.405	8.703	1.0343	0.3691520	
factor(AC2)	2	54.697	27.348	3.2502	0.0543543	.
factor(AE)	2	232.711	116.355	13.8282	7.333e-05	***
factor(AE2)	2	3.898	1.949	0.2316	0.7947942	
factor(CD)	2	61.454	30.727	3.6517	0.0394778	*
factor(BC)	2	4.238	2.119	0.2519	0.7791619	
factor(DE)	2	24.141	12.071	1.4345	0.2558124	
Residuals	27	227.187	8.414			

Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

The significant effects (at 5% level) are D,E, AE and CD. From the interaction plots, we select the best levels as D=2, E=0, A=0, C=2, D=2.

METHOD-2: LINER AND QUADRATIC CONTRASTS

FINAL OUTPUT OF STEPWISE REGRESSION WITH ALL LINEAR AND QUADRATIC EFFECTS

Call:
lm(formula = y ~ Al + Cl + Dl + El + Dq + Eq + AClq + ACql + ACqq + AElq + CDll + CFlq, data = design.sel)

Coefficients:

(Intercept)	Al	Cl	Dl	El	Dq
12.157	1.552	-1.424	2.082	-1.247	2.472
Eq	AClq	ACql	ACqq	AElq	CDll
1.707	2.045	-2.213	1.361	-3.507	5.023
CFlq					
2.391					

The above model is too large; and we should see whether all terms are significant by displaying the t-statistics.

```
> g=lm(y~Al+Cl+Dl+El+Dq+Eq+AClq+ACql+ACqq+AElq+CDll+CFlq)
> summary(g)

Call:
lm(formula = y ~ Al + Cl + Dl + El + Dq + Eq + AClq + ACql +
    ACqq + AElq + CDll + CFlq)

Residuals:
     Min       1Q   Median       3Q      Max
-4.31906 -1.66155  0.06925  1.86304  5.44715

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  12.1574     0.3610  33.676  < 2e-16 ***
Al            1.5517     0.6253   2.482 0.017271 *
Cl           -1.4240     0.6253  -2.277 0.028048 *
Dl            2.0820     0.6253   3.330 0.001846 **
El           -1.2473     0.6253  -1.995 0.052756 .
Dq            2.4722     0.6253   3.954 0.000297 ***
Eq            1.7067     0.6253   2.729 0.009304 **
AClq          2.0448     1.0830   1.888 0.066117 .
ACql         -2.2132     1.0830  -2.043 0.047467 *
ACqq          1.3611     1.0830   1.257 0.215958
AElq         -3.5069     1.4441  -2.428 0.019635 *
CDll          5.0231     1.2506   4.017 0.000246 ***
CFlq          2.3909     1.3017   1.837 0.073491 .
---
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

Residual standard error: 2.653 on 41 degrees of freedom
Multiple R-Squared: 0.7124,     Adjusted R-squared: 0.6282
F-statistic: 8.461 on 12 and 41 DF,  p-value: 9.639e-08
```

ITERATION 1

Thus we delete the terms that are not significant at 5% level (El, Aclq, Acqq, CFlq) and re-run the regression.

```
> g1=lm(y~Al+Cl+Dl+Dq+Eq+ACql+AElq+CDll)

> summary(g1)

Call:
lm(formula = y ~ Al + Cl + Dl + Dq + Eq + ACql + AElq + CDll)

Residuals:
     Min       1Q   Median       3Q      Max
-5.14441 -1.77221 -0.02101  1.65487  5.91435
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  12.1574     0.3936  30.886  < 2e-16 ***
Al            1.5517     0.6818   2.276  0.02766 *
Cl           -1.4240     0.6818  -2.089  0.04242 *
Dl            2.0820     0.6818   3.054  0.00379 **
Dq            2.4722     0.6818   3.626  0.00073 ***
Eq            1.7067     0.6818   2.503  0.01600 *
ACql         -2.2132     1.1809  -1.874  0.06741 .
AElq         -2.0355     1.3101  -1.554  0.12725
CDll          5.6603     1.3101   4.321 8.48e-05 ***
---
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

Residual standard error: 2.893 on 45 degrees of freedom
Multiple R-Squared: 0.6247,     Adjusted R-squared: 0.558
F-statistic: 9.362 on 8 and 45 DF,  p-value: 1.711e-07
```

ITERATION 2

We again delete the terms that are not significant at 5% level (ACql, Aelq) and re-run the regression.

```
> g2=lm(y~Al+Cl+Dl+Dq+Eq+CDll)
> summary(g2)

Call:
lm(formula = y ~ Al + Cl + Dl + Dq + Eq + CDll)

Residuals:
   Min     1Q Median     3Q    Max
-5.981 -1.395 -0.228  2.131  5.914

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  12.1574     0.4097  29.671  < 2e-16 ***
Al            1.5517     0.7097   2.186  0.03379 *
Cl           -1.4240     0.7097  -2.007  0.05057 .
Dl            2.0820     0.7097   2.934  0.00516 **
Dq            2.4722     0.7097   3.483  0.00108 **
Eq            1.7067     0.7097   2.405  0.02017 *
CDll          6.5417     1.2292   5.322 2.81e-06 ***
---
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

Residual standard error: 3.011 on 47 degrees of freedom
Multiple R-Squared: 0.5753,     Adjusted R-squared: 0.521
F-statistic: 10.61 on 6 and 47 DF,  p-value: 2.003e-07
```

ITERATION 3

We again delete the term that are not significant at 5% level (Cl) and re-run the regression.

```
> g3=lm(y~Al+Dl+Dq+Eq+CDll)
> summary(g3)

Call:
lm(formula = y ~ Al + Dl + Dq + Eq + CDll)

Residuals:
     Min       1Q   Median       3Q      Max
-5.73843 -2.01620 -0.09954  2.36921  6.83796

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept)  12.1574     0.4225  28.778  < 2e-16 ***
Al            1.5517     0.7317   2.121  0.03915 *
Dl            2.0820     0.7317   2.845  0.00650 **
Dq            2.4722     0.7317   3.379  0.00145 **
Eq            1.7067     0.7317   2.332  0.02392 *
CDll          6.5417     1.2674   5.162 4.64e-06 ***
---
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

Residual standard error: 3.104 on 48 degrees of freedom
Multiple R-Squared: 0.5389,     Adjusted R-squared: 0.4908
F-statistic: 11.22 on 5 and 48 DF,  p-value: 3.395e-07
```

Thus, the final model is

Y = 1.5517 Al + 2.0820 Dl + 2.4722 Dq + 1.7067 Eq + 6.5417 CDll

To maximize y, we choose Al = +1, Dl = +1, Dq = +1, Eq = +1 and CDll=+1 (which means either Cl=Dl=-1 or Cl=Dl=1).

Finally, we choose the optimum combination as follows:

i. Al=+1 means A=2.
ii. Dl=+1 and Dq=+1 means D=2.
iii. Eq=+1 means E=0 or E=2; looking at the main effects plot we choose E=0.
iv. Since Dl=1, we choose Cl=1 i.e., C=2.

Disagreement between results obtained from method-1 and method-2:

The only difference between the settings obtained from method-1 and method-2 is the setting of factor A, which is 0 in method 1 and 2 in method 2. This apparent anomaly can be explained very easily. Note that in the simple analysis, the setting of A was obtained on the basis of the AxE interaction plot. A careful observation of this plot tells ua that choosing A=0 or A=2 makes very little difference. However, if we have the main effect of A in the model, one would go for A=2. Thus, considering both the main effect of A and interaction AE, A=2 may be a slightly better choice.